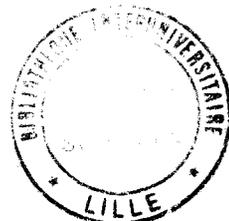


50376  
1994  
337

50376  
1994  
337

UNIVERSITE DES SCIENCES ET TECHNOLOGIES  
DE LILLE



THESE

présentée par **Jean-Luc BRUYELLE**

pour obtenir le titre de **DOCTEUR**

(Spécialité: **PRODUCTIQUE: AUTOMATIQUE ET INFORMATIQUE INDUSTRIELLE**)

**CONCEPTION ET REALISATION D'UN  
DISPOSITIF DE PRISE DE VUE  
STEREOSCOPIQUE LINEAIRE**

-

**APPLICATION A LA DETECTION  
D'OBSTACLES A L'AVANT DES  
VEHICULES ROUTIERS**

Soutenu le 2 décembre 1994

Composition du jury:

Pierre VIDAL	Président
Bernard DUBUISSON	Rapporteurs
Claude LAURGEAU	
Bertin DE BETTIGNIES	Examineurs
Jean-Pierre DEPARIS	
Marie DE SAINT BLANCARD	
Jack-Gérard POSTAIRE	Directeur de thèse



50376  
1994  
337

50376  
1994  
337



# TABLE DES MATIERES



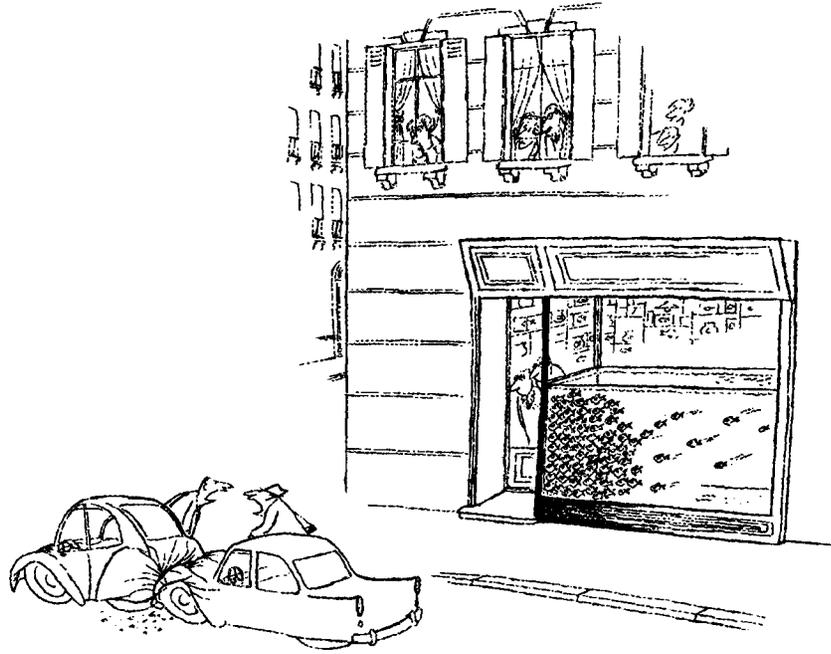
# Table des matières

<b>1. INTRODUCTION .....</b>	<b>1</b>
<b>1.1. Nécessité de la détection d'obstacles sur route.....</b>	<b>3</b>
1.1.1. Les accidents de la route.....	4
Importance des accidents de la route.....	4
Causes des accidents de la route .....	4
1.1.2. Amélioration de la sécurité routière.....	7
Actions sur le réseau routier .....	7
Action sur les conducteurs.....	7
Action sur les véhicules.....	8
Nouvelle approche de l'amélioration de la sécurité.....	9
<b>1.2. Le programme PROMETHEUS.....</b>	<b>10</b>
<b>1.3. Méthodes existantes pour la détection d'obstacles sur route .....</b>	<b>14</b>
1.3.1. Méthodes actives de détection d'obstacles .....	14
1.3.2. Méthodes passives de détection d'obstacle.....	16
<b>1.4. Considérations sur la stéréovision.....</b>	<b>18</b>
1.4.1. Prise de vue.....	19
1.4.3. Extraction des primitives.....	19
1.4.4. Mise en correspondance (ou appariement).....	21
1.4.5. Triangulation géométrique .....	23
Calibrage, calibration et étalonnage.....	23
1.4.6. Reconstruction de la scène 3D .....	24
<b>1.5. Organisation du mémoire.....</b>	<b>25</b>
<b>2. DISPOSITIF DE DETECTION AUTOMATIQUE D'OBSTACLES SUR LA CHAUSSEE.....</b>	<b>27</b>
<b>2.1. Les capteurs linéaires .....</b>	<b>29</b>
2.1.1. Structure des capteurs linéaires.....	29
2.1.2. Caractéristiques des capteurs linéaires.....	32
<b>2.2. Les caméras linéaires.....</b>	<b>33</b>
2.2.1. Constitution d'une caméra linéaire.....	33
2.2.2. Champ de visée d'une caméra linéaire .....	34
2.2.3. Exemples d'applications des caméras linéaires.....	35
Reconstitution de la dimension verticale par déplacement mécanique.....	35
Applications dans lesquelles la verticale est inutile.....	37
<b>2.3. Principe du stéréoscope linéaire embarqué .....</b>	<b>39</b>
2.3.1. Hypothèses .....	39
Hypothèses sur les caméras.....	39
Hypothèses sur la scène .....	41
2.3.2. Installation des caméras dans la scène .....	42
2.3.3. Reconstitution de la profondeur et de la position.....	44
Notion de disparité .....	45
Calcul de la profondeur à partir de la disparité .....	45
Calcul de la position à partir de la profondeur .....	45
2.3.4. Passage du repère lié au stéréoscope au repère lié à la scène .....	46
Position et orientation des caméras dans la scène.....	46
Passage au repère de travail.....	48
Détermination de la hauteur des obstacles .....	49
<b>2.4. Réalisation électronique du stéréoscope linéaire embarqué.....</b>	<b>50</b>
2.4.1. Caméras .....	51
2.4.2. Acquisition.....	51
2.4.3. Traitements informatiques.....	52

<b>2.5. Réalisation mécanique du stéréoscope linéaire embarqué.....</b>	<b>53</b>
2.5.1. Contraintes liées au calibrage du stéréoscope.....	53
2.5.2. Contraintes liées à l'environnement.....	54
Contraintes liées à la stéréovision linéaire.....	54
Contraintes de stabilité.....	54
Contraintes de facilité de réglage.....	55
2.5.3. Monture orientable d'une caméra.....	56
Platine inférieure et montage sur le support commun.....	57
Ensemble de vis tirant-poussant.....	57
Platine supérieure et réglage des angles de site et de roulis.....	58
Platine azimutale et support de caméra.....	60
2.5.4. Monture commune aux deux caméras.....	60
<b>2.6. Conclusion.....</b>	<b>61</b>
<b>3. CALIBRAGE DU STEREOSCOPE LINEAIRE.....</b>	<b>63</b>
<b>3.1. Problème et contraintes.....</b>	<b>65</b>
3.1.1. Problème à résoudre par le calibrage.....	65
Cas général - Solutions existantes en vision matricielle.....	65
Cas de la stéréovision linéaire.....	67
3.1.2. Contraintes.....	68
Contraintes liées à l'emploi de caméras linéaires.....	68
Contraintes liées à la calibration des images.....	68
Contraintes liées à l'environnement automobile.....	69
<b>3.2. Installation pour le calibrage.....</b>	<b>69</b>
<b>3.3. Mire de calibrage.....</b>	<b>72</b>
3.3.1. Constitution de la mire.....	72
Repères de calibrage vertical.....	72
Repères de calibrage horizontal.....	73
Lignes-guides.....	73
3.3.2. Dimensions de la mire.....	74
<b>3.4. Procédure de calibrage.....</b>	<b>75</b>
3.4.1. Procédure pratique de calibrage.....	75
Phase 1: mise en place.....	75
Phase 2: réglage grossier de l'angle de site.....	77
Phase 3: réglage fin des angles de site et de roulis.....	78
Phase 4: ajustement final du calibrage vertical.....	79
Phase 5: blocage du calibrage vertical.....	80
Phase 6: réglage de l'angle d'azimut.....	80
3.4.2. Justification théorique de la procédure.....	81
<b>3.5. Résultats expérimentaux.....</b>	<b>82</b>
3.5.1. Premiers essais de calibrage.....	82
Réalisation de la mire.....	82
Réalisation du support des caméras.....	83
Méthodes de mesure des positions et des orientations.....	84
Méthode d'étalonnage du stéréoscope après calibrage.....	84
3.5.2. Premiers essais du stéréoscope: précision de localisation.....	87
<b>3.6. Remarques pratiques.....</b>	<b>89</b>
<b>3.7. Conclusion.....</b>	<b>90</b>
<b>4. TRAITEMENT DES IMAGES STEREOSCOPIQUES LINEAIRES..</b>	<b>91</b>
<b>4.1. Introduction.....</b>	<b>93</b>
<b>4.2. Extraction des primitives.....</b>	<b>93</b>
4.2.1. Nature des primitives.....	93
4.2.2. Méthode d'extraction des primitives.....	94
Calcul du gradient de l'image.....	97

Extraction des points de contour .....	98
Attributs des points de contour .....	99
Algorithme de détection des points de contour .....	100
4.2.3. Représentation des primitives en mémoire .....	102
<b>4.3. Mise en correspondance .....</b>	<b>103</b>
4.3.1. Critère de mise en correspondance .....	103
4.3.2. Contraintes de mise en correspondance .....	104
4.3.3. Appariements multiples .....	107
4.3.4. Primitives non appariées .....	108
4.3.5. Représentation en mémoire .....	109
<b>4.4. Triangulation géométrique .....</b>	<b>110</b>
<b>4.5. Conclusion .....</b>	<b>112</b>
<b>5. INTEGRATION SUR UN VEHICULE, ESSAIS SUR ROUTE .....</b>	<b>113</b>
<b>5.1. Introduction .....</b>	<b>115</b>
<b>5.2. Enregistrement des séquences acquises par le stéréoscopique linéaire .....</b>	<b>115</b>
5.2.1. Caractéristiques du disque dur .....	116
5.2.2. Contrôle de prise de vue sur le moniteur de l'ordinateur hôte .....	117
<b>5.3. Enregistrement synchronisé de séquences d'images linéaires et vidéo .....</b>	<b>117</b>
5.3.1. Nécessité d'un enregistrement vidéo synchronisé avec les images linéaires ..	117
5.3.2. Le code temporel de commande .....	118
Constitution du code temporel de commande .....	120
Transmission et enregistrement du code temporel de commande .....	120
Enregistrement synchronisé par le code temporel de commande .....	120
5.3.3. Utilisation du code temporel de commande pour l'enregistrement synchronisé de séquences linéaires et vidéo .....	122
<b>5.4. Exemples de séquences enregistrées .....</b>	<b>123</b>
5.4.1. Conditions expérimentales des prises de vue .....	123
5.4.2. Lecture des séquences stéréoscopiques linéaires .....	124
5.4.3. Détection et localisation d'un piéton .....	126
5.4.4. Route bien dégagée - aucun obstacle .....	131
5.4.5. Influence du tangage sur la détection .....	133
5.4.6. Approche et dépassement d'un véhicule .....	137
5.4.7. Dépassement par un véhicule .....	140
5.4.8. Croisement d'un vélomoteur sur une route de campagne .....	143
<b>5.5. Conclusion .....</b>	<b>146</b>
<b>6. CONCLUSION .....</b>	<b>147</b>
<b>BIBLIOGRAPHIE .....</b>	<b>153</b>
<b>A. HYPERFOCALE ET PROFONDEUR DE CHAMP .....</b>	<b>163</b>
A.1. Choix pratique du cercle de confusion .....	168
<b>B. AMELIORATIONS POSSIBLES DE LA METHODE D'ETALONNAGE DU STEREOSCOPE LINEAIRE .....</b>	<b>169</b>
B.1. Mieux garantir la coplanéité des champs de visée .....	171
B.2. Déterminer la distance focale des objectifs .....	173
B.3. Déterminer la distorsion des objectifs .....	174
<b>C. MESURE DES HAUTEURS A L'AIDE DU STEREOSCOPE LINEAIRE .....</b>	<b>177</b>
C.1. Mesure directe par le stéréoscope .....	179

**C.2. Mesure à l'aide d'une caméra matricielle calibrée .....180**



# Chapitre 1

## INTRODUCTION

## 1.1. Nécessité de la détection d'obstacles sur route

Dans l'histoire, peu d'inventions ont eu autant d'importance que l'automobile, qui a profondément modifié notre mode de vie, voire de pensée, tant les notions de distance et de temps en ont été affectées dans l'esprit de l'humanité.

Son apparition a constitué un facteur primordial dans le démarrage de l'industrie moderne, d'abord, après le chemin de fer, en tant que moyen de transport des marchandises, qui a permis d'ouvrir des marchés nouveaux aux marchandises à mesure que le réseau routier s'étendait, atteignant des zones auparavant difficilement accessibles.

Le réseau routier a ainsi également constitué un élément important du développement de régions entières, voire de pays entiers à l'exemple des Etats-Unis, dont la structure urbaine (l'activité économique est concentrée dans les centres-villes, les logements étant situés en banlieue, parfois éloignée de plusieurs dizaines de kilomètres du lieu de travail) a été dès l'origine façonnée par les facilités de transport offertes par la possession d'une voiture particulière: [Rostow 60] résume cette évolution, qui prit forme dès les années 1920, en écrivant que "*Les Etats-Unis se sont mis sur roues. Ce fut vraiment la période de l'automobile pour tous. Grâce à elle, les familles américaines ont commencé une grande migration intérieure pour s'installer dans des villas neuves, situées dans les banlieues; et ces maisons neuves s'emplitrent de plus en plus de postes de radio, de réfrigérateurs et d'autres appareils ménagers [...]*", ces appareils étant eux-mêmes transportés par la route, au moins sur une partie du trajet voire, de plus en plus souvent, sur son intégralité (59,5% des transports de marchandises en France en 1990, exprimés en t.km, étaient effectués par route [Fremy 93]).

Ce mode de vie basé sur la mobilité, maintenant répandu dans tous les pays industrialisés, fait de l'automobile l'un des éléments essentiels de la société moderne.

D'autre part, la voiture représente elle-même une part importante de la consommation: 16% du budget des ménages français lui étaient consacrés en 1990 [Fremy 1993], cette somme comprenant, outre l'achat d'une voiture, son entretien, son assurance, l'achat de son carburant, etc. Ceci confère à l'industrie automobile un rôle de moteur vis-à-vis de l'ensemble des autres industries: la fabrication d'une voiture nécessite des matériaux (métaux ferreux et non ferreux, matières plastiques, tissus pour l'aménagement intérieur, etc.), et des savoir-faire très variés (mécanique, électricité, maintenant électronique et informatique), qui justifient l'apparition d'industries de sous-traitance qui font parfois vivre des régions entières [Sauvy 84], et en constituent parfois l'employeur quasi-unique, comme l'est Michelin à Clermont-Ferrand ou comme l'était General Motors à Flint, Michigan, avant la fermeture de ses usines [Moore 89].

Mais l'importance même de l'automobile génère de nombreux inconvénients liés à la densité du trafic routier. Les aller-retours quotidiens sur les voies d'accès aux grandes villes provoquent la saturation de celles-ci, génératrice d'embouteillages dans lesquels les Français passent 80 millions d'heures par an [Saguez 94]. Le stationnement d'un nombre toujours croissant de véhicules impose la création d'aménagements de plus en plus envahissants — et coûteux — au

coeur même des villes, l'automobiliste moyen n'acceptant pas de parcourir à pied plus de quelques centaines de mètres. Les dommages à l'environnement, longtemps négligés, sont également de plus en plus mal acceptés, notamment depuis que les effets nocifs du bruit et de la pollution de l'air ont été mis en évidence par les chercheurs [Mautner 94].

Le plus grave de ces inconvénients, cependant, reste la mortalité importante sur la route.

### 1.1.1. Les accidents de la route

#### *Importance des accidents de la route*

En 1990, 10 289 Français sont morts dans des accidents routiers, soit 2% de la mortalité globale de cette année. Le risque d'accident en automobile est ainsi, et de loin, le plus élevé de tous les moyens de transports: selon l'INRETS<sup>1</sup>, le risque d'accident mortel par million de voyages est de 185 en voiture, contre 48 en train, 2,45 en avion, et 0,34 en métro [Fremy 93]. On considéra longtemps ce risque comme une rançon inévitable des performances et du nombre croissants des véhicules; ce n'est qu'au début des années 70 que le public et les autorités prirent conscience du drame que représentait une telle hécatombe — en 1972, l'année la plus meurtrière de l'histoire de l'automobile, on compta 16 617 tués et 388 067 blessés — et qu'une véritable politique de sécurité routière fut mise en place. Les mesures qui furent prises depuis cette époque par les autorités politiques, les améliorations qui furent apportées aux véhicules par les constructeurs, sont le fruit d'un travail de réflexion commune mené à partir de l'analyse des accidents [Ferrandez 86, Malaterre 86, Ramet 86] menées dans le but d'en déterminer les causes, et autant que possible de trouver les moyens de les éviter.

#### *Causes des accidents de la route*

Il est souvent difficile d'attribuer une cause unique à un accident, qui est le plus souvent la conséquence d'une succession de circonstances dont aucune, si elle était arrivée seule, n'aurait eu de conséquences fâcheuses [Gentile 94]. De plus, les données recueillies après coup par la police ou par les constats d'accidents sont souvent insuffisantes pour tirer des conclusions sur les causes de l'accident rapporté. C'est pourquoi l'INRETS a établi un fichier plus détaillé comportant, pour un nombre réduit d'accidents, relève les circonstances et tous les facteurs qui sont intervenus, recueillis par une équipe d'enquête interdisciplinaire [Ramet 86]. C'est le plus souvent ce fichier qui sert de base aux études concernant les causes d'accidents. Le programme REAGIR<sup>2</sup> d'étude des causes d'accidents, notamment, a permis de recenser les causes les plus fréquentes d'accidents, en définissant leur fréquence d'apparition, et en tenant compte des causes multiples — par exemple, un conducteur alcoolisé qui roule trop vite cumule deux causes

---

<sup>1</sup> Institut National de Recherche sur les Transports et leur Sécurité, 2, avenue du Maréchal Malleret-Joinville, 94114 Arcueil Cedex.

<sup>2</sup> Réagir par des Enquêtes sur les Accidents Graves et les Initiatives pour y Remédier: programme lancé en 1982 par le Comité Interministériel de la Sécurité Routière.

d'accident. Les conclusions de ces enquêtes, citées dans [Gentile 94], sont les suivantes:

*"[...] Dans presque tous les accidents un facteur au moins est lié à l'utilisateur (94%), que dans un accident sur deux un facteur est lié à l'infrastructure et que dans un accident sur trois un facteur au moins est lié au véhicule."*

*"Les facteurs liés à l'utilisateur concernent l'aptitude de celui-ci (alcool, fatigue, inattention, inaptitude à la conduite, problèmes physiques ou psychologiques...) dans 67% des accidents et son comportement au volant (vitesse, manœuvre dangereuse, infraction caractérisée, mauvaise utilisation du véhicule...) dans 80% des accidents. De tous ces facteurs, ce la vitesse qui arrive en tête (48% des accidents, et même 62% pour les motos), suivie de l'imprégnation alcoolique (30% des accidents) avec un taux d'alcoolémie moyen de 1,8g, soit plus du double du maximum légal."*

*"Dans un tiers des accidents examinés, la conception de l'infrastructure a joué un rôle (tracé, abords...); dans un quart des accidents, on trouve comme facteurs l'entretien et l'exploitation (signalisation, éclairage, état des chaussées)."*

*"Enfin, pour ce qui concerne les véhicules, leur conception intervient pour 14% des accidents (résistance aux chocs, problèmes de visibilité, conception générale) et leur entretien pour 21% (pneus, éclairage, freins, direction...)."*

Analysons plus particulièrement les facteurs liés au conducteur qui sont, de loin, prépondérants dans les causes des accidents analysés par REAGIR.

- **La vitesse** est, par ordre d'importance, le premier facteur évoqué comme cause d'accident: elle est impliquée dans 48% des cas étudiés. La réponse à cet état de fait est, en principe, simple: la limitation de vitesse, imposée en France depuis la fin des années 60. L'efficacité de cette mesure, contestée par certains irréductibles, est cependant confirmée par les statistiques, puisque le déclin du nombre d'accidents, et de leurs victimes, coïncida avec l'imposition de la limitation de vitesse, et dépassa même les espérances des autorités [Gentile 93]. *A contrario*, lorsque 40 états américains augmentèrent la vitesse autorisée de 55 à 65 mph en 1987, on constata une augmentation importante (18%) de la mortalité sur les routes, malgré une différence de vitesse assez modeste: 16 km/h seulement. Le respect effectif de la limitation de vitesse reste cependant un problème difficile à résoudre, en raison des résistances conscientes ou inconscientes des conducteurs [Chich 87].

Le non respect de la limitation de vitesse n'est pas la seule cause d'accident liée à la vitesse. Tout d'abord, on constate qu'un nombre important de conducteurs apprécie mal le danger lié à l'environnement (brouillard, pluie, chaussée glissante, visibilité réduite, virage serré) et abordent les difficultés du parcours avec une vitesse excessive, bien qu'inférieure à la limite autorisée. Le non respect des distances de sécurité entre véhicules, notamment en cas de trafic dense, est également une cause commune d'accidents: c'est la cause presque unique des carambolages, parfois très meurtriers (53 tués près de Beaune le 31 juillet 1982). La vitesse est alors un facteur indirect, ne permettant pas de réagir lorsque le véhicule précédent freine.

- **L'alcoolémie** est le deuxième facteur d'accident par ordre d'importance. 38% des responsables présumés d'accidents mortels ont une alcoolémie supérieure à 0,80 g/l, et ils sont responsables de 41% des tués. De même, 51% des au-

tomobilistes tués dans un accident où ils sont seuls impliqués ont une alcoémie supérieure à 0,80 g/l. Le risque d'être impliqué dans un accident mortel est ainsi multiplié par 7 pour une alcoémie comprise entre 0,80 g/l et 1,19 g/l, et par 165 au-delà de 2 g/l ([Got 89] cité dans [Gentile 94]). Les raisons de l'influence de l'alcool sur le risque d'accident sont essentiellement la baisse de la vigilance et le ralentissement des réflexes. Mais d'autres effets jouent également un rôle, tels que le rétrécissement du champ visuel.

L'alcool n'est pas le seul facteur qui entraîne de tels effets: l'usage de certains médicaments a ainsi été considéré par certains auteurs, tels que C. Got: des substances telles que les somnifères, les tranquillisants, les antidépresseurs, mais aussi les médicaments destinés à lutter contre le mal des transports, ont des effets semblables, bien que leur incidence réelle soit encore mal connue [Gentile 94]. Notons également les défaillances physiques liées à la prise de médicaments, qui interviennent dans 3% des accidents.

- Un autre facteur très important est *la fatigue*, citée dans 18% des enquêtes REAGIR. Ce facteur se rencontre essentiellement la nuit, mais une étude récente, menée par le Laboratoire du Sommeil lors de départs en vacances [Philip 93], montrait que 50% des automobilistes présentaient un déficit de sommeil de 3h30 en moyenne, et que 10% avaient passé une nuit blanche avant de partir. 11% des conducteurs se plaignaient de somnolence sévère pendant leur voyage. La fatigue présente le même type de risques que l'alcool précédemment cité: baisse de la vigilance et allongement du temps de réaction, et est très difficile à prévenir (il est impossible de quantifier la fatigue, à la différence de l'alcool).
- *L'inattention* vient presque à égalité avec la fatigue. Elle est généralement causée par la manipulation d'accessoires situés à l'intérieur de l'habitacle, tels que l'autoradio, et surtout le radiotéléphone, qui mobilisent une fraction importante des capacités de concentration du conducteur, exigent des manipulations et un contrôle visuel, au détriment de la conduite [Petica 93].

On constate ainsi que les problèmes liés au manque d'attention et de réflexes du conducteur sont une cause d'accident extrêmement fréquente, qui mérite une attention particulière lorsque l'on veut en diminuer le nombre et la gravité.

- *L'inaptitude à la conduite* est citée dans 16% de accidents. Elle est généralement le fait de conducteurs débutants et/ou qui conduisent un véhicule avec lequel ils sont peu familiers, neuf ou de location (6% des cas). Le problème est ici un manque de formation, et la solution passe par une éducation des conducteurs à la prudence.
- Enfin, *les défaillances physiques* apparaissent dans 12% des cas. Outre les malaises (4%) et les effets secondaires de médicaments (3%), il s'agit essentiellement de déficiences visuelles et/ou auditives permanentes (4%), qui entrent, comme l'inattention, la fatigue et l'alcool, dans les causes liées à un défaut de perception de l'environnement.

## 1.1.2. Amélioration de la sécurité routière

Considérant les causes d'accidents déterminées par des enquêtes telles que REAGIR, on peut envisager de nombreux moyens pour y remédier. Ces moyens mettent en oeuvre des moyens de diverses nature (législatives, techniques, etc.)

### **Actions sur le réseau routier**

L'infrastructure routière n'est jamais la cause unique d'un accident. Elle joue cependant un rôle dans un tiers des accidents analysés par les enquêtes REAGIR, et mérite donc une attention particulière.

- La mise au point de revêtements spéciaux, destinés à améliorer l'adhérence des pneumatiques, particulièrement en cas de pluie, permet de réduire l'influence des conditions météorologiques.
- La séparation des automobiles et des autres moyens de transports (deux-roues, machines agricoles et de chantier, piétons). La construction d'autoroutes et de boulevards périphériques réservés aux automobiles a permis d'optimiser leurs caractéristiques (absence de croisements, douceur des courbes, élimination des pentes importantes) pour la conduite à grande vitesse afin d'en limiter le danger. Les autoroutes offrent ainsi une sécurité très supérieure au reste du réseau routier (5 500 accidents par an sur un total de 150 000), malgré une vitesse autorisée très supérieure. Cependant, la vitesse élevée, liée à un sentiment trompeur de sécurité, rend les accidents sur autoroute plus graves: le nombre de décès par accident y est moitié plus élevée que la moyenne.
- L'examen des rapports d'accidents met en évidence un nombre réduits de points où se produisent un nombre anormalement élevé d'accidents. On parle alors de *points noirs*. Il s'agit généralement d'endroits où les conditions de circulation changent brutalement sur un trajet, surprenant les conducteurs imprudents ou peu habitués. L'amélioration de la sécurité à ces endroits passe par l'aménagement des infrastructures afin d'améliorer les conditions de circulation. Lorsqu'un tel aménagement n'est pas possible (par exemple, dans le cas d'un virage très serré en zone urbaine), une diminution du nombre d'accidents est généralement obtenue par l'amélioration de la signalisation et par la limitation de la vitesse aux abords de la zone dangereuse.

Notons toutefois que l'amélioration de la circulation présente un effet pervers: les automobilistes ont tendance à adapter leur comportement à la perception du risque. Ainsi, lorsque la circulation est facilitée par un aménagement nouveau, la circulation augmente. Si la sécurité est renforcée, un sentiment de sécurité trompeur pousse les conducteurs à une prise de risque accrue. On constate ainsi une notion de *conduite à risque constant* [Wilde 90], l'utilisateur acceptant un certain niveau de risque, et se tenant à ce niveau quelque soient les conditions.

### **Action sur les conducteurs**

On l'a vu, les causes d'accidents les plus fréquentes sont liées au comportement (80%) et aux limitations (67% — notamment de la capacité d'attention)

des conducteurs. Les efforts destinés à améliorer la sécurité routière portent donc en grande partie sur ce dernier.

- **L'éducation et la formation** des conducteurs a commencé dès 1893, avec l'instauration du permis de conduire par le préfet Lépine. Le fait de subordonner la conduite à un examen permet d'assurer un minimum de connaissances du code de la route et d'aptitude pratique à l'ensemble des conducteurs. L'organisation du permis de conduire, qui a beaucoup évolué au cours du temps, comporte toujours des lacunes, dont la principale est son côté "bachotage": immédiatement après l'obtention du permis, 14% des conducteurs interprètent mal des panneaux courants. L'expérience ne semble pas combler ces lacunes: les conducteurs anciens font encore moins bien, avec 20 à 25% d'erreur: soit qu'ils ne reconnaissent pas les panneaux récents, soit qu'ils ont oublié ceux qui existaient déjà lorsqu'ils ont débuté ([Auto-Moto 84] cité dans [Gentile 94]).

Le permis à points, mis en place en 1992, est précisément destiné, au moins en partie, à remédier à ce problème par des stages de recyclage destinés aux conducteurs désireux de regagner des points perdus à la suite d'infractions. On espère obtenir ainsi une formation permanente des conducteurs les plus dangereux.

Les campagnes d'information par voie de presse sont également un moyen d'éducation important. Leur efficacité réelle n'a cependant jamais pu être démontrée. Il semble en revanche que la médiatisation d'accidents réels, tels que celui de Beaune en 1982, aient un effet sensible et assez prolongé [Gentile 94].

- **La répression** est également un moyen important pour modifier les comportements à risque, tels que les excès de vitesse ou, plus généralement, le non respect du code. Plus que la répression des infraction elle-même, c'est la peur du gendarme qui est recherchée pour dissuader les contrevenants en puissance. Ceci suppose la réunion de deux conditions: les contrôles doivent être suffisamment fréquents pour que le risque d'être pris sur le fait soit non négligeables, et les sanctions doivent être assez sévères pour être redoutées. En France, ces deux objectifs semblent difficiles à atteindre en raison, d'une part, de l'insuffisance quantitative et qualitative des moyens de police affectés à ces tâches [Documentation 89], d'autre part d'un manque chronique d'application des peines prévues par la législation. De plus, la lenteur des procédures judiciaires éloigne l'infraction de sanction, ce qui en diminue l'effet dissuasif [Gentile 94].

Aucune donnée chiffrée ne permet d'estimer l'effet de la répression des infractions sur le nombre et la gravité des accidents.

### **Action sur les véhicules**

Les véhicules sont peu représentés dans les causes d'accidents, qu'il s'agisse de leur conception (14%) ou de leur entretien (21%). L'action concernant les véhicules, concernant l'entretien, porte essentiellement sur l'éducation des conducteurs qui doivent être conscients du danger présenté par un véhicule en mauvais état. Quand à leur conception, elle peut intervenir dans les accidents par leur

évitement (sécurité active) ou par la limitation des dégâts causés sur leurs passagers et sur les autres usagers éventuellement impliqués (sécurité passive).

- **La sécurité active** concerne la capacité d'un véhicule à éviter un accident par son comportement sur la route. Le freinage, la direction interviennent dans la sécurité active, et leur efficacité dans ce cadre peut être améliorée par des dispositifs d'assistance, tels que l'ABS (Anti Blocking System, développé par Bosch).

Après avoir suscité beaucoup d'espoirs, ces dispositifs se sont avérés décevants, voire dangereux, et ceci pour deux raisons. Tout d'abord, le conducteur a naturellement tendance à modifier son comportement en fonction du supplément de sécurité attendu, dans le sens d'une plus grande prise de risque: "Je ne risque rien à rouler trop vite, j'ai un ABS". Les statistiques montrent ainsi que l'ABS a provoqué une augmentation du nombre d'accidents, et surtout de leur gravité [Gentile 94, p.30]. D'autres systèmes destinés à améliorer la sécurité active, notamment l'antipatinage à l'accélération (ABR) ont connu le même destin. D'autre part, et pour les mêmes raisons, une panne du système d'assistance peut avoir des conséquences catastrophiques: l'auteur de ces lignes a eu personnellement connaissance de deux accidents causés par une panne d'ABS, entraînant dans les deux cas une sortie de route à grande vitesse.

- **La sécurité passive** concerne la capacité d'un véhicule, lors d'un accident, à protéger ses passagers et les autres usagers éventuellement impliqués. Parmi les dispositifs de sécurité passive, on peut citer notamment la ceinture de sécurité et l'airbag. La conception générale du véhicule joue également un grand rôle dans la sécurité passive: les pare-brise en verre feuilleté, les barres de renfort latérales situées dans les portières, les systèmes évitant l'enfoncement de la colonne de direction, les châssis dont la déformation protège l'habitacle en cas de choc frontal, les dispositifs anti-incendies, sont quelques exemples des efforts développés dans ce sens par les constructeurs.

La sécurité passive a, elle aussi, ses effets pervers. Une étude de l'INRETS a montré que les mesures prises pour accroître la sécurité des passagers d'un véhicule ont souvent pour effet d'augmenter son poids, sa rigidité, et par conséquent son agressivité, c'est à dire les dégâts infligés aux autres usagers lors d'un accident.

### ***Nouvelle approche de l'amélioration de la sécurité***

Depuis les années 80, une nouvelle approche de la sécurité routière est apparue, faisant intervenir les technologies récemment développées — électronique, informatique, communications — pour prévenir le conducteur de tout risque d'accident, et pour aider à son évitement.

Une telle approche utilise le concept de *voiture intelligente* [Parey 87], dotée de capteurs qui lui permettent de percevoir son environnement, de tester son propre état (aide à la maintenance) et celui de son conducteur (éthylisme, somnolence, etc.). Une telle voiture serait également dotée de moyens de communication sophistiqués qui la renseignent sur sa position, les conditions de circulation sur l'itinéraire prévu, et autorisent la coopération avec les autres véhicules lors des manoeuvres de dépassement, ou pour faciliter la conduite en

convoi, afin notamment d'éviter les carambolages. Une interface utilisateur sophistiquée, prenant en compte les recherches en cours ou à venir dans le domaine de l'ergonomie, lui permettrait de communiquer à son conducteur les informations dont il a besoin, tout en sollicitant le moins possible son attention par des informations inutiles, ce qui lui permettrait de se concentrer sur la conduite. Des cartes routières, numérisées sur un support à grande capacité, permettraient de définir l'itinéraire le mieux adapté, adapté en fonction des dernières informations sur le trafic, captées en temps réel par l'autoradio. Au besoin, la voiture intelligente pourrait donner des conseils à son conducteur sur la conduite à adopter, voire prendre des initiatives dans le but d'éviter un accident imminent (par un freinage d'urgence par exemple), et même refuser de fonctionner si elle juge son conducteur inapte à la conduite [Dobias 87].

Ce rêve a donné naissance à plusieurs projets de recherche, d'abord au sein des constructeurs, puis dans le cadre plus vaste de projets Eurêka en Europe, IVHS (Intelligent Vehicle-Highway System) aux Etats-Unis et au Japon. En Europe, le programme le plus ambitieux dans ce domaine est Prometheus.

---

## 1.2. Le programme PROMETHEUS

Le programme PROMETHEUS (PROgraMme for an European Traffic with Highest Efficiency and Unprecedented Safety) est un projet Eurêka, impliquant la quasi-totalité des constructeurs automobiles européens (cf. Figure 1.1), ainsi que de nombreux organismes de recherche publique ou privée. Son but est le développement de systèmes destinés à améliorer la qualité et la sécurité de l'ensemble du système de transport routier. Plus précisément, les objectifs visés sont les suivants:

- ***Le développement de systèmes d'évitement des accidents.*** Il s'agit de fournir une assistance *pertinente* au conducteur dans les situations critiques, afin d'éviter que ces situations n'entraînent des accidents. On a vu en effet que les accidents sont rarement la conséquence d'une cause unique, mais plutôt d'une succession de circonstances dont aucune, prise isolément, ne pourrait provoquer un accident. Il faudra acquérir automatiquement l'ensemble des informations nécessaires à la détection d'un risque d'accident, et présenter ces informations, après traitement informatique, sous une forme adaptée aux besoins du conducteur, c'est à dire aussi complète et synthétique que possible, tout en évitant de saturer ses capacités d'attention en lui présentant des informations non pertinentes, par exemple ce qu'il peut voir par lui-même.
- ***Le développement de systèmes destinés à améliorer le confort du conducteur,*** notamment en réduisant son stress par la délégation des tâches de conduite élémentaires à des équipements intelligents, et par l'optimisation des systèmes de commande et d'information. Outre l'amélioration du confort, de tels dispositifs sont également susceptibles d'améliorer la sécurité en sollicitant le moins possible l'attention du conducteur, qui peut rester vigilant sur les dangers éventuels.
- ***Le développement de systèmes destinés à améliorer la fluidité du trafic.*** Il s'agit d'éviter la saturation du réseau routier, génératrice d'embouteillages,

par l'utilisation de dispositifs embarqués communiquant avec les autres véhicules et des installations basées au sol. Cette fonction nécessite de gros moyens matériels et humains, afin de suivre en temps réel l'évolution du trafic afin d'optimiser les trajets suivis par les différents véhicules.

Les autres objectifs de Prometheus ont moins de rapports avec la sécurité routière, et visent plutôt des considérations écologiques.

- **Le développement de systèmes destinés à réduire la consommation d'énergie**, par la réduction des manoeuvres d'accélération et de freinage dans les flux denses de circulation, par l'optimisation de la conduite en convoi obtenue par une coopération entre véhicules, et par la réduction des détours inutiles lors de la recherche de la destination (systèmes d'aide à la recherche du meilleur trajet). On remarquera le lien avec l'objectif précédent: la fluidité du trafic et les économies d'énergie sont évidemment en rapport étroit.

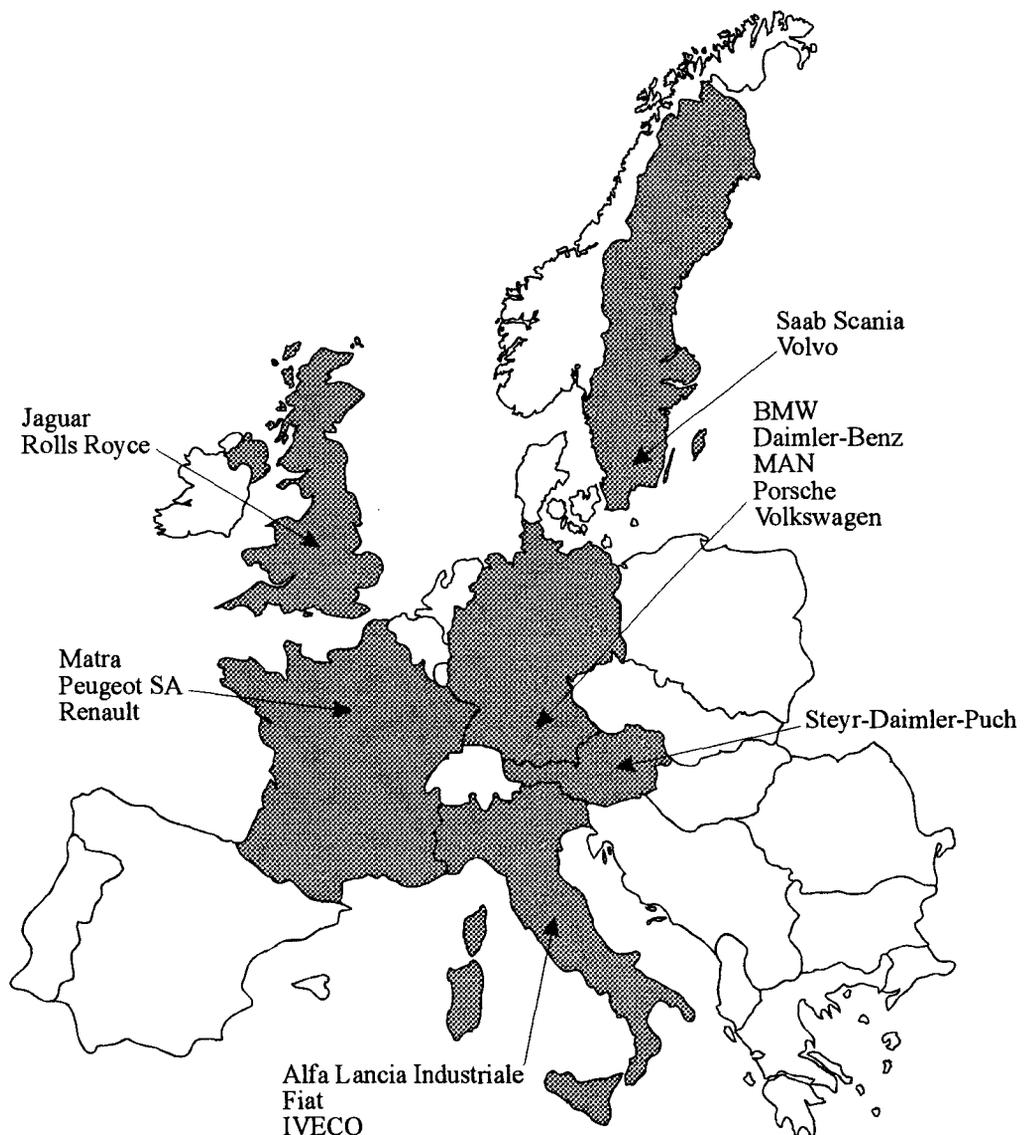


Figure 1.1: Les constructeurs automobiles participant au projet PROMETHEUS.

- **Le développement de systèmes destinés à réduire l'impact sur l'environnement**, par la réduction du trafic non nécessaire, et la diminution du besoin en nouvelles liaisons par une meilleure utilisation des liaisons existantes. On pourrait ajouter que la réalisation de ces deux objectifs entraîne également une diminution de la pollution atmosphérique et des nuisances sonores.

Pour atteindre ces objectifs, le comité de pilotage de PROMETHEUS a défini un ensemble de 23 fonctions, réparties en 4 groupes (cf. Figure 1.2), dont 3 concernent directement la sécurité (fonctions F1 à F14).

Notons que, comme dans le cas de l'ABS déjà mentionné, les fonctions de Prometheus peuvent avoir des effets pervers, d'ailleurs évoqués au sein de PROMETHEUS [Broughton 91, Malaterre 92]:

- Comme on l'a vu plus haut, certains auteurs [Wilde 90] ont mis en évidence une modification du comportement des conducteurs dotés d'équipements de sécurité active, tendant à compenser l'action de ces derniers par une plus grande prise de risque. Cet effet a déjà entraîné l'échec de l'ABS, et rien ne permet de penser qu'il sera absent des dispositifs conçus dans le cadre de PROMETHEUS.
- Il faut faire attention, en mettant en place de nouveaux dispositifs d'assistance, à la cohabitation avec les usagers qui n'en bénéficieront pas. Ainsi, la détection d'obstacles sera probablement plus efficace sur des véhicules équipés de dispositifs coopérants, tels que des transpondeurs radar, alors que les piétons, pourtant plus vulnérables, n'en seront pas équipés et seront peu ou pas détectables. Certains auteurs [Adams 88] ont également montré qu'une amélioration de la sécurité de certains usagers pouvaient augmenter le risque d'autres usagers plus vulnérables, probablement en raison d'une modification du comportement des usagers équipés.
- La présentation des données issues des dispositifs PROMETHEUS ne doit pas détourner l'attention du conducteur de sa tâche prioritaire, qui reste la conduite. Ce risque est surtout lié aux fonctions de confort, telles que l'aide à la navigation, mais il existe également pour les fonctions de prévention des accidents: des informations trop nombreuses ou difficiles à interpréter peuvent solliciter exagérément l'attention du conducteur, et causer l'accident qu'elles étaient censées éviter. On trouvera dans [Van Eslande 1992] une étude fort complète sur les obstacles à la prise en compte par l'utilisateur des informations qui lui sont présentées.
- Enfin, un dispositif de prévention des accidents n'est efficace que s'il fournit des informations *fiabiles*. Il est évident qu'un dispositif de sécurité qui omet de signaler un danger est inefficace; mais un dispositif qui provoque des fausses alertes est vite rejeté par l'utilisateur, parfois au risque d'ignorer des dangers qu'il aurait détectés lui-même sans l'agacement provoqué par de trop nombreuses fausses alertes, version *high-tech* de l'histoire de l'homme qui criait au loup. Ceci vaut en particulier pour la fonction F1 (détection d'obstacles), qui doit être particulièrement robuste.

	FONCTION
Amélioration de l'information au conducteur	<b>F1</b> Détection d'obstacles
	<b>F2</b> Surveillance de la route
	<b>F3</b> Surveillance de l'état du conducteur
	<b>F4</b> Surveillance de l'état du véhicule
	<b>F5</b> Amélioration de la vision
Aide active au conducteur	<b>F6</b> Détermination des marges de sécurité
	<b>F7</b> Détermination des trajectoires critiques
	<b>F8</b> Contrôle dynamique du véhicule
	<b>F9</b> Assistance au conducteur
Conduite coopérative	<b>F10</b> Manoeuvres intelligentes
	<b>F11</b> Régulation de vitesse intelligente
	<b>F12</b> Contrôle d'intersection intelligente
	<b>F13</b> Pré-information à moyenne portée
	<b>F14</b> Transmission d'alerte en cas d'accident
Gestion du trafic et des flottes	<b>F15</b> Routage statique
	<b>F16</b> Routage dynamique
	<b>F17</b> Planification de trajet
	<b>F18</b> Contrôle du réseau routier
	<b>F19</b> Gestion des places de stationnement
	<b>F20</b> Contrôle du flot de circulation
	<b>F21</b> Gestion des demandes de trajet
	<b>F22</b> Connexion avec les transports publics
	<b>F23</b> Gestion de flotte commerciale

Figure 1.2: Les différentes fonctions définies dans PROMETHEUS.

Le travail mené dans le cadre de PROMETHEUS par le Centre d'Automatique de Lille entre dans la fonction F1, destinée à concevoir des dispositifs embarqués de détection d'obstacle. Comme on l'a vu, le cahier des charges de cette fonction est sévère:

- Les dispositifs de détection d'obstacle mis en oeuvre doivent être assez petits et économes, en termes de consommation électrique, pour pouvoir être embarqués dans une voiture particulière sans en affecter les capacités ou le fonctionnement. Ceci vaut pour les capteurs et pour les moyens de traitement informatique qui leurs sont éventuellement associés.

- Ils doivent opérer très rapidement, afin de détecter les obstacles et émettre une alerte avant que le conducteur les aient vus par lui-même. En pratique, le temps de réaction admissible ne dépasse pas quelques dixièmes de secondes.
- Ils doivent être insensibles aux fausses détections, mais détecter tous les obstacles qui se présentent sur la chaussée.
- Ils doivent avoir une portée suffisante (plusieurs dizaines de mètres) pour détecter les obstacles à temps pour permettre son évitement par le conducteur.

## 1.3. Méthodes existantes pour la détection d'obstacles sur route

La détection automatique d'obstacles est depuis longtemps un problème très étudié par les chercheurs. Les méthodes développées sont de deux types: actives ou passives.

### 1.3.1. Méthodes actives de détection d'obstacles

Les détecteurs actifs émettent vers la scène un faisceau d'ondes radio (radar), d'ondes sonores (sonar) ou de lumière (lidar), dont les éventuelles réflexions indiquent la présence d'obstacles. Leur distance est alors connue par le temps de vol du faisceau renvoyé vers le récepteur (cf. Figure 1.3). Leur position peut également être déterminée, grâce à un perfectionnement consistant à mettre l'émetteur et le récepteur en rotation, comme par exemple dans les radars destinés à la surveillance du trafic aérien ou maritime.

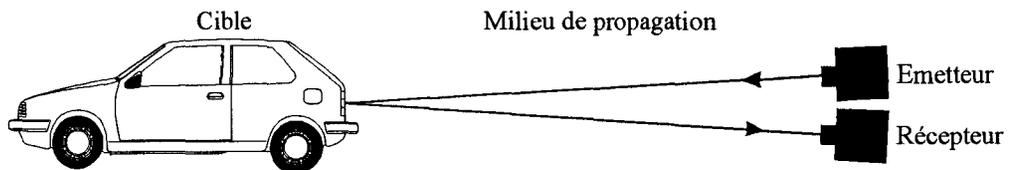


Figure 1.3 : Principe de fonctionnement d'un détecteur actif.

La portée des détecteurs actifs est conditionnée par:

- Les caractéristiques propres de l'émetteur (puissance) et du récepteur (sensibilité, sélectivité). L'angle d'ouverture du faisceau, à la sortie de l'émetteur, est également une caractéristique importante: plus on disperse l'énergie au départ, moins on en récupère à l'arrivée. Pour optimiser ce paramètre, conséquence des phénomènes de diffraction qui se manifestent également sur le trajet des faisceaux aller et retour, il est préférable de travailler à des fréquences élevées.
- L'absorption et la dispersion des faisceaux aller et retour par le milieu de propagation.
- Les propriétés réfléchissantes de la cible:
  - Si elle réfléchit l'intégralité du faisceau vers le récepteur, la détection est possible à grande distance. C'est notamment le cas si elle

est équipée d'un cataphote (pour les lidars) ou d'un transpondeur radar.

- Si, au contraire, elle absorbe le faisceau, le diffuse ou le réfléchit dans une autre direction, voire lui est transparente, la détection devient difficile, voire impossible, même à courte distance. Ces propriétés sont d'ailleurs exploitées pour rendre certains avions de guerre indétectables au radar: ils sont construits en matériaux non conducteurs qui sont transparents aux ondes radar, recouverts de peintures qui les absorbent, et toutes leurs surfaces sont inclinées pour renvoyer le faisceau à la verticale.

Les capteurs actifs sont classés selon leur principe de fonctionnement:

- Les **radars** sont bien adaptés à la détection des véhicules, en raison de leur masse métallique importante qui réfléchit très bien les ondes radio. De plus, la précision de localisation est bonne, et la largeur du faisceau, qui définit les limites de la zone de détection, peut être déterminée "à la demande" par les caractéristiques de l'antenne. Un faisceau étroit associé à un balayage horizontal permet de donner une carte de profondeur de la scène, mais avec une certaine lenteur et une fiabilité amoindrie par l'utilisation d'un dispositif mécanique [Deloof 90]. Des procédés de balayage électronique permettent d'éviter ces deux inconvénients, mais leur prix très élevé les réserve pour l'instant aux applications militaires.

Les êtres humains peuvent être détectés jusqu'à une distance de 80 m, mais les réflexions parasites causées par le fond de la scène (panneaux routiers ou piles de ponts par exemple) créent des fausses alarmes fréquentes, ce qui diminue l'intérêt du dispositif. La pluie introduit également des réflexions parasites. Il est néanmoins envisagé de limiter ces fausses alarmes par un usage intensif de transpondeurs sur les équipements routiers et un traitement du signal sophistiqué.

Notons enfin la possibilité, en modulant les faisceaux radar, d'utiliser ceux-ci comme un moyen de communication entre les véhicules. Cette possibilité n'est cependant envisagée que dans un futur relativement lointain [Deloof 90]. Les applications des radars au trafic routier sont notamment étudiées en Europe dans le cadre du projet SMILER (projet DRIVE V1002).

- Les **sonars** sont constitués d'un haut-parleur qui émet des impulsions ultrasonores, dont les réflexions sont captées par un microphone. Les avantages sur les radars sont un coût, une consommation électrique et un encombrement nettement inférieurs. De plus, tous les obstacles acoustiquement réfléchissants peuvent être détectés par ultrasons, ce qui représente la majorité des objets présents dans la nature. Il faut cependant remarquer, comme pour tous les détecteurs actifs, l'influence de l'angle d'incidence et de la nature (réfléchissante ou diffusante) du matériau constituant l'obstacle.

En raison de la diffraction, due à la longueur d'onde importante, la dispersion du faisceau est très forte: les capteurs fabriqués par Polaroid, par exemple, ont un angle d'ouverture minimal de  $10^\circ$  environ [Wane 93]. De plus, et pour la même raison, le récepteur est peu directif. Ces

deux caractéristiques limitent la portée du dispositif et empêchent une localisation précise des objets détectés [Boissier 85]. Sur route, la vitesse élevée induit un fort vent dans le capteur, ce qui peut le saturer (comme tout microphone) et limite encore la portée du dispositif. Dans la pratique, les capteurs acoustiques ont été abandonnés après quelques recherches, car les essais menés sur route avec ces dispositifs ont indiqué une portée utile de 12 m (40 m avec répondeur), même dans de bonnes conditions atmosphériques [Deloof 90]. Ils sont, en revanche, souvent utilisés sur des robots mobiles comme capteurs de proximité [Wane 93], par exemple pour suivre un mur à distance constante [Briot 79, Briot 81]. Dans ce domaine, des algorithmes permettant de modéliser l'environnement à partir de capteurs ultrasonores ont également été développés [Crowley 86, Khoumsi 86].

- Les *lidars* fonctionnent sur le même principe que les radars, mais utilisent un faisceau lumineux — généralement issu d'un laser — au lieu d'un faisceau d'ondes radio. La détection est possible sur tout type d'objets, métalliques ou non, mais les objets sombres et mats posent évidemment des problèmes (réflexion trop faible pour être détectée par le récepteur), de même que les surfaces transparentes (par exemple les vitres), qui sont traversées par le faisceau laser sans le réfléchir. Par temps clair, le lidar offre une bonne détection et, surtout, une localisation précise des obstacles. La présence de brouillard, de pluie ou de neige disperse et/ou absorbe les faisceaux aller et retour, et limite la portée pratique, tandis que le soleil peut parfois saturer le récepteur et le rendre inopérant. La poussière qui se dépose sur les lentilles frontales du lidar réduit également la portée [Deloof 90]. [Alizon 90] décrit une réalisation très performante de lidar destinée à être embarquée dans une voiture, conçue dans le cadre du projet Prometheus.

Notons que les détecteurs basés sur des lidars peuvent être également basés sur d'autres principes que le temps de vol: il s'agit des télémètres optiques à déphasage ou à triangulation [Monchaud 86].

### 1.3.2. Méthodes passives de détection d'obstacle

Les méthodes passives consistent à analyser l'image de la route captée par une ou plusieurs caméras (cf. Figure 1.4). Les avantages sont une meilleure résolution et une exploration plus rapide de la scène qu'avec les méthodes actives de détection. Cependant, l'interprétation des images implique un traitement informatique, généralement long. Le principe de fonctionnement des capteurs passifs suppose également un contraste suffisant entre les images des obstacles et celle de la chaussée: si on veut détecter un obstacle, il faut d'abord le voir! Ceci entraîne que, comme dans le cas du lidar, une mauvaise visibilité (due par exemple à la pluie ou au brouillard) compromet la détection.

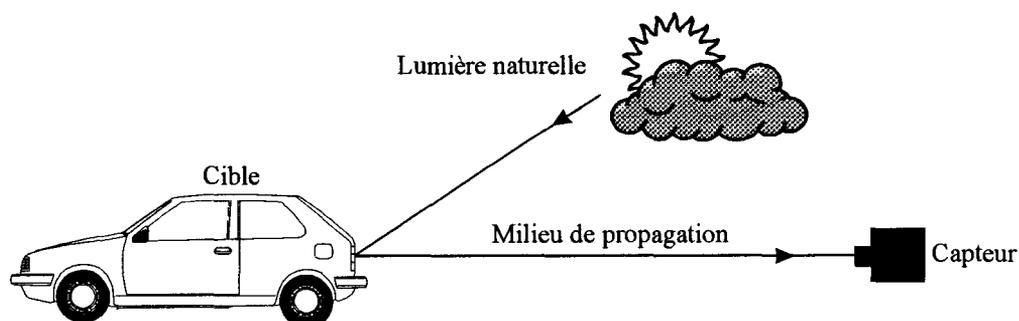


Figure 1.4 : Principe de fonctionnement d'un détecteur passif.

Il existe trois grandes approches pour la détection passive d'obstacles:

- La simple segmentation d'une image monoscopique permet de détecter les obstacles, mais elle entraîne de très nombreuses erreurs, car les ombres et les marquages de la chaussée sont systématiquement pris pour des obstacles. Cette approche purement bidimensionnelle est mal adaptée à la détection des obstacles, qui se distinguent essentiellement par leur relief et leur mouvement relatif.

[Graefe 90] présente une méthode de ce type pour détecter des obstacles sur route: l'image de la route devant le véhicule est scrutée afin d'y rechercher des obstacles potentiels qui sont signalés par les irrégularités du niveau de gris, la chaussée étant supposée d'un gris uniforme. Les fausses alarmes — très nombreuses — sont éliminées en ne conservant que les objets de forme rectangulaire. Ainsi, les ombres sont *souvent* éliminées (mais pas toujours), mais les obstacles de forme plus complexe sont également souvent ignorés. Les exemples présentés ne comportent d'ailleurs ni piétons, ni cyclistes.
- **L'analyse du mouvement** permet de différencier les objets verticaux (obstacles et autres véhicules) des objets horizontaux (chaussée). La démarche la plus répandue consiste à extraire le flot optique de l'image, avant de le segmenter pour ne conserver que les objets verticaux. On trouvera dans [Enkelmann 90] une présentation exhaustive des méthodes basées sur le flot optique appliquées à la détection d'obstacles.

Une variante de ce type de méthode [Carlsson 90] utilise une méthode de prédiction du mouvement apparent des formes détectées, en les supposant appartenir à la surface de la chaussée. L'image suivante prédite est ensuite comparée à l'image suivante réelle, saisie par la caméra: des différences importantes entre ces deux images indiquent la présence d'obstacles. Cette méthode fonctionne bien, mais présente une limitation gênante: elle détecte les obstacles, mais ne les localise pas.
- La **stéréovision** est une approche particulièrement intéressante: connaissant la position d'un objet dans l'espace, il est facile de déterminer s'il appartient à la surface de la chaussée. Ainsi, les marquages au sol et les ombres n'entraînent pas de fausses alarmes. La signalisation verticale, ou les piétons sur les trottoirs, ne sont pas non plus détectés comme des obstacles, car on peut déterminer qu'ils ne se trouvent pas sur la route. De plus, la connaissance de la position exacte d'un obsta-

cle est une information très importante pour définir une stratégie d'évitement.

La stéréovision présente cependant l'inconvénient majeur d'impliquer une quantité de calculs importante, incompatible avec la rapidité nécessaire à la détection d'obstacles en temps réel.

Nous avons cependant choisi cette approche pour concevoir un dispositif de détection d'obstacles. En effet, les moyens technologiques actuellement disponibles, correctement utilisés, permettent de résoudre ce problème dans certains cas. Nous verrons au Chapitre 2 que le problème de la détection d'obstacles sur route, qui est traité dans ce mémoire, se réduit à l'un de ces cas.

## 1.4. Considérations sur la stéréovision

Bien qu'il ne soit pas dans notre propos de faire un exposé détaillé sur ce sujet, il est indispensable de donner ici un ensemble de définitions et de mises au point qui seront utilisées dans la suite.

Les systèmes destinés à la reconstruction de la scène tridimensionnelle observée par un ensemble de caméras sont très variés. Cependant, leur organisation générale est toujours la même; elle est directement inspirée de l'étude de la vision animale et humaine, selon un concept connu sous le nom de *paradigme de David Marr* [Marr 82]. La Figure 1.5 donne la structure générale des traitements effectués dans un système de stéréovision. On y trouve plusieurs tâches successives:

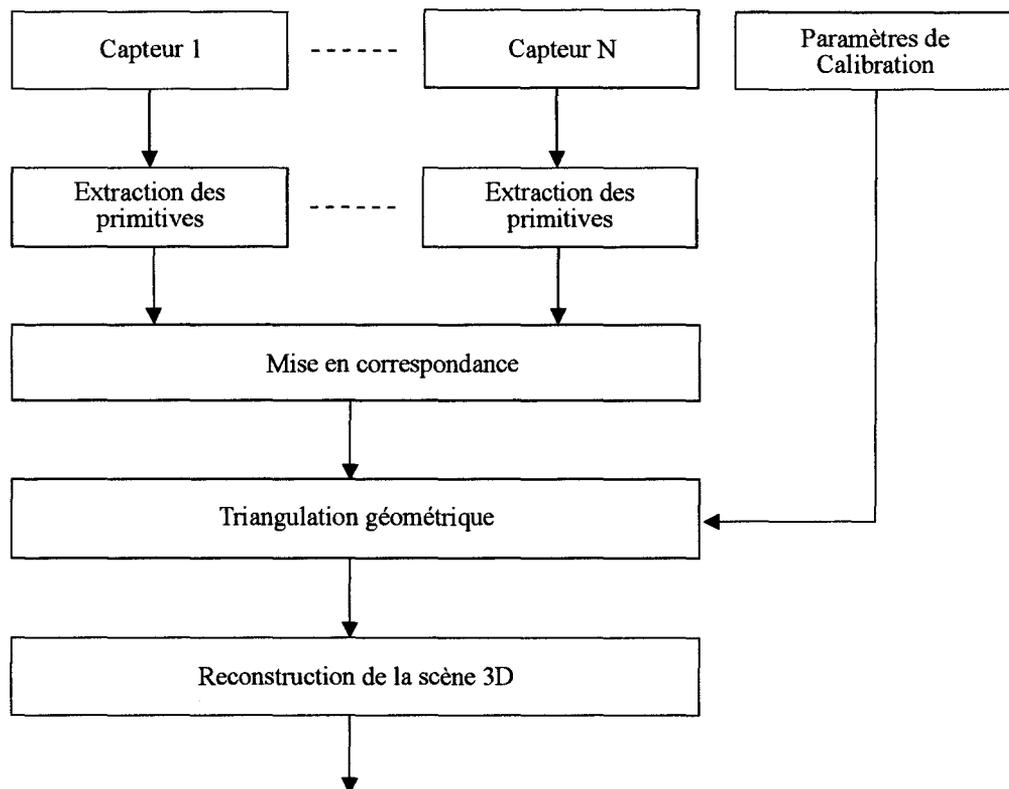


Figure 1.5 : Structure générale des traitements de stéréovision.

### 1.4.1. Prise de vue

Les différents *capteurs* d'image (caméras) constituent les entrées du système. Les dispositifs de prise de vue utilisés selon l'application sont très variables:

- Le *nombre de capteurs* est très variable. Bien qu'il soit plus évident d'utiliser deux caméras, certains auteurs préconisent l'utilisation de trois caméras pour améliorer la mise en correspondance (voir notamment [Ayache 89] et [Horaud 93-2]), alors que d'autres préfèrent n'en utiliser qu'une seule, ce qui simplifie le calibrage. Dans ce dernier cas, le relief peut être déterminé en faisant évoluer les objets [Dreschler 82] et/ou la caméra dans la scène [Adiv 85], selon des méthodes du type *depth from motion* ou *structure from motion* [Jerian 90]), ou en analysant les ombres des objets présents dans la scène en supposant connu l'éclairage (*structure from shading*).
- Leurs *caractéristiques géométriques* sont de deux types: *matricielles* (caméras vidéo par exemple) fournissant des images en deux dimensions, ou *linéaires*, fournissant des images à une seule dimension, comparables à des images vidéo comportant une seule ligne.

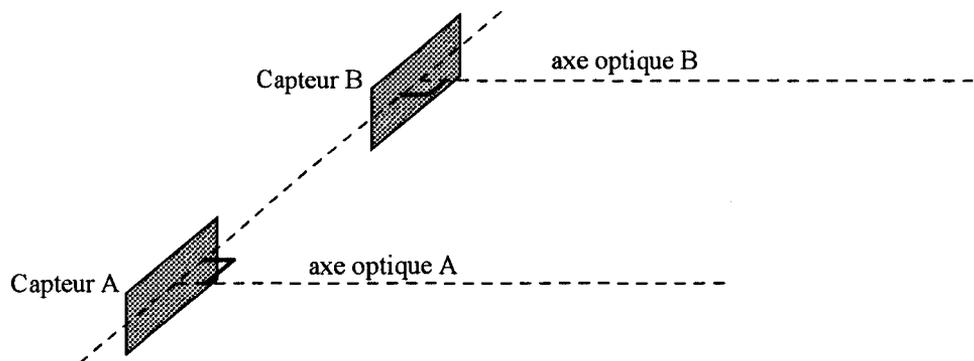


Figure 1.6: Disposition des caméras à axes parallèles.

- La *disposition* des différentes caméras est également une caractéristique importante du dispositif de prise de vue. La disposition la plus avantageuse est à axes optiques parallèles, tous les capteurs coplanaires, leurs lignes de balayage étant parallèles (cf. Figure 1.6). Cette disposition simplifie les calculs et assure la meilleure précision de localisation dans la scène tridimensionnelle. Cependant certains auteurs, jugeant trop difficile d'utiliser des moyens d'ajustage mécaniques, étudient le cas où la disposition et l'orientation des différentes caméras sont quelconques: voir [Ayache 89] et surtout [Horaud 93-2].

### 1.4.3. Extraction des primitives

Il est inenvisageable de trouver la position dans la scène de tous les points de l'image, et ceci pour au moins deux raisons:

- Ce serait beaucoup trop long: pour mettre en correspondance tous les pixels de deux images vidéo de 512x512 pixels, il faudrait comparer chaque pixel d'une image avec chaque pixel de l'autre image, ce qui revient à effectuer  $(512 \times 512) \times (512 \times 512)$ , soit plus de 68 milliards de

comparaisons. Les machines existantes les plus rapides (genre Cray) auraient besoin de plusieurs minutes pour effectuer un tel travail. Quant aux machines plus courantes, stations de travail ou ordinateurs personnels, il leur faudrait des heures.

- L'information locale contenue dans un pixel unique, c'est à dire son niveau de gris, est insuffisante pour permettre sa mise en correspondance sans ambiguïté. Par exemple, un point particulier d'une image donnée peut être associé indifféremment avec tous les points de toute autre image qui ont le même niveau de gris.

C'est pourquoi il est nécessaire de travailler sur des parties de l'image dont le contenu informationnel est plus riche. Ces parties, appelées *primitives*, peuvent être de différentes natures [Heraud 93-2]:

- Des *points d'intérêt* comme, par exemple, les points anguleux des contours, sont souvent utilisés. L'utilisation des points d'intérêt en stéréovision est avantageuse, car la "forme" de ces primitives ne change pas selon l'angle de prise de vue. En contrepartie, l'information attachée à ces points est relativement faible, ce qui oblige, pour la mise en correspondance, à considérer les niveaux de gris des pixels situés leur voisinage [Barnard 80] ou les segments joignant les points d'intérêt [Bhanu 84].
- Des *contours*, le plus souvent segmentés sous forme de segments de droites [Ayache 89, Heraud 88, Medioni 85] ou de portions de coniques pour aboutir à une représentation plus synthétique. L'inconvénient de ces primitives est que, en raison des effets de perspective, les images du même contour vu par plusieurs caméras n'ont généralement pas la même forme, et de tels effets sont complexes à prendre en compte lors de la mise en correspondance. Dans la pratique, l'utilisation des contours n'est réellement intéressante qu'avec des objets polyédriques [Heraud 93-2].
- Des *régions homogènes*, c'est à dire des régions de l'image dans lesquelles le niveau de gris est relativement uniforme, sont des primitives peu utilisées en stéréovision. En effet, comme pour les contours, leur mise en correspondance est souvent rendue difficile par les déformations dues aux effets de perspective sur les images saisies par des caméras ayant des angles de prise de vue différents [Lee 90]. Par contre, des scènes fortement texturées peuvent plus être facilement décrites par quelques régions de texture homogène que par des primitives de type contour, qui sont dans ce cas très nombreuses. [Lumia 83, Mitchell 78] sont deux exemples de méthodes de segmentation d'image utilisant l'analyse de texture. [du Buf 90] passe en revue d'autres méthodes de ce type. Notons cependant que, si les primitives du type "région de texture homogène" occupent moins de place en mémoire dans certains cas, le temps de calcul n'en est pas nécessairement réduit, car l'analyse de la texture est souvent très lente en raison de l'utilisation de la transformée de Fourier ou des matrices de co-occurrence.

La recherche de primitives dans une image, très rapide chez les êtres vivants (quelques centièmes de secondes selon David Marr), est très lente avec les ar-

chitectures informatiques et les algorithmes actuels: [Horaud 93-2] donne un temps typique de 30 secondes environ pour extraire les contours d'une seule image de 512x512 pixels sur Sun 4 (12,5 Mips). L'un des enjeux fondamentaux de la vision artificielle est, encore actuellement, de concevoir une méthode rapide et efficace permettant d'extraire sans erreur les primitives d'une image. Des architectures matérielles spécialisées ont été conçues dans ce but (notamment les *réтины artificielles*: voir par exemple [Keast 93]), mais leur production industrielle semble encore problématique.

#### 1.4.4. Mise en correspondance (ou appariement)

A partir des ensembles de primitives trouvées dans N images différentes de la même scène vue par N caméras, cette tâche détermine l'ensemble des N-uplets de primitives (une dans chaque image) correspondant à un même objet dans la scène (cf. Figure 1.7).

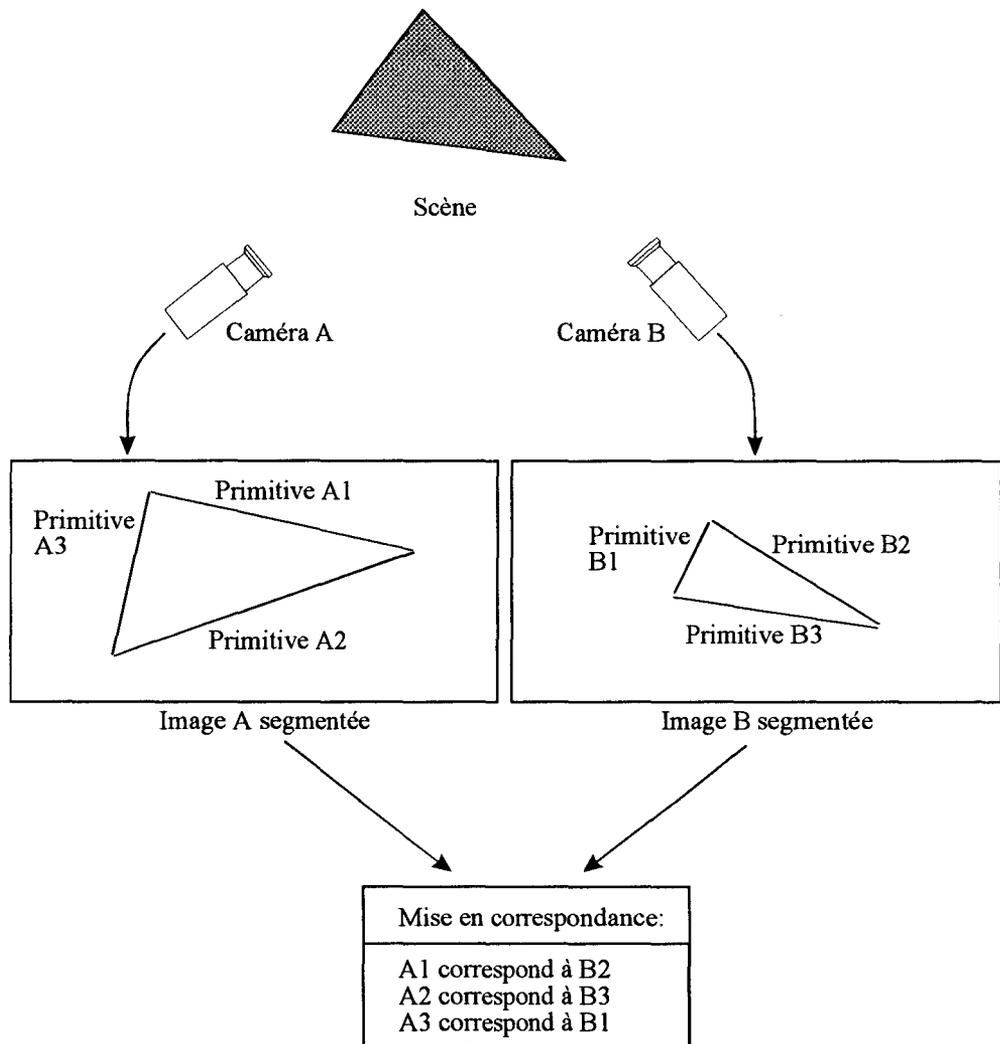


Figure 1.7: Mise en correspondance de primitives (ici, sur 2 images).

Selon la nature des primitives et le domaine d'application, il existe de nombreuses méthodes pour effectuer la mise en correspondance, qui se regroupent pour la plupart en quatre grandes approches [Heraud 93-2]:

- Les *méthodes pyramidales* représentent les images sous plusieurs résolutions, de la plus fine (celle du capteur) à la plus grossière. Plusieurs mises en correspondance successives sont alors effectuées, en partant de la plus faible résolution. Les résultats trouvés à un niveau de résolution donné sont utilisés pour initialiser la mise en correspondance au niveau suivant. On pourra trouver des exemples de méthodes pyramidales dans [Aubert 89] et dans [Heraud 93-2].
- Les méthodes utilisant la *programmation dynamique* représentent les images gauche et droite comme les deux dimensions d'une matrice contenant tous les appariements possibles. L'algorithme cherche alors les appariements qui donnent le "meilleur" chemin pour aller d'une extrémité à l'autre de la matrice, en minimisant une fonction de coût. Le principe de base de la programmation dynamique (toute partie d'un chemin optimal est elle-même optimale) impose deux contraintes: le sens de parcours de la matrice doit toujours être le même (les retours en arrière sont interdits), et chaque primitive d'une image ne peut intervenir qu'une fois. Une méthode très connue de ce type est décrite dans [Ohta 85].
- Les méthodes faisant appel à la *relaxation* commencent, dans une phase d'initialisation, par attribuer à chaque appariement possible un indice de confiance dépendant des appariements voisins. Ensuite, dans une phase itérative, les appariements sont comparés entre eux pour modifier les indices de confiance dans le but de réduire les incohérences d'appariement entre points voisins. A la fin du processus, un seul appariement est retenu pour chaque primitive issue de chaque image. On trouvera des algorithmes de relaxation dans [Ballard 82, Lee 90, Bhanu 84] ou dans [Pollard 85], dont [Heraud 93-2] donne un résumé.
- L'utilisation de l'*isomorphisme de graphes* utilise une représentation des images gauche et droite sous forme de graphes, dont les noeuds représentent les primitives, tandis que les arcs représentent les relations (de position notamment) entre ces primitives. Dans l'idéal, les graphes construits à partir des deux images devraient être identiques. Dans la réalité cependant, les occlusions, les différences de segmentation (un même contour étant par exemple divisé en un nombre différent de primitives dans les deux images), etc., introduisent des différences entre les deux graphes. On trouvera notamment dans [Skordas 88, Heraud 88, Heraud 89] comment on peut malgré tout trouver des appariements dans ce cas.

Ces approches, bien que très différentes, font cependant toutes appel à trois outils essentiels:

- Un *critère de ressemblance* entre les primitives que l'on cherche à mettre en correspondance. Le correspondant sera la primitive la plus "ressemblante" — c'est à dire celle qui maximise la valeur du critère — parmi toutes celles qui satisfont aux contraintes. Le critère de ressemblance est aussi appelé *fonction de coût*.

- Des *contraintes*, qui expriment les règles géométriques de formation des images d'un objet sur les différents capteurs, et qui permettent de réduire le temps de calcul et d'augmenter la robustesse, en limitant la zone de recherche des primitives correspondantes ou en imposant des conditions sur la position relative des appariements trouvés dans la scène. [Ayache 89, Horaud 93-2], et de manière générale tous les ouvrages traitant de la mise en correspondance, donnent des exemples de contraintes géométriques couramment utilisées.
- Des *méthodes de vérification de la cohérence des résultats*. En effet, il arrive souvent qu'une primitive donnée soit, par l'application du seul critère de ressemblance, mise en correspondance avec plusieurs autres (on parle alors d'*appariements multiples*). En faisant appel aux différentes contraintes, il faut choisir, parmi ces prétendantes, laquelle correspond effectivement au même objet. La vérification de la cohérence des résultats réside en général dans l'algorithme de mise en correspondance lui-même.

Quelque soit la méthode utilisée, la mise en correspondance des primitives consomme une puissance de calcul importante: [Meygret 90] indique, avec sa méthode, un temps de 15 secondes sur Sun 4 pour mettre en correspondances deux images de 512x144 pixels (1445 primitives). Il s'agit là d'un algorithme rapide: de nombreux auteurs citent des temps de calculs encore plus longs, atteignant souvent plusieurs dizaines de minutes, comme par exemple [Ohta 85].

#### 1.4.5. Triangulation géométrique

Les calculs de triangulation géométrique permettent, en fonction des positions dans les images des différentes primitives mises en correspondance, de déterminer la position de ces dernières dans la scène tridimensionnelle. Ceci est réalisé, selon le modèle géométrique des caméras, par une simple multiplication matricielle ou par la résolution d'un système d'équations, polynomiales ou non [Zhou 92].

Cette procédure utilise un modèle géométrique du système de prise de vue. Ce modèle décrit la formation de l'image sur les capteurs des caméras, dont les caractéristiques sont décrites par les *paramètres de calibration*. Ces paramètres expriment la position et l'orientation des caméras dans le repère de travail (*paramètres extrinsèques*), et les caractéristiques de la projection d'une image sur leurs capteurs (*paramètres intrinsèques*). Ils seront décrits plus précisément dans le Chapitre 2.

#### **Calibrage, calibration et étalonnage**

Il est nécessaire de faire ici une remarque concernant le vocabulaire. En effet, il est d'usage courant d'appeler "calibration" trois opérations de natures très différentes, ce qui peut poser au lecteur des problèmes de compréhension.

Cette situation regrettable est liée au fait que le mot anglais "calibration", qui n'existe pas en français, a un sens très général qui recouvre trois notions très différentes:

- Le **calibrage**, qui consiste à mettre un dispositif en conformité avec une norme. Par exemple, on *calibre* une entrée d'oscilloscope pour qu'une tension d'un volt donne une déviation verticale d'une division sur le calibre 1V/div.
- L'**étalonnage**, qui consiste à mesurer la réponse d'un dispositif — en général non calibrable — pour en déduire sa précision ou ses limites d'utilisation. Par exemple, on *étalonne* la réponse photométrique d'une pellicule photographique pour déterminer la plage d'illumination qui donne les meilleures images.
- Enfin, la prise en compte dans les calculs du calibrage d'un dispositif (en calculant la tension d'après la déviation de la trace dans le premier exemple) ne porte pas de nom général en français.

La confusion entre ces trois notions nuit à la compréhension de nombreux travaux publiés en français sur la stéréovision. C'est pourquoi nous avons choisi d'utiliser dans ce mémoire un vocabulaire plus précis:

- La détermination des paramètres de calibration est effectuée lors d'une phase préliminaire appelée **calibrage**. Le calibrage du stéréoscope linéaire fera l'objet du Chapitre 3.
- La mesure de la précision géométrique du système calibré est également faite lors d'une phase préliminaire à l'exploitation du dispositif stéréoscopique. Nous appelons cette phase l'**étalonnage** du dispositif. Nous réaliserons un étalonnage dans le Chapitre 5, afin de déterminer la précision du stéréoscope linéaire que nous aurons construit.
- La prise en compte des paramètres de calibration dans la phase de calibration géométrique, lors de l'exploitation du dispositif stéréoscopique, sera appelée **calibration**, comme on le verra au Chapitre 2. Ce choix se justifie par le fait que, dans la littérature, ce nom est le plus souvent employé dans ce sens. Le lecteur ne devrait donc pas être trop perturbé dans ses habitudes.

#### 1.4.6. Reconstruction de la scène 3D

La reconstruction de la scène 3D utilise les positions tridimensionnelles des primitives mises en correspondance pour reconstituer une représentation plus synthétique et plus abstraite de la scène (cf. Figure 1.8). Cette tâche permet de reconstituer les différents objets présents dans la scène alors que, dans les étapes antérieures, seule existait la notion de primitive, sans relation immédiate avec les objets physiques. Un modèle de la scène, c'est à dire un ensemble de connaissances et de règles sur les objets à reconstruire et sur leurs relations, est nécessaire.

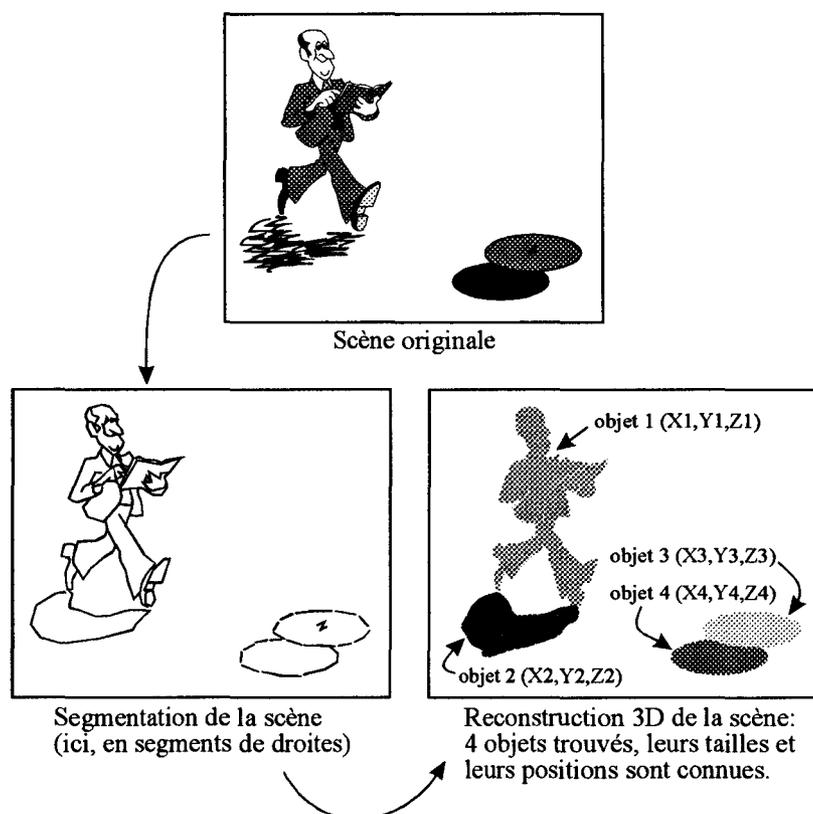


Figure 1.8: Reconstruction 3D de la scène à partir des primitives mises en correspondance.

## 1.5. Organisation du mémoire

Le travail effectué au Centre d'Automatique dans le cadre du projet Prometheus comprend des aspects très variés, puisqu'il s'agit, d'une part, de développer un nouveau type de capteur délivrant des séquences d'images stéréoscopiques linéaires, et d'autre part d'utiliser ce capteur pour détecter des obstacles devant un véhicule. Ces deux aspects ont justifié un découpage du travail de recherche en deux parties distinctes, qui correspondent à deux thèses différentes. La présente thèse est axée sur le développement du capteur stéréoscopique, sa mise en oeuvre et sa validation. La deuxième thèse, qui sera soutenue ultérieurement par Jean-Christophe Burie, portera davantage sur les algorithmes de détection d'obstacle.

Le **chapitre 2** de ce mémoire expose le principe du capteur d'images utilisé dans notre dispositif de détection d'obstacles (cf. § 1.4.1). Une attention particulière est portée à la définition de sa géométrie, ainsi qu'à la présentation de ses principaux composants: les caméras linéaires. C'est également dans ce chapitre qu'est présentée la conception mécanique de ce stéréoscope linéaire.

Le **chapitre 3** développe la méthode de calibrage géométrique du stéréoscope, méthode spécialement conçue pour le cas particulier de la stéréovision linéaire, et qui permet de déterminer les paramètres de calibration (cf. § 1.4.5).

Le **chapitre 4** concerne le traitement informatique des images linéaires, dont la structure générale a été donnée dans les § 1.4.3 à 1.4.5. Seule une version pré-

liminaire de ce traitement sera donnée ici, le but étant de vérifier le bon fonctionnement du stéréoscope linéaire et d'évaluer ses performances. Toutefois, cette première version servira de point de départ à l'élaboration des traitements définitifs, destinés, cette fois, à la détection des obstacles.

Le **chapitre 5** présente les séquences d'images réalisées avec le stéréoscope linéaire durant des essais en conditions réelles, et les résultats de la détection d'obstacles effectuée durant ces essais. La méthode d'enregistrement des images stéréoscopiques linéaires est également décrite dans ce chapitre.

Enfin, le **chapitre 6**, conclusion de ce mémoire, propose des évolutions possibles du stéréoscope linéaire pour l'adapter à d'autres domaines d'application que la détection d'obstacles. En effet, les méthodes développées ici privilégient la rapidité, la précision de localisation n'étant pas un facteur critique en détection d'obstacles. Des méthodes de calibrage et de triangulation géométrique plus évoluées — mais plus lentes — peuvent améliorer sensiblement la précision de la localisation, par exemple pour des applications de contrôle industriel ou de métrologie.



## Chapitre 2

# DISPOSITIF DE DETECTION AUTOMATIQUE D'OBSTACLES SUR LA CHAUSSEE

Ce Chapitre décrit le principe de fonctionnement du capteur que nous avons développé dans le cadre du programme Prometheus, et que nous appellerons *stéréoscope linéaire* dans la suite. Après une description générale des capteurs (cf. § 2.1), puis des caméras linéaires (cf. § 2.2), nous verrons comment utiliser ces caméras en couple stéréoscopique pour déterminer la position d'un objet dans l'espace (cf. § 2.3). A la fin de ce chapitre sont présentés les détails de la conception électronique (cf. § 2.4) et mécanique (cf. § 2.5) du stéréoscope.

## 2.1. Les capteurs linéaires

### 2.1.1. Structure des capteurs linéaires

Le capteur linéaire est une version simplifiée des capteurs CCD matriciels utilisés dans les caméras vidéo. Il comporte essentiellement une rangée de photodétecteurs (ou *photosites*), des registres à décalage, des amplificateurs de sortie, et des circuits annexes destinés à la gestion des registres à décalage (cf. Figure 2.1).

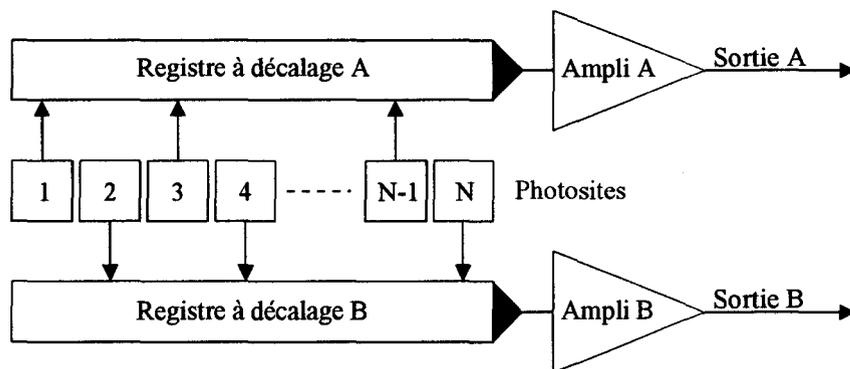


Figure 2.1 : Structure d'un capteur CCD linéaire.

On remarque sur la Figure 2.1 la présence de deux registres à décalage analogiques, de part et d'autre de la rangée de photosites, qui reçoivent, l'un les pixels pairs, l'autre les pixels impairs. Cette disposition, commune à la quasi-totalité des capteurs linéaires, permet d'optimiser les performances du capteur sans augmenter ses dimensions, en réduisant le nombre de décalages en augmentant la dynamique admissible dans les registres à décalage. Les deux amplificateurs de sortie — non systématiques mais souvent adoptés — permettent la sortie simultanée de deux pixels successifs, ce qui rend possible des acquisitions très rapides avec des composants standards, en faisant appel à deux chaînes d'acquisition fonctionnant en parallèle. En contrepartie, les dispersions des caractéristiques de ces deux chaînes de traitement parallèles engendrent un bruit particulier (*bruit de pair-impair*) qui donne à l'image d'une plage uniforme un aspect en peigne très reconnaissable (cf. Figure 2.2). Heureusement, ce bruit est généralement d'amplitude faible, et altère très peu le signal de sortie.

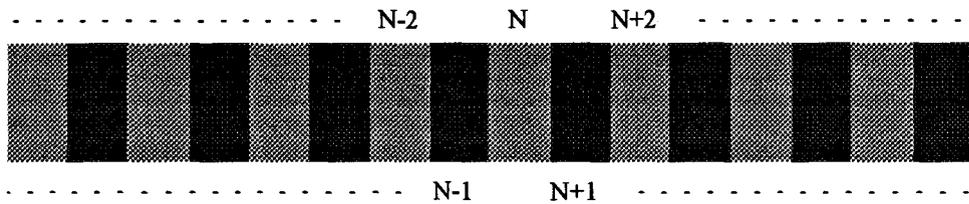


Figure 2.2 : Image d'une plage uniforme avec bruit de pair-impair.

Un photosite a la structure représentée sur la Figure 2.3, dérivée de celle d'un transistor MOS-FET. Les photons reçus sur la surface du photosite traversent la grille, qui est trop mince pour être opaque, et transmettent leur énergie aux atomes du silicium en brisant les liaisons covalentes de celui-ci. Ceci entraîne l'apparition de paires électron-trou, qui sont séparées par le champ électrique imposé par la grille maintenue à un potentiel positif  $V_{PG}$ , de telle sorte que les électrons viennent s'accumuler sous cette dernière. Le puits de potentiel imposé par  $V_{PG}$  empêche les charges de s'échapper dans le substrat. Une deuxième grille, dite *grille de transfert*, fait office de "porte de sortie". En imposant  $V_{TG} > V_{PG}$ , on abaisse suffisamment la barrière de potentiel pour permettre l'écoulement des charges vers une cellule du registre à décalage (cf. Figure 2.4).

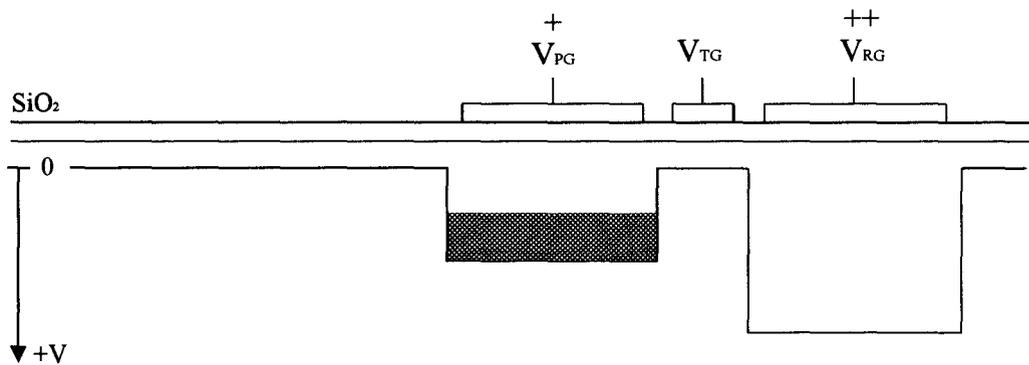


Figure 2.3 : Structure d'un photosite.

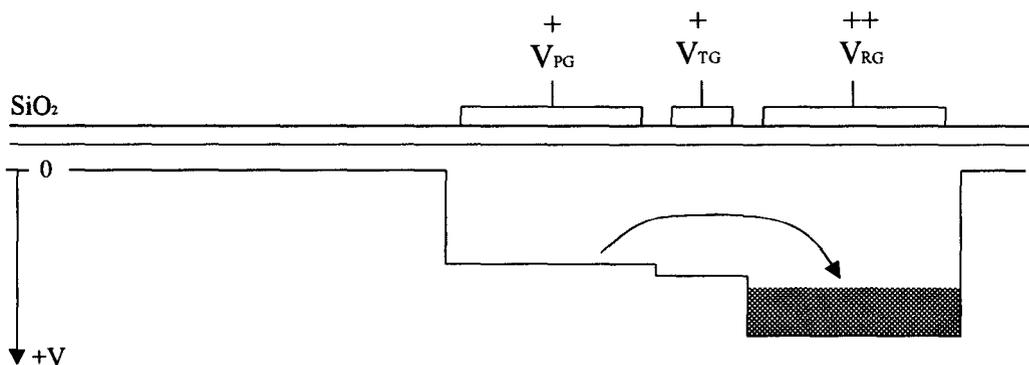


Figure 2.4 : Vidage du photosite dans le registre à décalage.

Les premières versions de capteurs CCD présentait un défaut appelé *smear* : si la charge accumulée dans une cellule devenait trop importante, en raison d'un éclairage trop intense, elle pouvait déborder le puits de potentiel imposé par  $V_{PG}$ , et l'excédent se déversait dans le registre à décalage ou dans les cellules

voisines. L'effet sur l'image était que les zones claires semblaient baver sur les zones sombres (cf. Figure 2.5). Les capteurs récents intègrent des dispositifs — différents selon les constructeurs — destinés à empêcher le smear. Par exemple, les capteurs fabriqués par Fairchild comprennent, pour chaque photosite, une grille supplémentaire maintenue à une faible tension  $V_{CE}$  qui impose un potentiel plus bas que celui de la grille de chargement du photosite (cf. Figure 2.6). Lorsque les charges accumulées dans le photosite excèdent ce potentiel, elles se déversent sous une quatrième grille, maintenue à un potentiel  $V_{SINK}$  élevé qui impose un puits de potentiel très profond, sans affecter les autres photosites, ni le registre à décalage [Murphy 87].

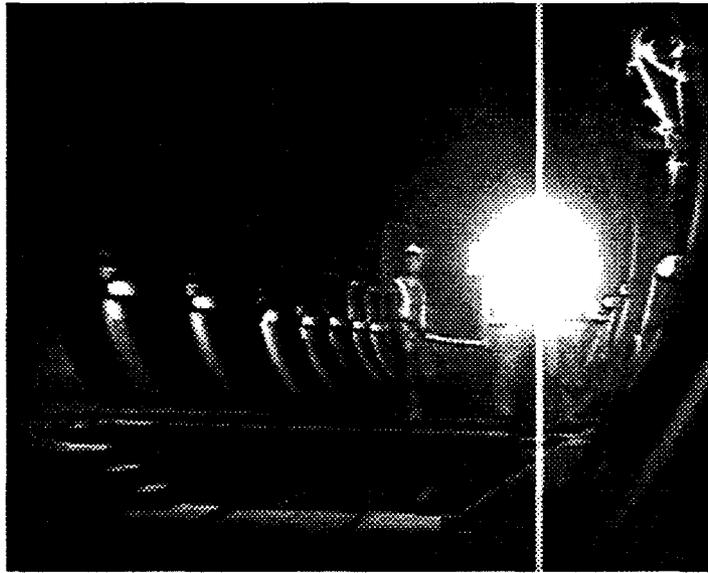


Figure 2.5 : Effet du smear sur une image vidéo.

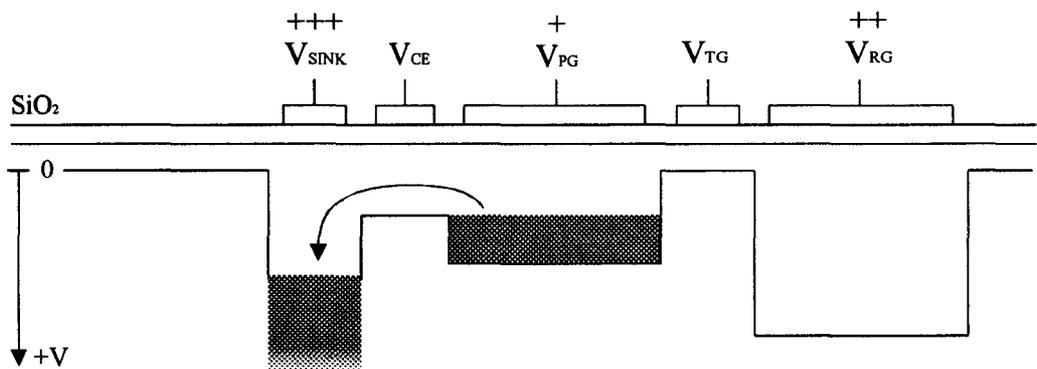


Figure 2.6 : Structure d'un capteur anti-smear.

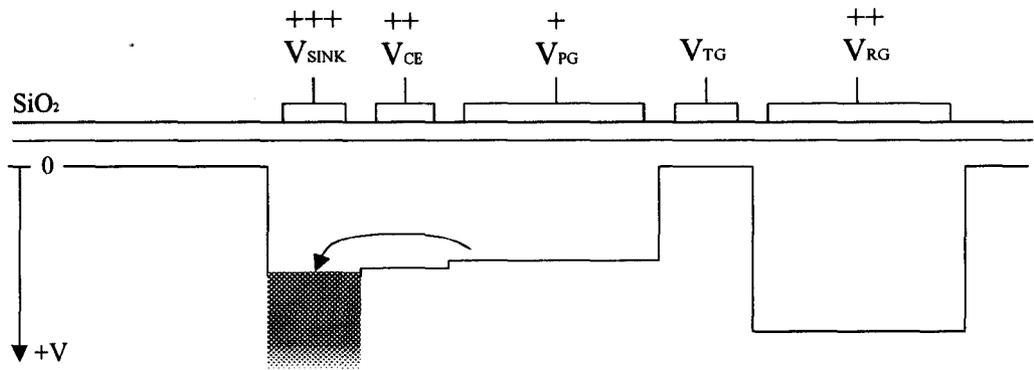


Figure 2.7 : Structure d'un capteur avec obturateur électronique.

Une évolution du dispositif anti-smear permet de contrôler le temps d'exposition: en imposant sur la grille anti-smear un potentiel  $V_{CE}$  suffisamment élevé, toute la charge accumulée dans le photosite se déverse dans le puits de potentiel créé par  $V_{SINK}$  (cf. Figure 2.7). Quand on ramène  $V_{CE}$  à sa valeur normale, les charges recommencent à s'accumuler dans le photosite. Cette fonction, qui permet d'ajuster le temps d'exposition avec une grande précision, est nommée *obturateur électronique*, ou parfois *contrôle d'exposition*.

### 2.1.2. Caractéristiques des capteurs linéaires

Outre les qualités propres à tous les capteurs CCD (faible consommation, haute sensibilité, grande dynamique, absence de marquage par les sources de lumière intense, maîtrise parfaite de la géométrie), les capteurs linéaires présentent des caractéristiques qui les rendent particulièrement intéressants dans nombre d'applications:

- **Grande résolution horizontale.** La structure très simple des capteurs linéaires permet d'intégrer un grand nombre de photosites, en augmentant la largeur de la barrette. La résolution horizontale typique d'un capteur linéaire est de 2048 points, contre 500 points pour les capteurs matriciels courants, soit un gain d'un facteur 4 en moyenne. Il existe sur le marché des capteurs linéaires de 6000 points et plus.
- **Faible quantité d'information.** Un capteur linéaire typique délivre 2048 pixels sur sa ligne unique. En comparaison, un capteur matriciel d'excellente qualité délivre 582 lignes de 750 points, soit 440.000 pixels. Les caméras linéaires délivrent donc environ 200 fois moins d'information à traiter, tout en offrant une résolution horizontale supérieure.
- **Compatibilité avec les optiques photographiques.** Grâce à la grande largeur du capteur — généralement comprise entre 20 et 28 mm, on peut utiliser des objectifs photographiques dont les caractéristiques de résolution, de distorsion et de vignettage sont très supérieures à celles des objectifs vidéo. La grande résolution des objectifs photographiques est indispensable pour tirer partie de celle des capteurs linéaires. Leur faible distorsion (typiquement 0,4% ou moins) peut généralement être négligée, ce qui évite les coûteux calculs de correction informatique en

ligne lors de l'exploitation. Enfin, leur faible vignelage peut éviter des problèmes lors de la mise en correspondance des images en stéréovision.

Les caméras vidéo, quant à elles, utilisent généralement des objectifs destinés à la surveillance, de qualité très médiocre. Quant aux objectifs destinés aux caméras de télévision, ce sont tous des zooms de grande amplitude, dont l'utilisation en stéréovision pose de nombreux problèmes de calibrage et est à proscrire dans la pratique. Le manque d'optiques de bonne qualité, exploitables en traitement d'image, réduit encore les performances des caméras matricielles en vision artificielle.

- **Souplesse des paramètres de prise de vue.** Les signaux de commande d'une caméra linéaire ne sont pas liés à une norme rigide comme dans le cas des caméras vidéo. L'opérateur peut ajuster librement le temps d'exposition et l'intervalle entre deux images successives. Il peut aussi synchroniser chaque prise de vue sur un événement extérieur, même survenant 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 (jusqu'à 20 Mpixels/s).

---

## 2.2. Les caméras linéaires

### 2.2.1. Constitution d'une caméra linéaire

Une caméra linéaire est constituée:

- d'un capteur linéaire (cf. § 2.1),
- d'un objectif (généralement un objectif photographique, cf. § 2.1.2),
- d'une électronique de synchronisation et de conditionnement du signal de sortie du capteur.

Ces éléments, ainsi que leurs interconnexions, sont très proches de leurs équivalents dans les caméras matricielles. Les différences concernent essentiellement les liaisons électriques. En effet, il n'existe aucune normalisation des signaux de commande, qui varient selon la marque et, souvent aussi, selon le modèle de la caméra. De façon générale, ils sont transmis sous forme d'impulsions binaires sur plusieurs liaisons RS-422 séparées. Ces signaux sont les suivants:

- **Horloge de transmission des pixels.** Cette horloge contrôle la fréquence de transmission des pixels successifs de la caméra vers la carte d'acquisition. A chaque période d'horloge, un nouveau pixel est transmis. En général, cette horloge synchronise également les autres signaux de commande.
- **Commande de transmission des pixels:** une impulsion commande le transfert dans les registres à décalage des charges accumulées dans les photosites, puis leur transmission vers la sortie vidéo en synchronisme avec l'horloge de transmission des pixels.

- **Commande d'exposition:** une impulsion commande l'obturateur électronique décrit au § 2.1.1. Le temps d'exposition est égal au temps qui sépare la fin de cette impulsion et le début de l'impulsion de commande de transmission des pixels. Ceci permet d'ajuster le temps d'exposition en fonction des conditions d'éclairage.
- **Sortie vidéo:** cette sortie, qui permet de transmettre les images acquises par la caméra, peut exister sous plusieurs formes. Elle peut consister en une sortie analogique unique, transmettant alternativement les pixels issus des registres à décalage pair et impair. Elle peut aussi fournir simultanément, sur deux connecteurs séparés, les pixels pairs et impairs, ce qui autorise des débits plus élevés. Enfin, il existe des sorties vidéo numériques parallèles, qui délivrent un signal numérisé à l'intérieur même de la caméra. Cette dernière solution évite les parasites de transmission dans des environnements fortement bruités, mais au prix d'une complexité et d'un coût accrus.

Notons que le mot "vidéo" ne signifie pas ici que le signal issu de cette sortie peut être visualisé sur un moniteur vidéo — il n'est pas conforme aux normes de balayage — mais simplement que son contenu représente l'image acquise.

Par ailleurs, aucune caméra linéaire ne peut fonctionner sans un ordinateur équipé d'une carte spécialisée (*carte d'acquisition linéaire*). Toutes les fonctions de la caméra, y compris la visualisation vidéo et la synchronisation sur une source externe (genlock), sont accessibles uniquement par le biais de la carte d'acquisition et de son logiciel de gestion.

### 2.2.2. Champ de visée d'une caméra linéaire

Le champ de visée se réduit à un plan, nommé *plan de visée*, que nous supposons horizontal pour plus de clarté. En conséquence, la notion de verticale est absente, ce qui ne constitue pas toujours un inconvénient et peut même être avantageux dans certaines applications (cf. § 2.2.3).

Les caractéristiques géométriques du plan de visée sont les suivantes (cf. Figure 2.8):

- Ouverture angulaire du champ de visée: .....  $\alpha = 2 \arctan \frac{l}{2f}$
- La largeur du champ à une distance  $D$  : .....  $L = \frac{l f}{D}$
- Largeur couverte par un pixel à une distance  $D$  : ..  $P = \frac{p f}{D}$

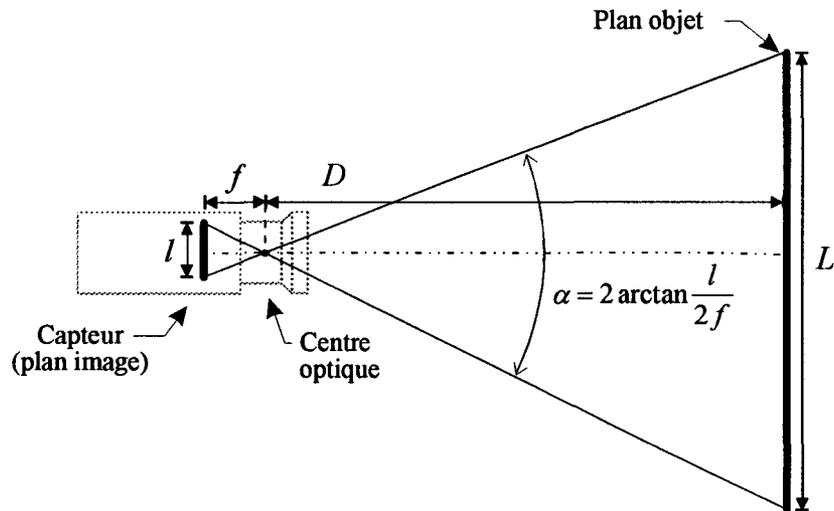


Figure 2.8 : Champ de visée d'une caméra linéaire.

### 2.2.3. Exemples d'applications des caméras linéaires

On pourrait croire *a priori* que l'image linéaire n'apporte rien par rapport à l'image matricielle, et que son intérêt se limite à des cas exceptionnels. Voici au contraire quelques applications très courantes, dans lesquelles les caméras linéaires remplacent avantageusement des caméras matricielles plus classiques.

#### Reconstitution de la dimension verticale par déplacement mécanique

Les capteurs linéaires sont très appréciés dans des applications pour lesquelles la résolution des caméras matricielles est trop faible pour être utilisable. Un dispositif de balayage mécanique permet alors de reconstituer la dimension manquante, au prix d'une certaine lenteur. Les paragraphes qui suivent passent en revue certaines de ces applications.

#### **Numérisation de documents**

Les capteurs linéaires sont la pièce essentielle des *scanners*, dans lesquels un document est filmé avec une très grande résolution pour être intégré en tant qu'image dans la composition d'un document, par exemple en PAO (publication assistée par ordinateur). L'information verticale (c'est-à-dire dans la direction normale au plan de visée) est reconstituée par un balayage mécanique qui est assuré en général par le déplacement du capteur (cf. Figure 2.9).

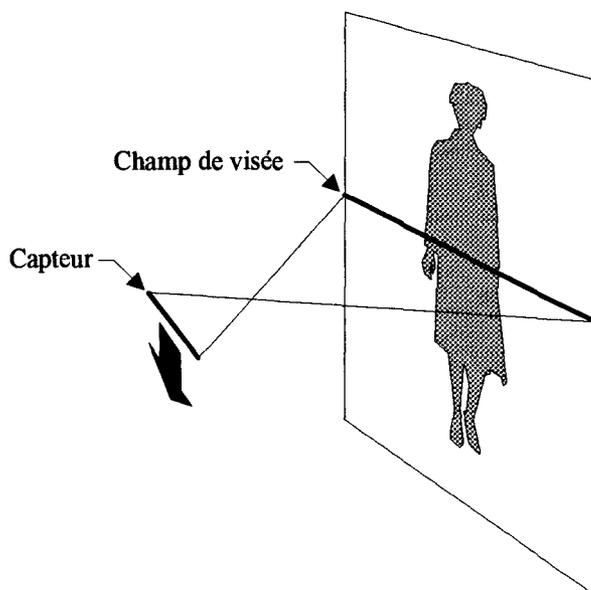


Figure 2.9 : Balayage de l'image dans un scanner.

### Contrôle de qualité

Dans l'industrie, le contrôle de qualité fait souvent appel aux capteurs linéaires, notamment dans la fabrication de tissus, de tôles, et plus généralement de tout produit plan fabriqué en continu. Dans ce cas, le capteur est perpendiculaire au sens de défilement du produit: c'est le mouvement de ce dernier qui permet de reconstituer la dimension manquante et d'obtenir une image complète (cf. Figure 2.10).

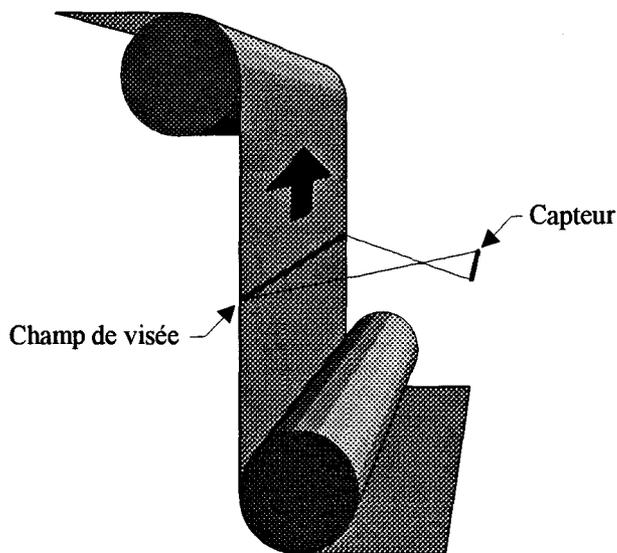


Figure 2.10 : Balayage de l'image en contrôle qualité.

### Téledétection

Les satellites d'observation terrestre Spot utilisent comme capteur d'image des capteurs linéaires de grande résolution, plus fiables et plus rapides que les radiomètres à balayage mécanique qui équipaient les satellites plus anciens tels que les Landsat [Couzy 81]. C'est le mouvement orbital du satellite qui fournit

le balayage (cf. Figure 2.11). La société Fairchild a étudié, pour le compte de l'armée américaine, un dispositif similaire embarqué sur des avions [Dyck 89].

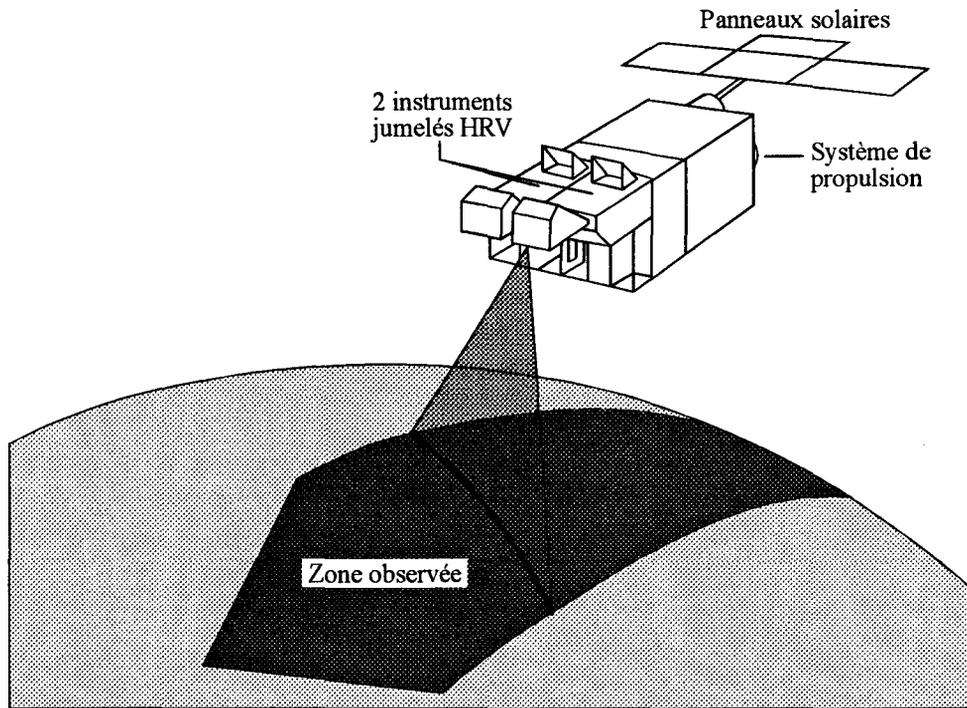


Figure 2.11 : Balayage de l'image en imagerie satellitaire.

## Applications dans lesquelles la verticale est inutile

### Métrie

Dans ce domaine, l'intérêt de la vision artificielle linéaire réside dans sa précision ( $\pm 0,05\%$  dans le cas d'un capteur 2048 points), mais surtout dans l'absence de tout contact mécanique susceptible de perturber la mesure. Un exemple de ce type d'application est le suivi des déformations d'une éprouvette soumise à des contraintes mécaniques [Duvieubourg 93]. La caméra vise la longueur de l'éprouvette, sur laquelle ont été tracés deux repères dont on mesure l'écartement afin de suivre leur évolution au fur et à mesure que la déformation augmente. Cette mesure se fait simplement en comptant les pixels sombres séparant les traits clairs sur l'image (cf. Figure 2.12).

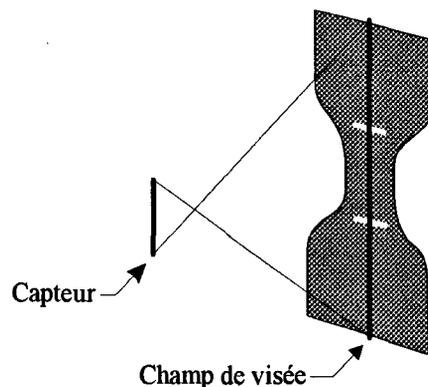


Figure 2.12 : Mesure de dimension par caméra linéaire.

### Spectroscopie

Dans le domaine de la spectroscopie, un capteur linéaire saisit l'image du spectre et la fournit à l'ordinateur qui en assure l'interprétation. Des capteurs spécifiques ont été développés pour cette application, qui exige une grande résolution, une absence totale de distorsion géométrique, une grande sensibilité, une réponse spectrale très étendue, et une excellente uniformité de la réponse des différents photosites [Buil 89].

### Navigation en robotique mobile

En robotique mobile, certains auteurs font appel à des caméras linéaires pour percevoir l'environnement d'un robot. A cet effet, deux caméras sont montées en couple stéréoscopique pour pouvoir déterminer la distance et la position des objets qui constituent l'environnement du robot.

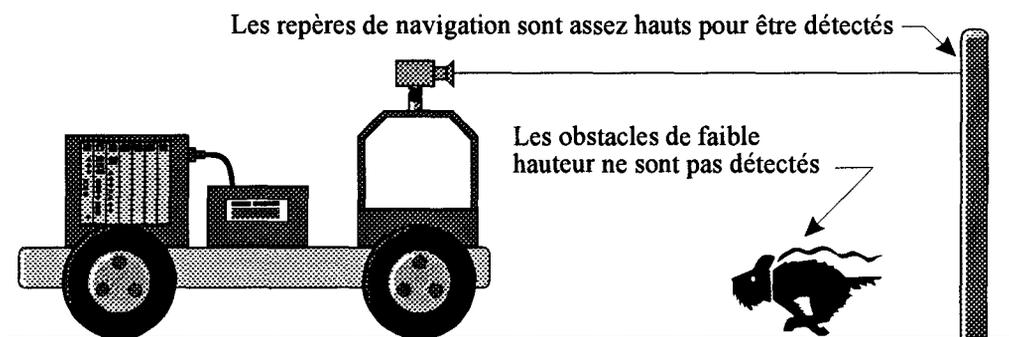


Figure 2.13 : Système de guidage de robot mobile à caméras horizontales [Colle 90].

[Colle 90] propose un dispositif de stéréovision linéaire destiné à guider un robot mobile dans un environnement simple, constitué d'objets aux contours verticaux (cf. Figure 2.13). Les caméras visent à l'horizontale, ce qui leur permet de voir à n'importe quelle distance des repères placés à leur hauteur. Mais cette disposition les empêche de voir les objets de faible hauteur, et interdit donc la détection d'obstacles. De plus, les hypothèses simplificatrices utilisées (les objets présents dans la scène ont des arêtes verticales [Kriegman 89]) limitent strictement l'emploi de ce système à certaines scènes d'intérieur, et le rendent inexploitable dans le cas de l'environnement routier.

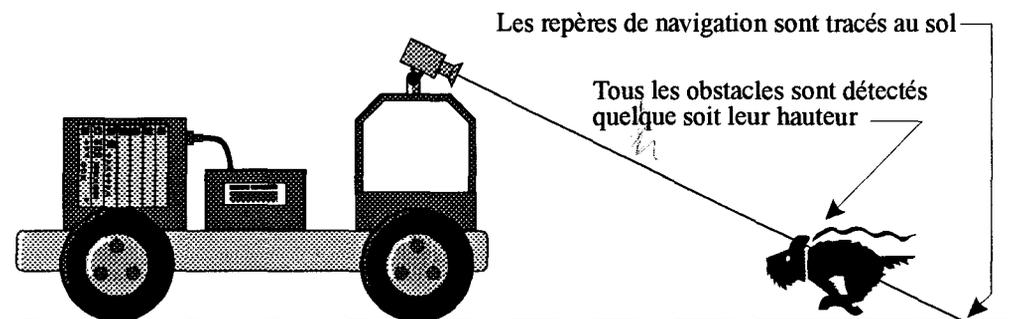


Figure 2.14 : Système de guidage de robot mobile à caméras inclinées [Inigo 85, Inigo 87].

[Inigo 85, Inigo 87] utilise un système stéréoscopique linéaire incliné vers le sol pour guider un robot mobile le long d'une bande de peinture tracée au sol, tout en détectant les obstacles placés sur son chemin (cf. Figure 2.14). Dans ce but, le couple de caméras linéaires est incliné vers le sol, ce qui présente l'avantage

de fournir un fond relativement homogène — et donc des images plus faciles à interpréter — mais limite la portée du dispositif. Comme dans beaucoup d'autres travaux réalisés dans le domaine de la robotique mobile, l'environnement envisagé est très simple, et suppose un fond uniforme, et des objets bien contrastés. Les problèmes causés par les ombres et les occlusions ne sont pas pris en compte, ce qui ne permet pas l'emploi de cette méthode dans un environnement réel tel qu'une route.

Le dispositif que nous décrivons ci-dessous utilise la stéréovision linéaire pour résoudre le problème de la détection d'obstacles sur route. En ceci, il constitue une évolution de [Inigo 85, Inigo 87], dont les caméras ont la même configuration inclinée vers la chaussée. Cependant, l'environnement routier impose des contraintes supplémentaires de précision et de robustesse (formes complexes, fond hétérogène, faible contraste, éclairage variable...), qui interdisent certaines hypothèses simplificatrices courantes en robotique mobile (par exemple [Kriegman 89]: toute l'information utile est formée par les arêtes verticales). Par ailleurs, les méthodes de traitement des images fournies par les caméras doivent être à même d'éliminer les fausses détections, tout en permettant une bonne détection des objets de faible contraste et/ou de petites dimensions.

---

## 2.3. Principe du stéréoscope linéaire embarqué

### 2.3.1. Hypothèses

#### Hypothèses sur les caméras

Notre dispositif de prise de vue n'est pas différent dans son principe de tous les systèmes de prise de vue stéréoscopique à deux caméras. Cependant, la nature particulière des caméras linéaires, qui ne restituent pas l'information verticale, impose une configuration particulière sans laquelle les deux caméras du couple stéréoscopique ne verraient pas la même scène, ce qui rendrait impossible la reconstruction 3D:

#### ***Les plans de visée des deux caméras sont confondus***

Cette première hypothèse signifie que:

- les deux capteurs doivent être dans le prolongement l'un de l'autre (cf. Figure 2.15);
- la ligne contenant les deux capteurs doit être parallèle à la ligne joignant les centres optiques des deux objectifs (cf. Figure 2.16).

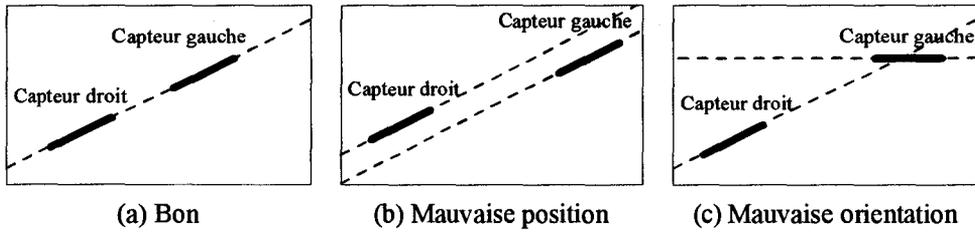


Figure 2.15 : Les capteurs doivent être dans le prolongement l'un de l'autre.

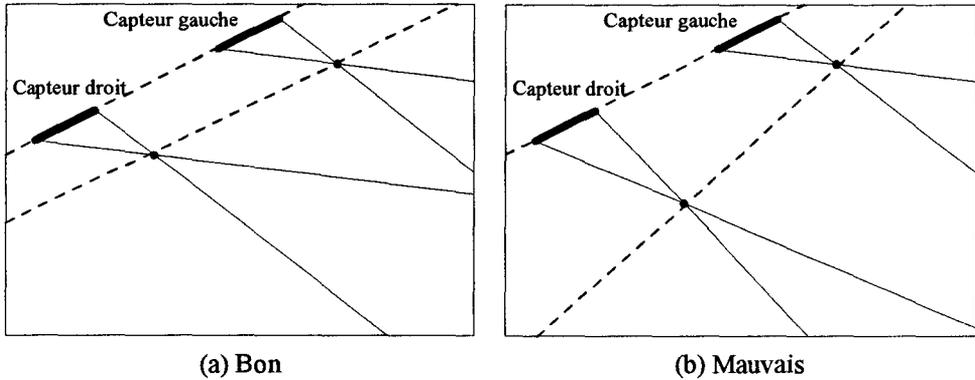


Figure 2.16 : La ligne joignant les deux capteurs doit être parallèle à la ligne joignant les deux centres optiques.

On remarquera que ces deux conditions entraînent également le parallélisme des axes optiques des deux caméras (cf. Figure 2.17). Ce n'est pas absolument indispensable pour permettre la mise en correspondance des images, mais cette configuration simplifie les équations et, surtout, réduit considérablement la quantité de calculs requis pour la reconstruction 3D, ce qui est essentiel pour un système devant fonctionner en temps réel.

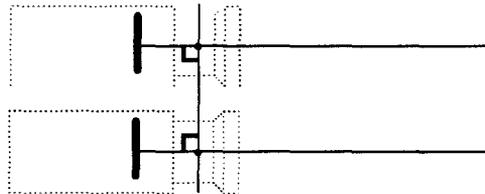


Figure 2.17 : Configuration des axes optiques dans le stéréoscope linéaire.

Nous verrons au Chapitre 3 comment une procédure de calibrage adaptée permet de remplir ces deux conditions. Dans la suite, le plan qui contient les capteurs et les centres optiques des deux objectifs sera appelé *plan de visée du stéréoscope linéaire*.

### Les objectifs sont des sténopés

Il est courant d'assimiler un objectif à un *sténopé* (trou étroit, en grec), c'est à dire un trou de diamètre infiniment petit, confondu avec le centre optique de l'objectif (cf. Figure 2.18). On néglige les effets de diffraction. Le fait d'assimiler un objectif à un sténopé signifie en pratique que:

- Sa distorsion est nulle. Autrement dit, que les points image et les points objet correspondants sont reliés par des lignes droites passant par le centre optique. Dans le cas d'un objectif photographique de bonne qualité, la distorsion est proche de 0,4%, ce qui signifie que

l'erreur de position d'un point objet, proportionnelle à sa distance de l'axe optique, est inférieure à  $(4 \cdot 10^{-3} \times 1024) = 4$  pixels pour un capteur linéaire courant de 2048 pixels. Nous verrons au Chapitre 3 que cette erreur peut être négligée sans sacrifier sensiblement la précision, du moins dans l'application envisagée ici.

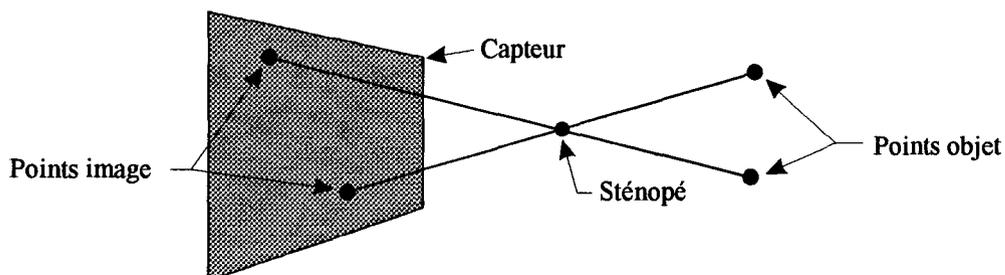


Figure 2.18 : Construction de l'image par un sténopé.

- Sa profondeur de champ est grande devant la profondeur de la scène observée, c'est-à-dire que toute la scène est parfaitement focalisée. Il en résulte que chaque point de la scène est projeté par l'objectif sur un photosite unique du capteur. Dans la pratique, cette condition est remplie si le diaphragme de l'objectif est suffisamment fermé, et si la distance de mise au point est convenablement choisie (cf. Annexe A).

Ces deux hypothèses simplifient considérablement les calculs de projection inverse qui permettent de déterminer la profondeur et la position des points observés dans la scène tridimensionnelle.

### **Les deux objectifs ont la même distance focale**

Il est toujours possible de considérer le cas où les facteurs d'échelle des deux caméras, fonctions des distances focales des deux objectifs, sont différents. Par exemple, [Ayache 89] considère indépendamment les paramètres de calibration de chaque caméra, ce qui permet en théorie de prendre en compte des ensembles de caméras ayant des focales quelconques. Cependant, la complexité des algorithmes utilisés pour la reconstruction de la scène en est notablement augmentée, ce qui entraîne un très fort accroissement des temps de calcul. De plus, l'échantillonnage spatial réalisé par les capteurs CCD peut entraîner des incertitudes très importantes sur la position des points sur les images rectifiées.

Les normes de fabrication autorisent un écart de  $\pm 5\%$  par rapport à la valeur nominale de la distance focale. En réalité, grande série aidant, les constructeurs font généralement beaucoup mieux. En particulier, il est rare que deux objectifs du même modèle, ayant des numéros de série très proches, présentent une différence de distance focale perceptible.

Nous ferons donc en sorte d'utiliser des objectifs de focales très proches, et nous les supposons égales dans la suite.

### **Hypothèses sur la scène**

La reconstruction d'une scène complexe, sans hypothèse particulière sur celle-ci, est un travail difficile. Elle nécessite une grande quantité de calculs qui pénalisent la rapidité d'exécution des procédures et mobilisent beaucoup de

mémoire. Bien que les méthodes existantes permettent d'obtenir des résultats satisfaisants lorsque le temps n'est pas une contrainte, il ne serait pas réaliste d'envisager une réalisation fonctionnant en temps réel sans certaines hypothèses simplificatrices, même avec de gros moyens informatiques.

Heureusement, la chaussée constitue un environnement relativement homogène, sur lequel plusieurs hypothèses peuvent être faites pour réduire la complexité des traitements.

La plus classique de ces hypothèses, qui constitue la base de nombreuses méthodes de détection d'obstacles sur route, est que *la chaussée est plane* [Carlsson 90, Enkelmann 90]. Nous utiliserons cette hypothèse quand nous aborderons la reconstitution de la hauteur des objets détectés sur la chaussée (cf. § 2.3.4).

### 2.3.2. Installation des caméras dans la scène

Les caméras sont installées sur le véhicule, à un emplacement qui sera précisé au Chapitre 5.

Leurs axes optiques sont parallèles, écartés d'une distance  $E$ . Ces deux axes définissent un *plan de visée*, dans lequel visent les caméras. Le parallélisme des axes optiques n'est pas indispensable pour la reconstruction, mais cette configuration simplifie les équations et, surtout, réduit considérablement la quantité de calculs requis pour la reconstruction 3D, ce qui est essentiel pour un système devant fonctionner en temps réel.

Le plan de visée est incliné par rapport à l'horizontale d'un angle négatif  $\theta$  (cf. Figure 2.19), de façon à rencontrer la chaussée à une distance  $D_{max}$  à l'avant des caméras, qui définira la portée du dispositif (ceci en supposant la chaussée plane, comme on l'a vu plus haut). Cette disposition permet de s'assurer que même des obstacles de faible hauteur pourront être détectés.

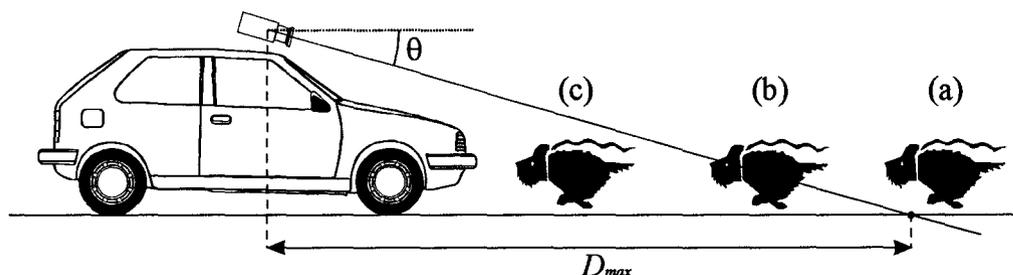


Figure 2.19 : Vision d'un objet mobile par le stéréoscope linéaire.

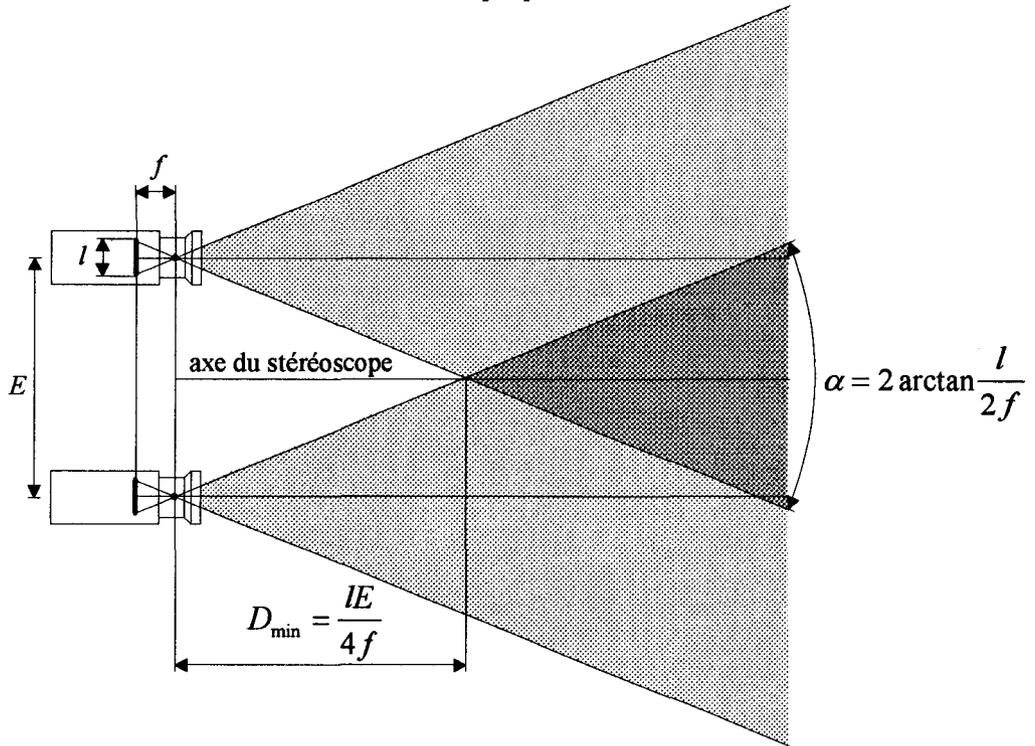
La partie commune aux champs de visée des caméras gauche et droite est appelée *secteur de stéréovision*. Dans ce secteur, les objets sont vus simultanément par les deux caméras, et on peut donc reconstituer leur profondeur par des méthodes de triangulation. Les caractéristiques géométriques du secteur de stéréovision sont précisées sur la Figure 2.20:

- L'axe du secteur est confondu avec l'axe du stéréoscope.
- Distance minimale au centre optique:.....  $D_{min} = \frac{lE}{4f}$

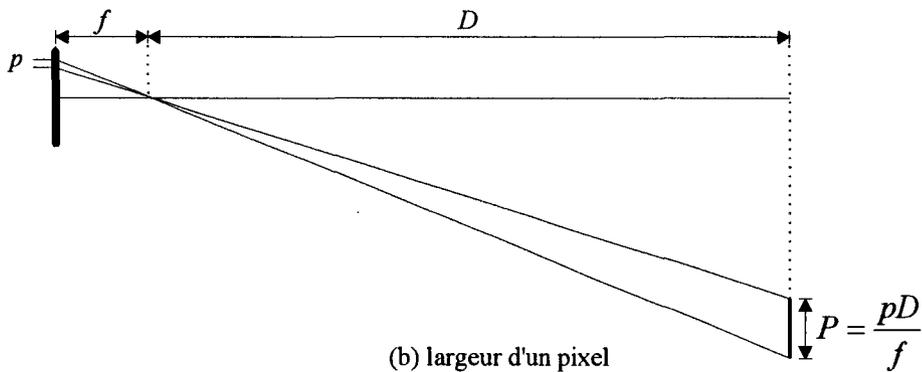
- Ouverture angulaire:.....  $\alpha = 2 \arctan \frac{l}{2f}$
- Champ à une distance  $D$ :.....  $L = \frac{lD}{f}$
- Largeur couverte par un pixel à une distance  $D$ : ...  $P = \frac{pD}{f}$

où:

- $l$  est la largeur du capteur CCD,
- $N$  est le nombre de pixels du capteur,
- $p$  est la largeur d'un photosite ( $p = \frac{l}{N}$ ),
- $f$  est la distance focale des objectifs,
- $E$  est l'écartement des axes optiques des deux caméras.



(a) secteur de stéréovision



(b) largeur d'un pixel

Figure 2.20 : Géométrie du stéréoscope linéaire.

On notera ici encore que la planéité du champ de visée interdit d'obtenir directement d'information dans le plan vertical. C'est lorsque les caméras — ou l'objet filmé — se déplacent que l'on peut reconstituer la vision d'un objet dans son ensemble. Par exemple, le chien de la Figure 2.19 est d'abord invisible lorsqu'il n'a pas encore atteint le secteur de stéréovision (a). Puis, au fur et à mesure qu'il le traverse, le système détecte ses pattes avant, puis son nez, puis son dos, et enfin sa queue (b). Finalement, le chien redevient invisible pour le système puisque sa hauteur est inférieure à celle des caméras (c).

### 2.3.3. Reconstitution de la profondeur et de la position

Considérons un point objet observé par le stéréoscope linéaire (cf. Figure 2.21). Chaque capteur perçoit de ce point une image qui est sa projection à travers le centre optique de l'objectif associé.

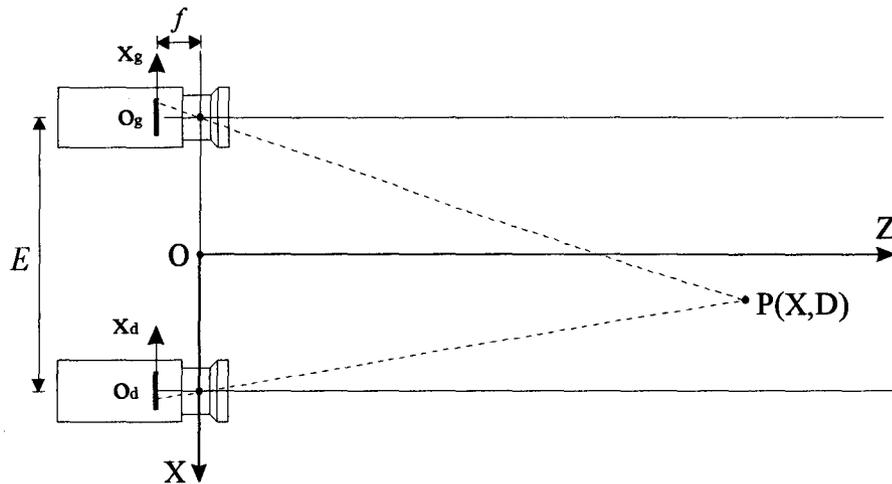


Figure 2.21 : Projection d'une image sur les capteurs du stéréoscope linéaire.

Définissons un repère lié au dispositif de prise de vue:

- L'origine  $O$  est au milieu du segment de droite joignant les centres optiques des deux objectifs.
- L'axe des  $X$  (abscisses) est la droite qui joint les centres optiques des deux objectifs.
- L'axe des  $Z$  (profondeurs) est contenu dans le plan optique, normal à l'axe des  $X$ , et dirigé vers la scène. Il est confondu avec l'axe du stéréoscope.

Soit  $P = \begin{pmatrix} X \\ D \end{pmatrix}$  la position du point objet  $P$  dans ce repère, où  $X$  est l'abscisse de  $P$  selon l'axe  $OX$  et  $D$  sa profondeur selon l'axe  $OZ$ .

Définissons, de plus, un repère local lié à chaque capteur:

- L'origine est au centre du capteur. On la note  $o_g$  pour le capteur gauche,  $o_d$  pour le capteur droit.
- L'axe des abscisses,  $x_g$  ou  $x_d$ , est dirigé vers la gauche du capteur quand on regarde vers la scène.

Soit  $x_g$  la position du point image sur le capteur gauche, et soit  $x_d$  sa position sur le capteur droit.

Exprimons les positions des deux points objets dans leurs repères locaux respectifs:

$$\begin{cases} x_g = \frac{f\left(\frac{E}{2} + X\right)}{D} \\ x_d = \frac{f\left(\frac{E}{2} - X\right)}{D} \end{cases} \quad \text{Equation 2.1}$$

### Notion de disparité

La *disparité* du point  $P$  est la différence de position  $\delta$  des images de  $P$  sur les deux capteurs (cf. Figure 2.21):

$$\delta = |x_g - x_d|$$

Dans le cas présent, elle prend la forme particulière:

$$\delta = \frac{E \cdot f}{D}, \forall X \quad \text{Equation 2.2}$$

où  $f$  est la distance focale des deux objectifs et  $E$  l'entraxe séparant leurs centres optiques.

Ce résultat indique que la disparité est la même pour les images de tous les points situés à une même profondeur, indépendamment de leurs abscisses.

### Calcul de la profondeur à partir de la disparité

La profondeur se déduit directement de la disparité et des paramètres du stéréoscope à partir de la relation précédente:

$$D = \frac{E \cdot f}{\delta} \quad \text{Equation 2.3}$$

### Calcul de la position à partir de la profondeur

Nous avons vu plus haut comment obtenir la position des points objets sur les capteurs gauche et droit (cf. Equation 2.1).

Nous pouvons utiliser ces équations pour déterminer la position  $X$  d'un point objet à partir de sa profondeur  $D$ :

$$\begin{cases} X = \frac{x_g \cdot D}{f} - \frac{E}{2} \\ X = \frac{x_d \cdot D}{f} + \frac{E}{2} \end{cases}$$

Equation 2.4

L'une ou l'autre de ces deux équations permet de calculer la position  $X$  du point objet à partir de sa profondeur  $D$  et de la position du point image correspondant sur le capteur gauche ( $x_g$ ) ou sur le capteur droit ( $x_d$ ).

### 2.3.4. Passage du repère lié au stéréoscope au repère lié à la scène

Les traitements géométriques se font dans un repère lié au capteur. Or, dans le cas général, on cherche à connaître la position des objets observés dans un autre repère (dit *repère de travail*), généralement lié à la scène. C'est pourquoi le changement de repère décrit dans ce paragraphe est nécessaire.

#### Position et orientation des caméras dans la scène

Dans la suite de ce paragraphe, il sera beaucoup question des paramètres de calibration extrinsèques, c'est à dire de la position et de l'orientation des caméras dans le repère lié à la scène. Ces paramètres sont au nombre de six: trois pour la position, et trois pour la rotation. Leurs noms varient selon les auteurs, c'est pourquoi il est nécessaire de définir les termes que nous utiliserons ici.

#### Paramètres de position

La position d'une caméra dans la scène 3D s'exprime, comme celle d'un point, par sa *hauteur* ( $Y$ ), sa *profondeur* ( $Z$ ) et son *abscisse* ( $X$ ) (cf. Figure 2.22). Le repère  $OXYZ$  est un repère orthonormé indirect.

#### Paramètres d'orientation

L'orientation d'une caméra est exprimée en coordonnées polaires dans un repère centré sur le centre optique de son objectif. Les trois axes de rotation sont le *site* ( $\theta$ ), l'*azimut* ( $\psi$ ) et le *roulis* ( $\rho$ ) (cf. Figure 2.22), qui sont définis comme suit:

- L'axe de site est parallèle à l'axe des abscisses. L'angle de site est positif si la caméra vise vers le haut, négatif si elle vise vers le bas.
- L'axe d'azimut est parallèle à l'axe des hauteurs. L'angle de site est positif si la caméra vise vers la gauche, négatif si elle vise vers la droite.
- L'axe de roulis est confondu avec l'axe optique de la caméra. L'angle de site est positif si la caméra penche à gauche, négatif si elle penche à droite.

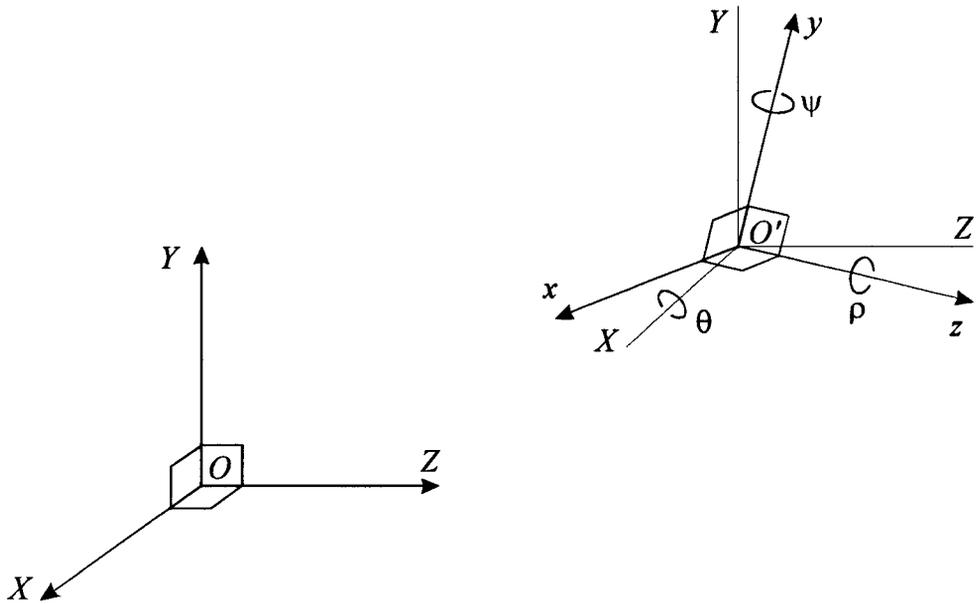


Figure 2.22 : Repères liés à la scène (OXYZ) et à la caméra (O'xyz).

Enfin, l'ordre dans lequel les trois rotations sont effectuées n'est pas indifférent (le changement de repère est généralement calculé comme une multiplication de matrices). Par convention, on effectue d'abord la rotation en site, puis la rotation en azimut, et enfin la rotation en roulis (cf. Figure 2.23).

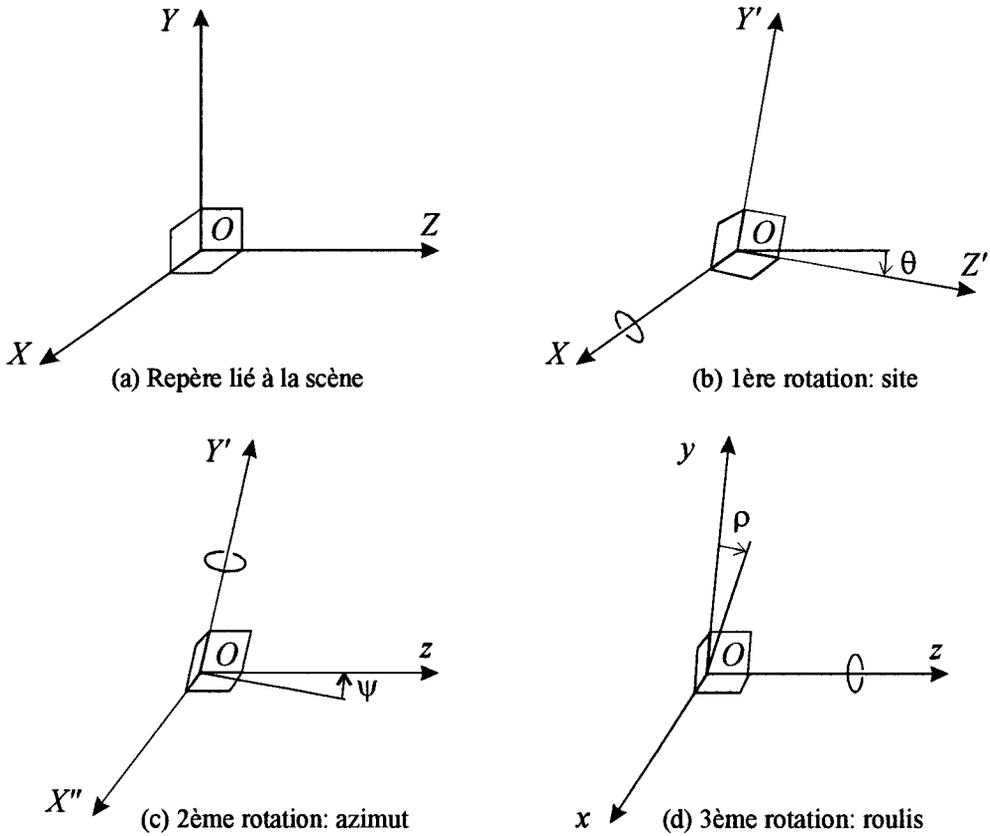


Figure 2.23 : Changement de repère lié à l'orientation de la caméra dans la scène.

### Passage au repère de travail

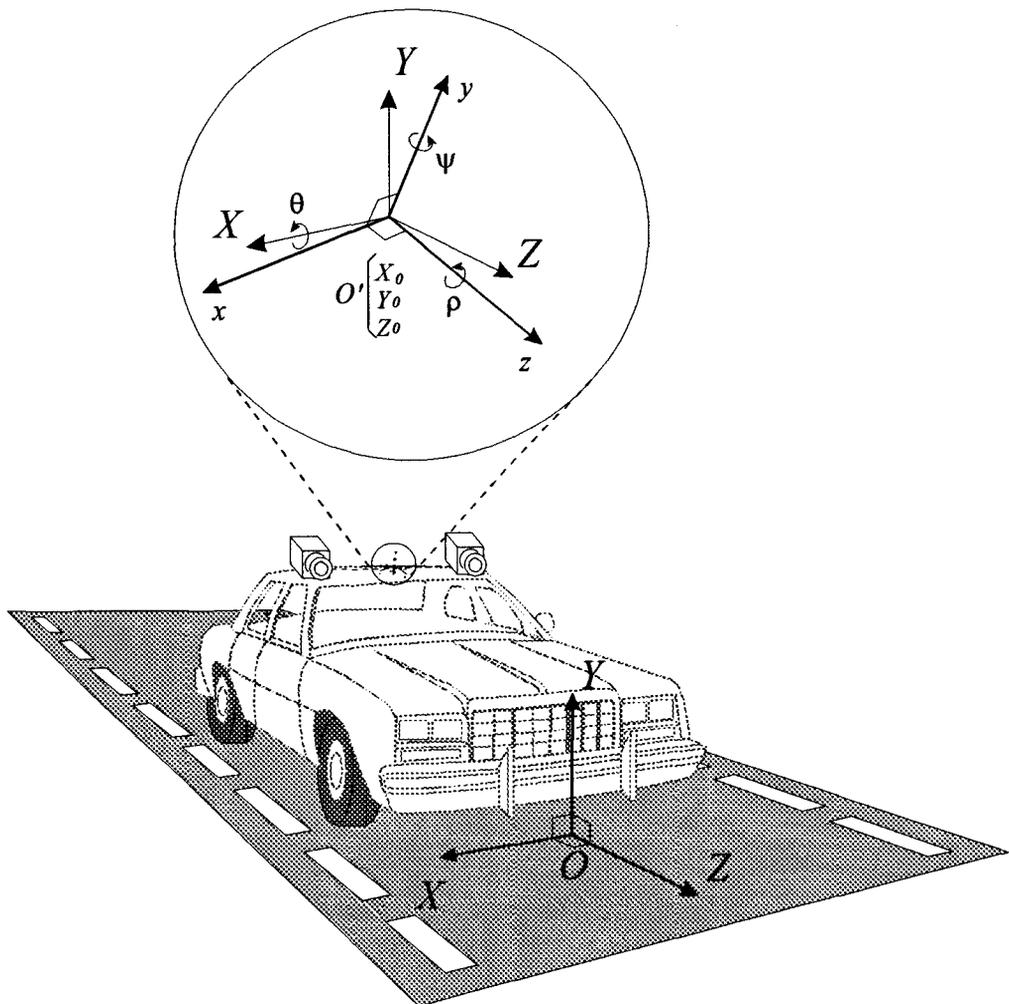


Figure 2.24 : Repère lié au stéréoscope ( $O'xyz$ ) et repère de travail ( $OXYZ$ ).

Dans le cadre du programme ProArt, afin de pouvoir fusionner les résultats en provenance des différents capteurs, le repère de travail est imposé. Il est lié au véhicule et a les caractéristiques suivantes (cf. Figure 2.24):

- L'origine est à la hauteur de la chaussée, à la verticale du milieu du pare-chocs avant.
- L'axe des profondeurs ( $Z$ ) est dirigé dans l'axe longitudinal du véhicule. Il est situé dans le plan de la chaussée (cf. § 2.3.1).
- L'axe vertical ( $Y$ ) est perpendiculaire à la surface de la chaussée supposée plane, dirigé vers le haut.
- L'axe des abscisses ( $X$ ), l'axe des ordonnées ( $Y$ ) et l'axe des profondeurs ( $Z$ ) forment un repère orthonormé.

Dans ce système de coordonnées, le repère lié au stéréoscope linéaire a les caractéristiques suivantes:

Origine  $O'$  située en:  $\begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix}$

Orientation:  $\begin{pmatrix} \theta: \text{angle de site } (\theta < 0) \\ \psi: \text{angle d'azimut} \\ \rho: \text{angle de roulis} \end{pmatrix}$

Le passage du repère lié au stéréoscope au repère de travail se fait par l'équation matricielle suivante (cf. Equation 2.5), dont la justification est donnée dans [Plastock 87]:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \rho \cos \psi & -\cos \rho \sin \psi \cos \theta + \sin \rho \sin \theta & X_0 \\ \sin \psi & \cos \theta \cos \psi & Y_0 \\ -\sin \rho \cos \psi & \sin \rho \sin \psi \cos \theta + \cos \rho \sin \theta & Z_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ z \\ 1 \end{pmatrix}$$

Equation 2.5

On remarquera que  $y$  n'apparaît pas dans cette expression. En effet, puisque les objets observés par le stéréoscope linéaire sont, par définition, dans son plan de visée,  $y$  est toujours égal à 0.

**Détermination de la hauteur des obstacles**

L'Equation 2.5 permet d'exprimer dans le repère de travail tridimensionnel la position d'un point localisé dans le plan de visée. On constate l'apparition de la hauteur  $Y$  du point, notion qui faisait défaut dans le plan de visée.

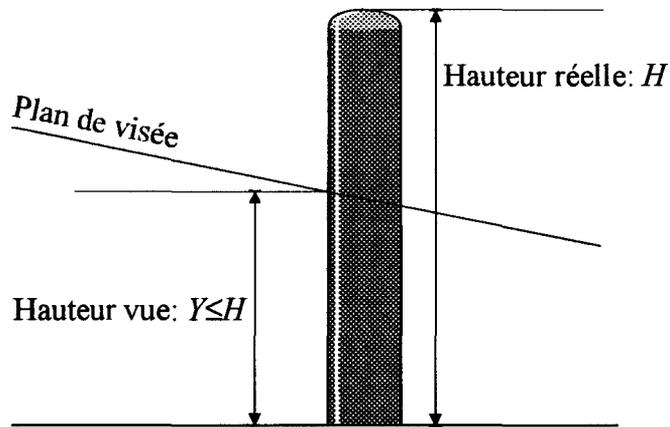


Figure 2.25 : La hauteur vue n'est pas la hauteur réelle de l'objet.

Il est cependant important de remarquer que la hauteur  $Y$  que l'on détermine par l'Equation 2.5 n'est pas la hauteur de l'obstacle, mais seulement celle des points qui en sont vus par le stéréoscope linéaire (cf. Figure 2.25). La hauteur  $H$  de l'obstacle est généralement supérieure à  $Y$ . On verra plus de détails sur ce sujet dans l'Annexe C.

## 2.4. Réalisation électronique du stéréoscope linéaire embarqué

Le stéréoscope est réalisé sous la forme d'un ensemble électronique dont le support est un ordinateur de type PC-AT (cf. Figure 2.26). Ce type de machine, imposé par le cahier des charges Prometheus, présente de nombreux avantages:

- Sa puissance de calcul et son coût peu élevés sont jugés réalistes pour le type d'application envisagée. Il faut cependant signaler l'apparition récente de calculateurs beaucoup plus puissants, basés sur des processeurs RISC, dont le coût baisse rapidement et devrait rejoindre rapidement celui des PC actuels.
- Ses dimensions et sa consommation électrique sont modestes, comparées par exemple à celles des racks VME qui servent souvent de base à ce type de réalisation. L'adaptation à un véhicule automobile en est facilitée.
- Son bus ISA permet la réalisation simple de cartes d'extension pour répondre aux besoins particuliers de l'application. De plus, de nombreuses cartes existent dans le commerce, ce qui permet généralement de faire l'économie du développement de cartes spécifiques.

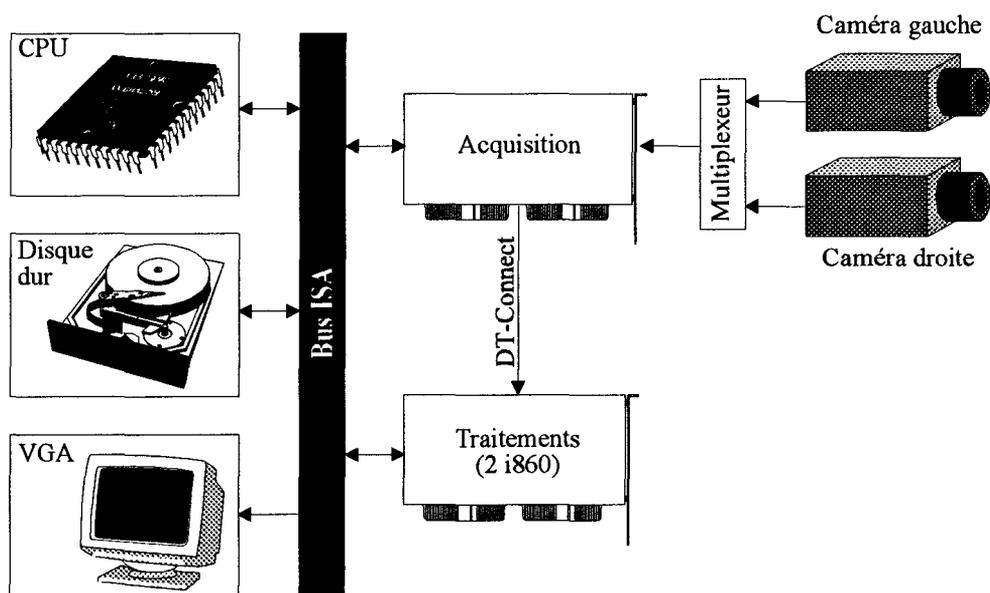


Figure 2.26 : Architecture électronique du stéréoscope linéaire.

### 2.4.1. Caméras

Nous avons choisi le modèle Fairchild CAM 1840N, dont les caractéristiques principales sont les suivantes:

- Dimensions (l x h x p, sans objectif) : 81,3 x 82,5 x 95,9 mm
- Définition: 2592 points.
- Largeur du capteur: 25,9 mm.
- Dimensions des photosites : 10 $\mu$ m x 10 $\mu$ m.
- Dispositif anti-smear.
- Obturateur électronique.
- Sensibilité équivalente: 200 ISO environ.
- Monture d'objectif compatible Nikon AI.
- 1 sortie vidéo (multiplexage des 2 sorties du capteur).
- Signaux de synchronisation: cf. § 2.2.1.
- Alimentation:
  - +5 V, 450 mA (=V<sub>PG</sub>, cf. § 2.1.1)
  - +15 V, 280 mA (=V<sub>SINK</sub>, cf. § 2.1.1)
  - 15 V, 80 mA

Ces tensions sont délivrées par deux alimentations à découpage — une par caméra — enfichées dans le bus ISA du calculateur-hôte.

### 2.4.2. Acquisition

La carte d'acquisition linéaire est une DT2856 de Data Translation (cf. Figure 2.27).

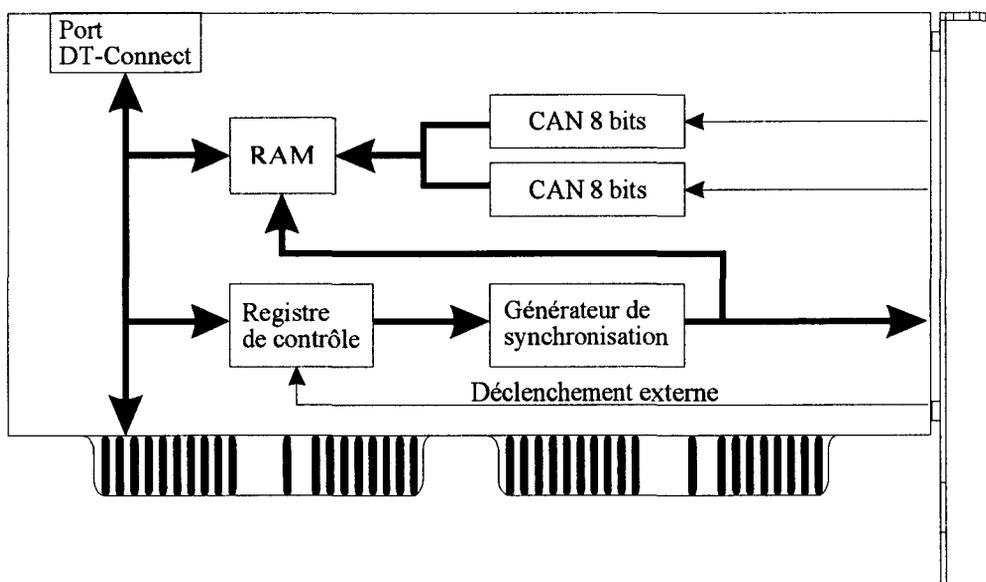


Figure 2.27 : Synoptique simplifié de la carte d'acquisition.

Cette carte, prévue pour une seule caméra, possède deux entrées vidéo pour les caméras disposant de deux sorties différentes pour les pixels pairs et impairs (cf. § 2.2.1). Ces deux entrées sont utilisées pour connecter les sorties de nos deux caméras, les signaux de synchronisation étant modifiés à cet usage par un montage électronique externe. Ceci permet d'économiser le coût, mais aussi les interconnexions et la consommation électrique d'une deuxième carte d'acquisition.

Chaque entrée vidéo est reliée à un convertisseur analogique-numérique d'une résolution de 8 bits.

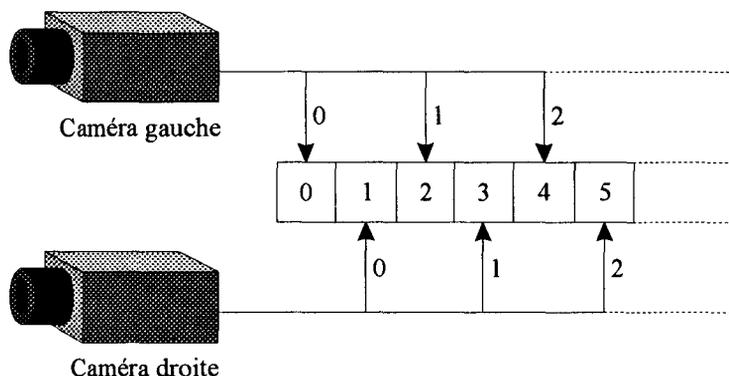


Figure 2.28 : Organisation des pixels en mémoire.

Après chaque prise de vue, la mémoire interne de la carte d'acquisition contient un tableau de  $2 \times 2592$  octets. Les éléments d'indice pair correspondent aux pixels de la caméra de gauche, tandis que les éléments d'indice impair correspondent aux pixels de la caméra de droite (cf. Figure 2.28).

Ces pixels sont transmis vers la carte de traitement via une liaison DT-Connect.

Le bus ISA permet au programme de contrôle, tournant sur le processeur du PC-AT, de sélectionner les paramètres de prise de vue, notamment le temps d'exposition et le temps entre deux acquisitions successives.

### 2.4.3. Traitements informatiques

Les algorithmes de traitement sont implantés sur une carte spécialisée, de type Alacron 860AL-XP, comportant:

- 2 processeurs RISC i860 cadencés à 50 MHz.
- 4 Mo de mémoire RAM partagée par les deux processeurs.
- 1 interface DT-Connect assurant la liaison avec la carte d'acquisition.

Les programmes tournant sur cette carte sont téléchargés par le PC-AT. Les données proviennent du bus ISA (paramètres des traitements) ou de la liaison DT-Connect (images). Les résultats sont envoyés au programme de gestion tournant sur le PC-AT.

Les programmes de la carte de traitement sont développés sur le PC en langage C. Le parallélisme est géré par un mécanisme très simple basé sur l'utilisation de sémaphores.

## 2.5. Réalisation mécanique du stéréoscope linéaire embarqué

### 2.5.1. Contraintes liées au calibrage du stéréoscope

Comme nous l'avons vu au § 2.3.1, la stéréovision linéaire suppose que les plans de visée des différentes caméras soient rigoureusement confondus. Cette contrainte est en particulier essentielle dans le cas où l'objet observé présente une faible extension verticale, ou s'il présente de fortes différences de niveau de gris dans la direction verticale (cf. Figure 2.29). En effet, dans ce cas, la mise en correspondance de ses différentes images deviendrait problématique, voire impossible, si celles-ci ne correspondaient pas à la même "tranche" du sujet, c'est-à-dire à une même section par un plan unique.

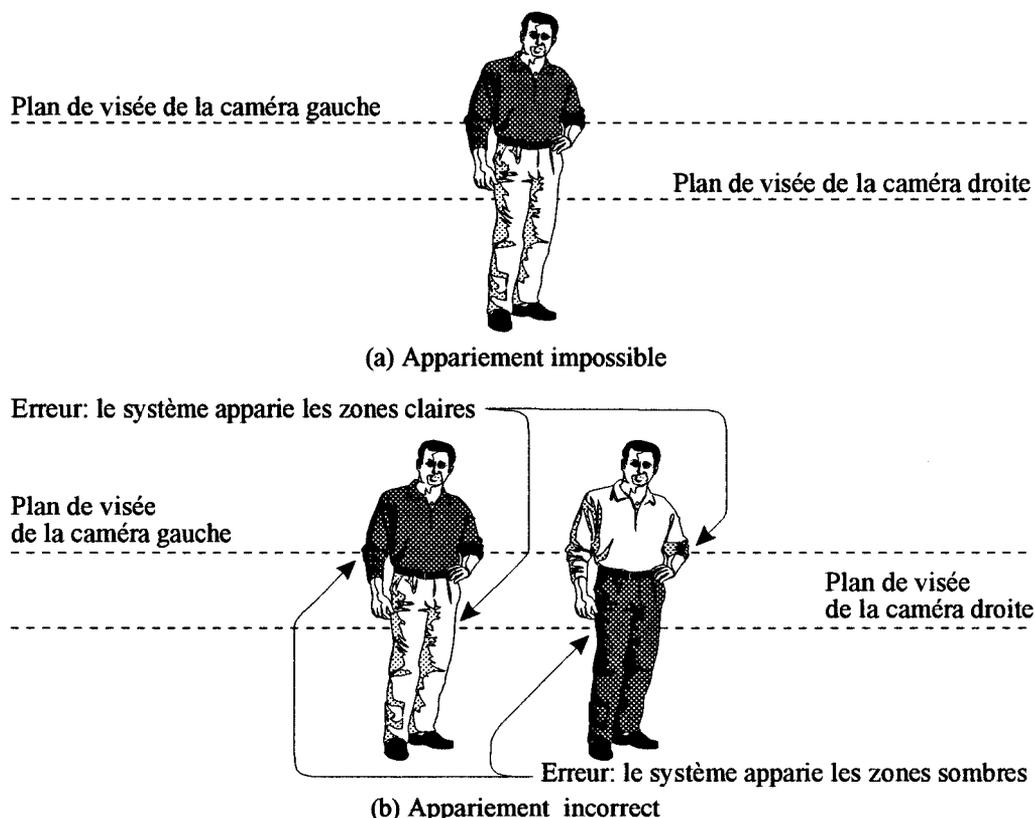


Figure 2.29 : Effet d'un décalibrage affectant la coplanéité des plans de visée.

Dans notre cas, il est donc essentiel que le système mécanique de positionnement et d'orientation des caméras permettent un réglage extrêmement précis de la coplanéité des champs de visée. Ceci implique un ajustement simultané et particulièrement minutieux des angles d'inclinaison et de roulis, et de la hauteur relative des centres optiques des objectifs.

La précision sur l'angle d'azimut est moins critique, du fait qu'un dérèglement dans cet axe ne nuit pas à la coplanéité des deux champs de visée, et peut donc être compensé par une simple translation de l'une des deux images, tout comme dans le cas de la stéréovision matricielle.

## 2.5.2. Contraintes liées à l'environnement

### Contraintes liées à la stéréovision linéaire

Comme on l'a vu au Chapitre 1, tout système de stéréovision impose la connaissance des paramètres géométriques intervenant dans la formation de l'image sur chaque capteur. Dans le cas du stéréoscope linéaire, ces paramètres de calibration sont les coefficients de la matrice définie au § 2.3.4.

L'ajustement mécanique des paramètres de calibration est généralement considéré comme trop difficile par les auteurs, qui préfèrent déterminer les paramètres de calibration de chaque caméra par un étalonnage préalable, sans tenter de les retoucher, puis corriger géométriquement les images prises par chaque caméra (opération appelée *rectification*) avant la mise en correspondance [Ayache 89]. Cette méthode évite tout réglage mécanique, mais ajoute au traitement des images une quantité de calculs importante.

En pratique, pour limiter la perte de précision induite par les paramètres de rotation et l'échantillonnage spatial de l'image, les opérateurs s'efforcent d'ajuster les angles de roulis des caméras pour les rendre aussi proches que possible. Cependant, l'ajustement recherché reste relativement grossier, et ne vise qu'à améliorer la précision sur la position des points observés: les termes de rotation de la matrice de calibration ne sont pas annulés et, par conséquent, la rectification des images reste nécessaire.

Dans le cadre de la détection d'obstacles sur route par stéréovision linéaire, de telles solutions sont inapplicables, et ceci pour deux raisons:

- Nous devons détecter des obstacles **en temps réel**, et les calculs matriciels mis en oeuvre par la rectification des images consomment un temps important qui pénalise cet objectif.
- La nature particulière des images linéaire interdit toute correction en aval de la caméra. En effet, un dérèglement dans l'axe vertical fait que l'image ne contient pas l'information recherchée.

C'est pourquoi nous avons conçu et réalisé un support mécanique spécifique à la stéréovision linéaire, qui permet d'assurer simultanément la coplanéité des champs de visée et la détermination des paramètres de prise de vue en ajustant les orientations des deux caméras. La méthode de calibrage du stéréoscope linéaire, qui sera développée dans le Chapitre 3, utilise les caractéristiques particulières de cette monture pour assurer la coplanéité des champs de visée des deux caméras et ajuster les paramètres de calibration.

### Contraintes de stabilité

La très grande précision requise sur les réglages géométriques du stéréoscope linéaire est généralement jugée incompatible avec un ajustement définitif à la fabrication. En effet, un couple de prise de vue stéréoscopique est généralement un système de dimensions importantes, dans lequel les positionnements et les orientations des caméras peuvent varier de façon non négligeable dans le temps, en raison notamment des contraintes mécaniques auxquelles il est soumis au cours de son utilisation. De ce fait, il peut être nécessaire de le recalibrer de

temps à autre. Dans ce but, notre monture comporte les différents réglages nécessaires.

Comme, dans notre cas, le système de prise de vue doit être embarqué sur un véhicule roulant à grande vitesse sur une chaussée génératrice de vibrations dont l'amplitude et la fréquence sont *a priori* inconnues, sa structure mécanique doit être étudiée pour éviter l'apparition de phénomènes de résonance qui mettraient à mal le calibrage. La partie mécanique de notre prototype est donc largement dimensionnée afin d'augmenter sa rigidité et son inertie et, par conséquent, diminuer le risque de résonance.

### Contraintes de facilité de réglage

Le calibrage d'un dispositif de prise de vue stéréoscopique est connu pour être un travail long et fastidieux, qui constitue certainement l'une des principales difficultés de la stéréovision. La complexité des réglages mécaniques est ici un élément important qui doit être pris en compte.

De plus, l'objectif final de Prometheus étant l'installation du dispositif en série dans les véhicules automobiles, des contraintes supplémentaires apparaissent, liées à l'entretien normal de ces véhicules:

- **L'équipement nécessaire au calibrage doit être de dimensions réduites.** Il est en particulier hors de question d'utiliser une mire tridimensionnelle. En effet, un tel objet remplirait facilement un garage de taille normale, ce qui n'est évidemment pas possible dans la pratique.
- **Le temps requis par le calibrage doit être aussi réduit que possible.** Le prix de la maintenance automobile est, en effet, proportionnel au temps passé. Un temps de calibrage important imposerait un coût de maintenance élevé qui dissuaderait l'acheteur éventuel d'un tel dispositif. La méthode de calibrage décrite au Chapitre 3 répond à cet impératif.

L'accessibilité des réglages joue notamment un rôle très important dans le temps pris par le calibrage. C'est pourquoi nous avons choisi de placer tous les réglages sur le dessus de la monture.

De plus, l'expérience montre que le découplage entre les différents réglages, c'est-à-dire l'absence d'interactions parasites qui modifient les paramètres déjà ajustés lorsque l'opérateur agit sur d'autres paramètres, est un facteur extrêmement important pour la rapidité, mais aussi pour la précision. En effet, l'expérience montre que l'opérateur préfère souvent s'en tenir à un réglage approximatif s'il risque de tout compromettre en tentant de l'améliorer. La monture est donc conçue de façon à obtenir le meilleur découplage possible entre ses différents axes mécaniques.

- **L'équipement et les méthodes de calibrage doivent pouvoir être maîtrisées par un mécanicien.** La plupart des méthodes de calibrage existantes impliquent de la part de l'opérateur des connaissances poussées en mathématiques. Or, nous ne pouvons pas demander ce genre de connaissances à un garagiste. Il est beaucoup plus réaliste

d'adapter l'équipement et les méthodes de calibrage utilisées au savoir-faire courant en mécanique automobile. En particulier, nous ne mettrons pas en jeu de connaissances mathématiques approfondies. En contrepartie, nous pourrions faire appel à l'habileté manuelle de l'opérateur. La méthode de calibrage (cf. Chapitre 3), alliée à la monture que nous décrivons dans les paragraphes suivants, tient compte de cette exigence.

### 2.5.3. Monture orientable d'une caméra

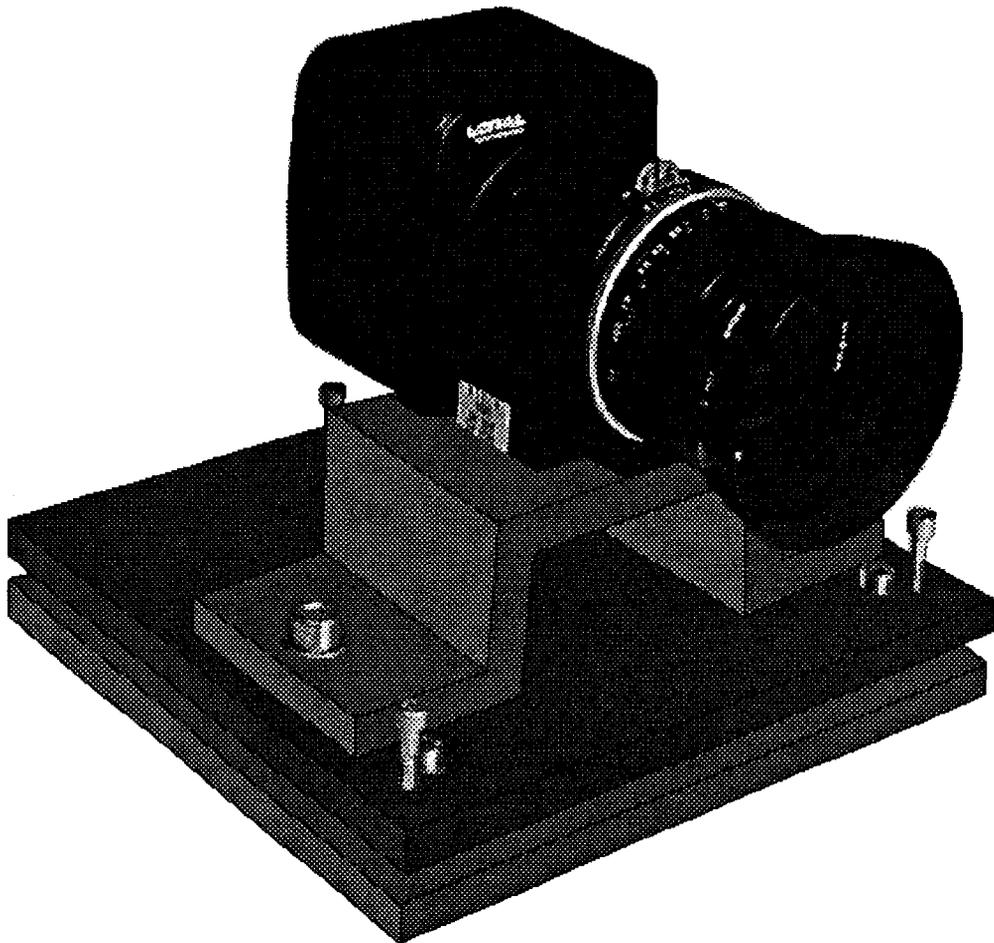


Figure 2.30 : Caméra linéaire sur sa monture ajustable.

Chaque caméra est montée sur un support indépendant (cf. Figure 2.30), orientable avec une grande précision selon les angles de site, d'azimut et de roulis. La hauteur du support est également ajustable. Les caractéristiques essentielles de ce support sont:

- Ajustage par 4 vis.
- Blocage efficace des réglages par l'utilisation de systèmes à vis tirant-poussant. La vis poussant assure le réglage, la vis tirant le bloque.
- Excellente accessibilité des réglages, tous situés sur le dessus.

- Faible couplage des différents réglages. De plus, les relations qui définissent les interactions entre réglages sont connues (cf. Page 59), et prises en compte dans la procédure de calibrage (cf. Chapitre 3).
- Faible hauteur, ce qui améliore la stabilité.

Chaque monture de caméra est constituée de 3 platines métalliques superposées, visibles sur la Figure 2.30: la *platine inférieure* fixe, la *platine supérieure* orientable en site et en roulis, et la *platine azimutale*.

### Platine inférieure et montage sur le support commun

Cette platine sert de base aux réglages des angles d'azimut et de roulis par 3 jeux de vis tirant-poussant. La stabilité du montage est assurée en limitant au minimum la longueur de ces vis, et en leur donnant une section aussi importante que possible.

### Ensemble de vis tirant-poussant

Ce type d'assemblage est très répandu dans les applications qui nécessitent à la fois une grande précision de réglage et une grande rigidité, notamment en optique de précision [Texereau 52]. Un ensemble de réglage à vis tirant-poussant (cf. Figure 2.31) permet de bloquer l'écartement et l'orientation relative de deux plans dans une position donnée sans perturber le réglage. Il se compose de deux platines dont l'écartement  $h$  est ajustable, d'une vis poussant qui assure cet ajustement, et d'une vis tirant qui maintient la vis poussant en contact. Dans le cas où la platine inférieure est réalisée dans un matériau tendre (par exemple en aluminium), on évite son écrasement par la vis poussant en intercalant une touche en matériau plus dur (généralement en acier ou en laiton), qui répartit l'effort sur une surface plus large.

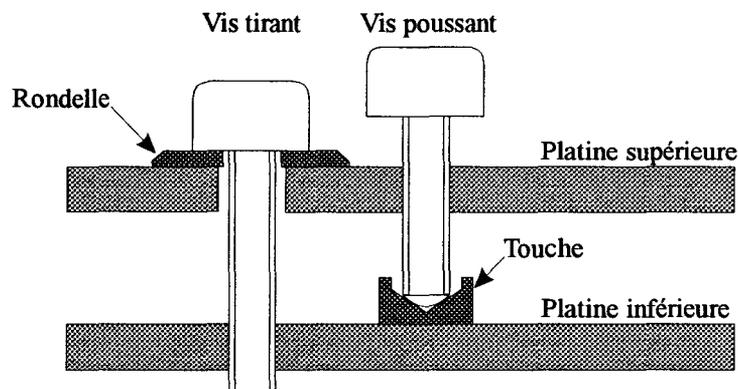


Figure 2.31 : Ensemble de réglage à vis tirant-poussant.

Le réglage est possible lorsque la vis tirant est desserrée. On règle alors l'écartement en tournant la vis poussant. Après avoir resserré la vis tirant, il est possible de déplacer l'assemblage sans perturber le réglage, grâce à l'effort de la vis tirant qui bloque les deux platines l'une contre l'autre.

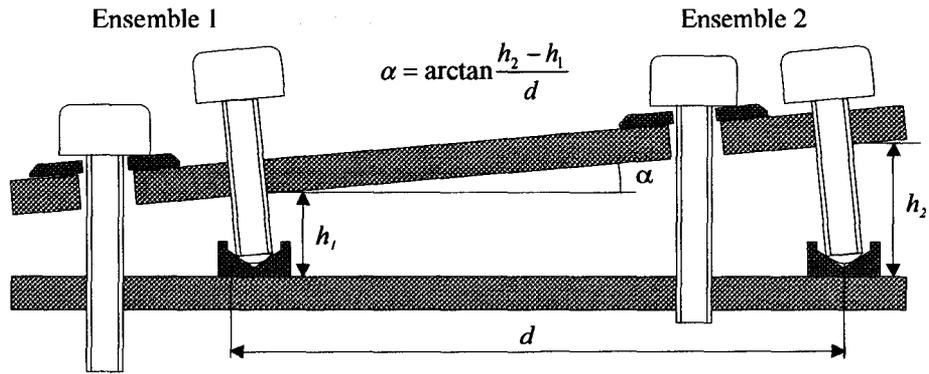


Figure 2.32 : Détermination d'un angle à l'aide d'ensembles de vis tirant-poussant.

En utilisant deux ensembles de réglage, il est possible d'ajuster l'angle que font entre elles les platines inférieure et supérieure (cf. Figure 2.32). Nous utiliserons cette propriété dans la suite.

### Platine supérieure et réglage des angles de site et de roulis

La platine supérieure repose sur la platine de base par l'intermédiaire de 3 systèmes de vis tirant-poussant décrits ci-dessus. Elle comporte notamment 3 trous lisses de passage destinés aux vis tirant, et 3 trous filetés destinés aux vis poussant.

Ces trous sont écartés de  $120^\circ$  sur deux couronnes concentriques, la couronne extérieure portant les vis poussant pour obtenir la meilleure précision de réglage possible (cf. Figure 2.33). Pour ne pas compromettre l'efficacité du réglage, la couronne intérieure portant les vis de blocage est aussi proche que possible de la couronne extérieure.

Pour cette raison, mais aussi pour maximiser le couple de serrage et faciliter l'ajustement, les vis de serrage et de blocage ont une tête CHc à 6 pans creux. La manipulation de ces vis se fait à l'aide de clés Allen, qui permettent d'obtenir un bras de levier important et une grande précision angulaire.

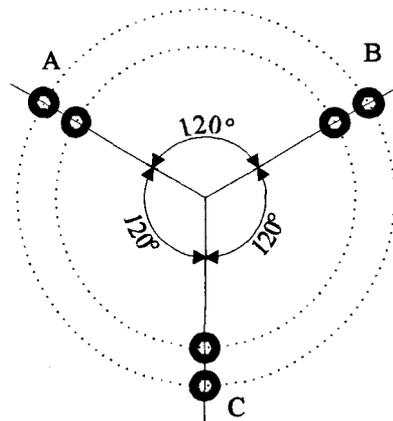


Figure 2.33 : Disposition des ensembles de réglage entre la platine inférieure et la platine supérieure.

La disposition des 3 ensembles de réglage, écartés de  $120^\circ$ , n'est pas arbitraire: la platine supérieure doit reposer sur trois points pour former une structure isostatique, dont la position est complètement définie par rapport à la platine

inférieure. L'angle de  $120^\circ$  entre les 3 vis poussant leur donne une configuration en triangle isocèle, qui assure un facteur d'échelle (exprimé en  $^\circ/\text{tour}$ ) identique sur les 3 réglages. De plus, les efforts sont mieux répartis, ce qui assure une stabilité maximale.

Les trois ensembles de réglage sont repérés par les lettres A, B et C (cf. Figure 2.33). Dans la suite, nous nous référerons à un ensemble de réglage par la lettre associée.

Les angles de site et de roulis sont ajustés par les différences de hauteur des 3 ensembles de réglage (cf. Figure 2.34). Les ensembles de réglage A (à gauche) et B (à droite) agissent essentiellement sur l'angle de roulis. Cependant, leur influence sur l'angle de site n'est pas nulle, et la procédure de calibrage du stéréoscope devra en tenir compte.

L'ensemble de réglage C (au centre) se trouve dans l'axe de la caméra, derrière elle pour éviter que la clé de réglage ne puisse se trouver dans le champ de visée lors du calibrage. Cet ensemble n'agit que sur l'angle de site.

La configuration des trois ensembles de réglage fait apparaître trois axes de rotation, représentés sur la Figure 2.34:

- Une action sur la vis poussant A entraîne un mouvement de rotation autour de l'axe a.
- Une action sur la vis poussant B entraîne un mouvement de rotation autour de l'axe b.
- Une action sur la vis poussant C entraîne un mouvement de rotation autour de l'axe c.

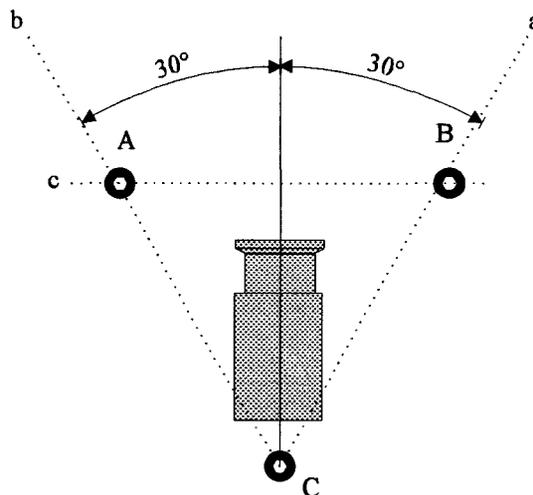


Figure 2.34 : Axes de rotation des ensembles de réglage A, B et C.

Les axes a, b, et c sont écartés de  $120^\circ$ . L'axe c est perpendiculaire à l'axe optique de la caméra, tandis que les axes b et c forment avec ce dernier deux angles égaux de  $30^\circ$  (cf. Figure 2.34). On notera également que le centre optique de l'objectif n'est éloigné de la vis poussant C que de quelques centimètres: étant données les dimensions de la scène (quelques dizaines de mètres), nous pourrions sans inconvénient négliger cette distance et considérer que l'intersection de a et b coïncide avec le centre optique. Cette constatation

sera utilisée dans le raisonnement géométrique qui mène à la procédure de calibrage du stéréoscope (cf. Chapitre 3).

### Platine azimutale et support de caméra

La platine supérieure supporte une troisième platine (cf. Figure 2.35) dont la fonction est de permettre le réglage en azimut de la caméra. Cette troisième platine est montée directement sur la platine supérieure du montage décrit plus haut.

Le dispositif de réglage se compose de deux vis à tête CHc:

- L'une, placée dans un trou rond avec un faible jeu, joue le rôle de centre de rotation.
- L'autre, placée dans un trou oblong, bloque le réglage.

On notera que ce système de réglage en azimut est d'une conception beaucoup plus simple que le système de réglage en site et en roulis. Sa précision est également moins bonne, ce qui n'a pas une grande importance puisqu'un léger décalibrage en azimut peut être facilement compensé (cf. § 2.3.1).

A la suite d'essais, l'amplitude  $\Delta\Psi$  du réglage en azimut a été fixée à la valeur, jugée suffisante dans tous les cas, de  $\pm 2^\circ$ .

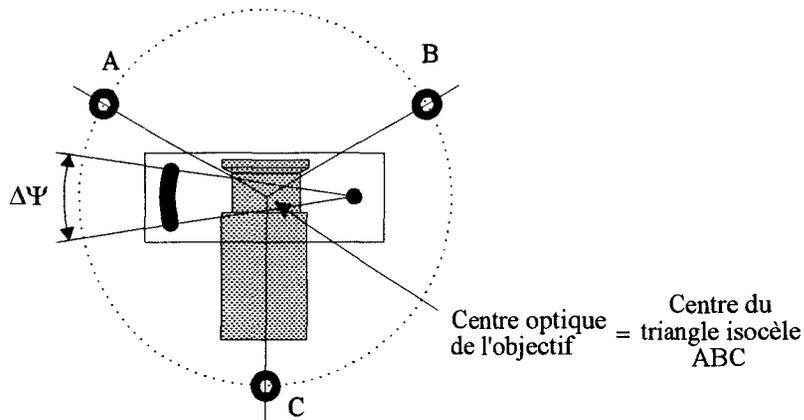


Figure 2.35 : Réglage azimutal de la caméra.

#### 2.5.4. Monture commune aux deux caméras

Les deux montures de caméra décrites ci-dessus ne sont pas montées directement sur le véhicule. En effet, la rigidité de celui-ci est insuffisante pour assurer le maintien du calibrage des caméras durant les prises de vue. Nous avons donc choisi de monter les deux caméras, avec leurs supports, sur une monture commune très rigide qui repose sur le véhicule.

Pour les essais, la monture se compose d'une poutre en profilé d'aluminium, dont les dimensions (90x180 mm de section, 1,20 m de long) et la masse (22 kg) assurent une inertie et une rigidité très importantes à notre stéréoscope linéaire expérimental (cf. Figure 2.36). Nous sommes ainsi assurés que la structure mécanique de celui-ci n'introduira pas de vibrations parasites durant les prises de vue sur route.

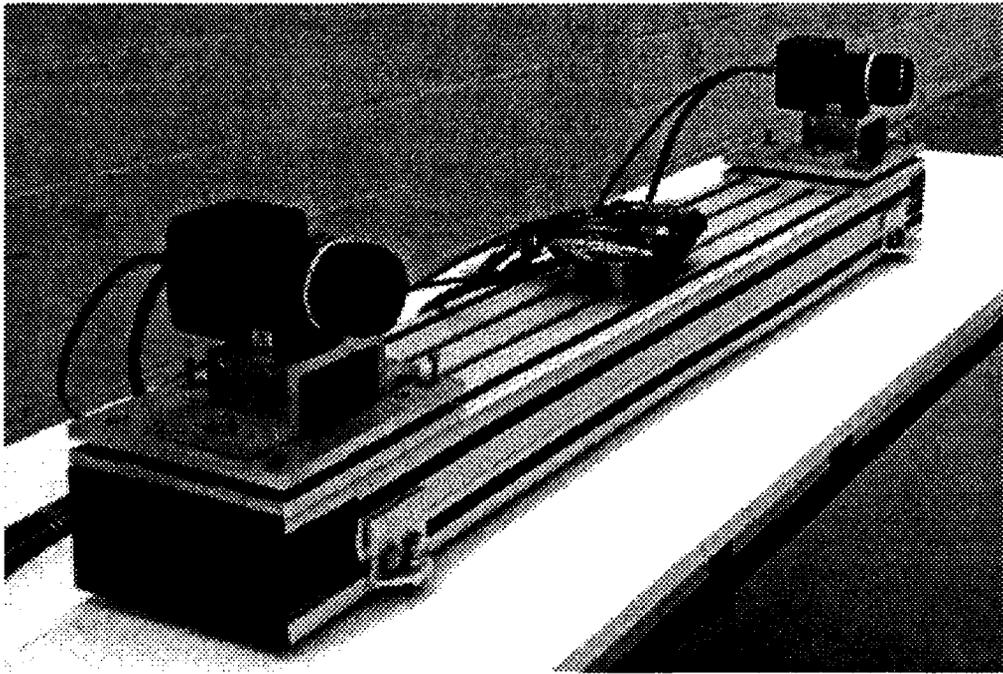


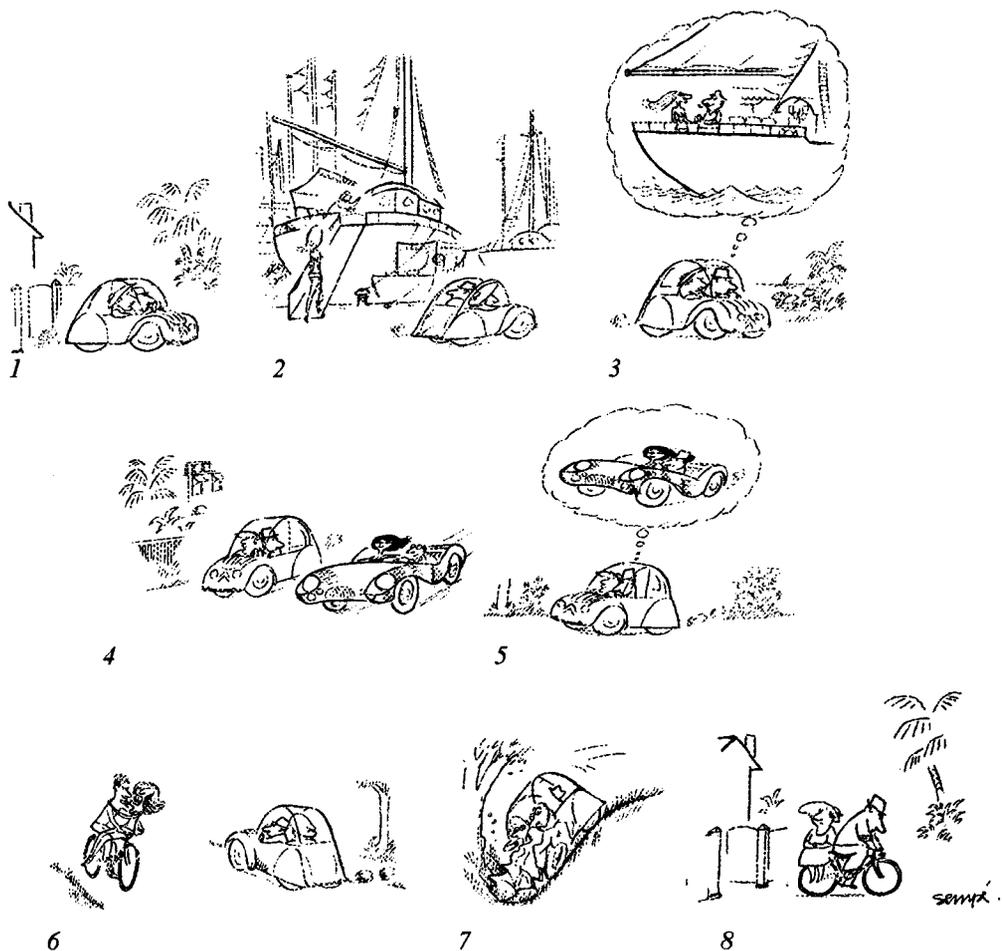
Figure 2.36: Le premier prototype du stéréoscope linéaire sur sa monture.

## 2.6. Conclusion

La détection d'obstacles en temps réel est un problème difficile à résoudre en stéréovision lorsque l'on utilise des caméras matricielles, en raison de la grande quantité d'information fournie par les caméras. Cependant, on a vu que l'utilisation de caméras linéaires, dans une configuration adaptée, permet de diminuer le flux de donnée par un facteur 200 environ, tout en améliorant la résolution horizontale et donc les performances de détection et de localisation des obstacles. Le problème peut donc être traité à l'aide de moyens technologiques actuels.

Nous avons conçu un capteur stéréoscopique linéaire — ou *stéréoscope linéaire* — dans ce but. Nous avons également déterminé les relations qui permettent la reconstruction géométrique de la scène à partir des images fournies par ce capteur.

Cependant, ces relations prennent en compte la géométrie du stéréoscope linéaire, exprimée par un certain nombre de paramètres. Ces *paramètres de calibration* doivent donc être déterminés avant l'utilisation du stéréoscope linéaire, lors d'une phase préalable de calibrage. Cette phase fait l'objet du Chapitre suivant.



## Chapitre 3

# CALIBRAGE DU STEREOSCOPE LINEAIRE

## 3.1. Problème et contraintes

Nous disposons maintenant d'un dispositif qui nous permet d'obtenir des images stéréoscopiques de la scène (cf. Chapitre 2). Nous connaissons également les relations géométriques qui permettent de déterminer la position des objets à partir de leurs images sur les différents capteurs. Nous avons vu que ces relations utilisent des paramètres (appelés *paramètres de calibration*) qui décrivent la position du dispositif de prise de vue dans la scène (*paramètres extrinsèques*) et ses caractéristiques propres (*paramètres intrinsèques*). La détermination de ces paramètres, lors de la phase de calibrage, est un problème souvent délicat à résoudre en vision matricielle, et encore plus dans le cas de la vision linéaire où il a fait l'objet de très peu de publications. Ce chapitre expose la méthode de calibrage que nous avons développée pour notre stéréoscope.

### 3.1.1. Problème à résoudre par le calibrage

#### Cas général - Solutions existantes en vision matricielle

Comme nous l'avons vu, les calculs de projection inverse utilisés pour déterminer la position d'un point dans la scène utilisent des paramètres de calibration, qui expriment les caractéristiques géométriques des caméras (cf. § 1.4). Ces paramètres sont de deux types [Tsai 85]:

- Les *paramètres intrinsèques* constituent des caractéristiques propres aux caméras, qui ne sont pas susceptibles de changer lorsque ces dernières sont déplacées. Il s'agit essentiellement des caractéristiques des capteurs (dimensions, résolution, etc.) et de l'objectif (distance focale, distorsion, etc.).

Notons que, pour simplifier les équations de reconstruction 3D, il est courant de négliger certains paramètres intrinsèques dans les calculs - notamment les distorsions des objectifs - quand la précision de localisation n'en souffre pas de manière trop sensible pour l'application envisagée. Il existe ainsi plusieurs modèles géométriques des caméras, dont le plus simple (mais le moins précis) est le *modèle du sténopé* que nous utilisons dans ce travail (cf. Chapitre 2). Il existe d'autres modèles, dont les plus utilisés sont décrits en détail dans [Zhou 92].

- Les *paramètres extrinsèques* expriment la position et l'orientation des caméras par rapport à la scène. Ils sont indépendants du modèle géométrique de la caméra, mais varient lorsque les caméras se déplacent, ce qui peut poser des problèmes pratiques, en particulier dans les applications de robotique mobile. Dans ce cas, il est généralement possible, après chaque mouvement des caméras, de déterminer les nouveaux paramètres extrinsèques, au prix de temps de calcul importants, en utilisant des mires disposées dans la scène [Kite 90, Wang 91] ou à partir des images de la scène elle-même [Chang 93].

Il est rare que l'on puisse déterminer tous les paramètres de calibration — notamment les paramètres intrinsèques — directement à partir des images de la scène observée pendant l'exploitation normale. Ils doivent donc être déterminés lors d'une phase préparatoire aux prises de vue, appelée phase de *calibrage*, ou d'*étalonnage*, des caméras.

Le calibrage consiste, après avoir mis en place les caméras dans les conditions de prise de vue, à installer dans la scène des objets particuliers constituant une *mire de calibrage* dont les formes, les dimensions, les localisations et les orientations sont connues. Des mesures effectuées sur les images de cette mire permettent de déterminer les paramètres de calibration intrinsèques et extrinsèques des caméras.

La nature de la mire de calibrage utilisée varie selon la méthode de calibrage, qui dépend elle-même du modèle géométrique des caméras, de la précision requise sur la reconstruction de la scène 3D, des dimensions de la scène et des autres contraintes éventuelles de l'application et, parfois, du sens pratique des auteurs:

- Certaines mires comportent de nombreux points de repère répartis dans l'espace 3D. L'installation de ces mires est longue et délicate, la précision requise sur la localisation de chaque point de référence étant élevée. De plus, les paramètres de calibration obtenus n'étant réputés valables que dans les limites de la mire (dans le cas contraire on extrapole les résultats obtenus, avec une précision souvent médiocre [Zhou 92]), les dimensions de celle-ci doivent être du même ordre que celles de la scène. Les problèmes d'encombrement sont évidents dans de nombreux cas, en particulier celui des scènes de très grandes dimensions, telles qu'une rue. La mire de la Figure 3.1, par exemple, a été utilisée pour une application d'archivage des équipements routiers.

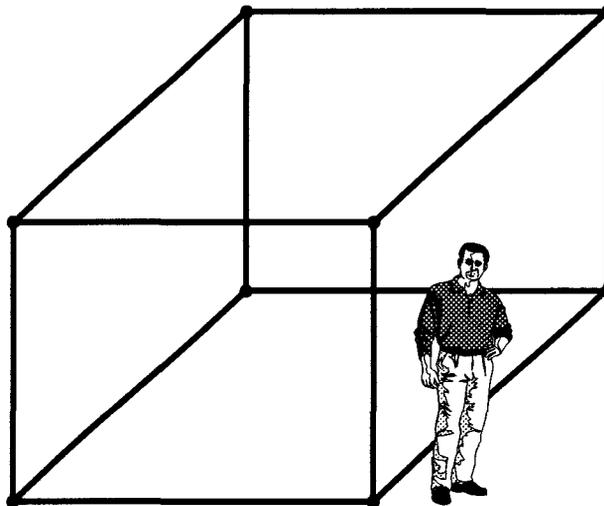


Figure 3.1: Mire utilisée par [Lee 91]; le personnage donne l'échelle.

- D'autres auteurs ont proposé des mires de calibrage planes, dont les avantages pratiques sont évidents: plus grande facilité de construction et d'installation, notamment.

Afin de reconstituer une mire tridimensionnelle, il est courant de placer la mire plane sur un support mobile qui permet une translation dans le sens de la profondeur (cf. Figure 3.2) Ce dispositif de translation doit être très précis pour ne pas affecter la précision du calibrage. [Lenz 88], [Heraud 93-2], et surtout [Martins 81] et [Zhou 92], décrivent des méthodes de calibrage qui exploitent ce montage. Comme dans le cas précédent, la précision décroît très vite lorsque l'on sort des limites de la mire.

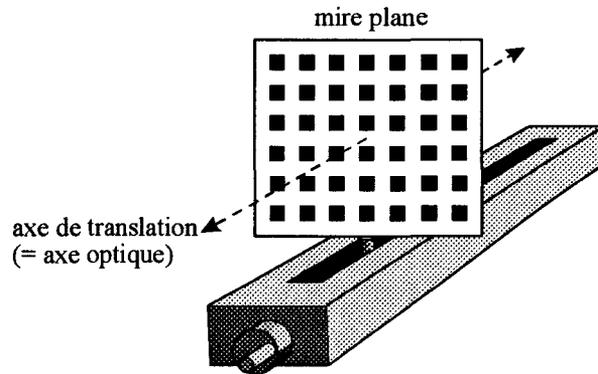


Figure 3.2 : Mire plane sur table de translation (d'après [Heraud 93-2]).

D'autres méthodes font appel à une mire plane fixe. Le calibrage est alors moins précis, mais le dispositif est moins encombrant, et il peut être laissé à demeure dans la scène pour effectuer des recalibrages "au vol" lorsque les caméras bougent. [Wang 91] est un bon exemple simple d'une méthode de ce type, tandis que [Nomura 92] et [Chang 93] fournissent en plus la distorsion de l'objectif. Par ailleurs, le modèle géométrique du sténopé qui est retenu dans ces méthodes semble donner une meilleure précision, hors des limites de la mire, que le modèle des deux plans qui est souvent utilisé avec les mires tridimensionnelles ou à mire plane mobile.

### Cas de la stéréovision linéaire

Le stéréoscope linéaire est composé de caméras linéaires, qui ne peuvent constituer un couple stéréoscopique que si leurs champs de visée sont situés dans le même plan, appelé *plan de visée* du stéréoscope (cf. § 2.2.2).

En effet, le champ de visée vertical des caméras linéaires, limité à une seule ligne, ne permet pas la compensation passive d'un écart de position ou d'orientation de l'une des caméras, par application de transformations géométriques réalisées par des procédures logicielles. Une méthode au moins a été développée pour déterminer les paramètres de calibration d'une caméra linéaire [Heraud 93-1], mais la précision obtenue est très insuffisante pour garantir la coplanéité des plans de visée.

De plus, nous avons vu que la portée du dispositif de détection est conditionnée par l'angle de site du stéréoscope (cf. § 2.3.4). La simple mesure des paramètres de calibration permettrait de déterminer la portée, mais non de l'ajuster à la valeur voulue, ce qui est pourtant indispensable dans notre application.

Ces raisons font que nous ne pouvons nous contenter, comme dans les méthodes de stéréovision matricielle, de déterminer de façon passive les paramètres de position, d'orientation et d'échelle des caméras, pour appliquer ensuite les corrections géométriques obtenues par traitement informatique des images. Nous devons appliquer une méthode active de calibrage, afin d'assurer la coplanéité des champs de visée des deux caméras, et de mettre celles-ci dans une configuration géométrique prédéterminée. Cette approche n'a, à notre connaissance, jamais été utilisée en vision matricielle, et seulement de façon empirique en vision linéaire [Inigo 85, Colle 90]. Aucune méthode de ce type n'a jusqu'ici été décrite dans la littérature.

### 3.1.2. Contraintes

#### Contraintes liées à l'emploi de caméras linéaires

Comme nous l'avons vu, les caméras linéaires ont un champ vertical limité à une ligne (cf. § 2.2.2). Pour cette raison, il n'est possible que de voir ou non un objet donné. Dans le cas où l'on ne le voit pas, il est impossible de savoir si le plan de visée se trouve au-dessus ou en dessous de lui.

La mire de calibrage devra donc fournir le moyen de déterminer la position du plan de visée par rapport aux repères tracés sur la mire. Dans le cas contraire, le calibrage serait irréalisable dans la pratique.

#### Contraintes liées à la calibration des images

La méthode de calibrage doit également tenir compte des paramètres de calibration utilisés dans les calculs de reconstruction géométrique de la scène 3D. Ces paramètres sont la position  $(X, Y, Z)$  et l'orientation  $(\theta, \psi, \rho)$  de la caméra dans la scène, auxquelles s'ajoute le facteur d'échelle de la projection sur les capteurs des caméras, déterminé par la distance focale des objectifs, la largeur des capteurs et la taille de chaque pixel.

L'application envisagée ici exige de fixer ces paramètres (ou, du moins, certains d'entre eux) à des valeurs prédéterminées, afin que les paramètres de prise de vue soient ceux déterminés lors des simulations. Il s'agit donc de réaliser l'opération inverse de l'étalonnage (cf. § 1.4): au lieu de mesurer les caractéristiques d'un système de prise de vue, on lui fait prendre les caractéristiques que l'on souhaite.

Enfin, pour permettre l'exploitation en temps réel du stéréoscope, les calculs de calibration lors de la prise de vue doivent être aussi réduits que possible. On a vu au chapitre 2 que nous négligeons certains paramètres (la distorsion notamment) dont la prise en compte nécessiterait beaucoup de calculs sans pour autant améliorer sensiblement la précision.

### Contraintes liées à l'environnement automobile

Rappelons brièvement les contraintes analysées au § 2.5.2 :

- Les besoins matériels doivent être aussi réduits que possible.
- L'encombrement de l'installation de calibrage doit être limité.
- L'opérateur ne doit pas avoir besoin de connaissances théoriques poussées en mathématiques ou en informatique.
- Le temps nécessaire au calibrage doit être maintenu dans des limites raisonnables.
- Les différents degrés de liberté mécaniques permettant les réglages doivent pouvoir être ajustés indépendamment les uns des autres.

Certaines de ces contraintes ont déjà été prises en compte dans la conception de la monture ajustable décrite au Chapitre 2. Les autres seront prises en compte dans la méthode de calibrage décrite dans ce Chapitre.

## 3.2. Installation pour le calibrage

La mire de calibrage est plane, afin de prendre le moins de place possible (cf. § 3.1.2). Nous pourrions tracer cette mire sur le sol, reconstituant ainsi la scène filmée lors de l'exploitation normale du stéréoscope, mais cette solution ne serait pas très pratique car la distance entre les caméras et la mire serait alors égale à  $D_{max}$ , soit 50 m dans notre cas. Les locaux de cette taille ne sont pas courants, et il est très rarement possible de les libérer pour tracer des mires au sol dans le seul but de calibrer des stéréoscopes! Nous avons donc préféré utiliser une mire verticale, située plus près du stéréoscope à calibrer, et de dimensions plus réduites (cf. Figure 3.3). La conception et les caractéristiques de cette mire sont décrites dans les paragraphes suivants.

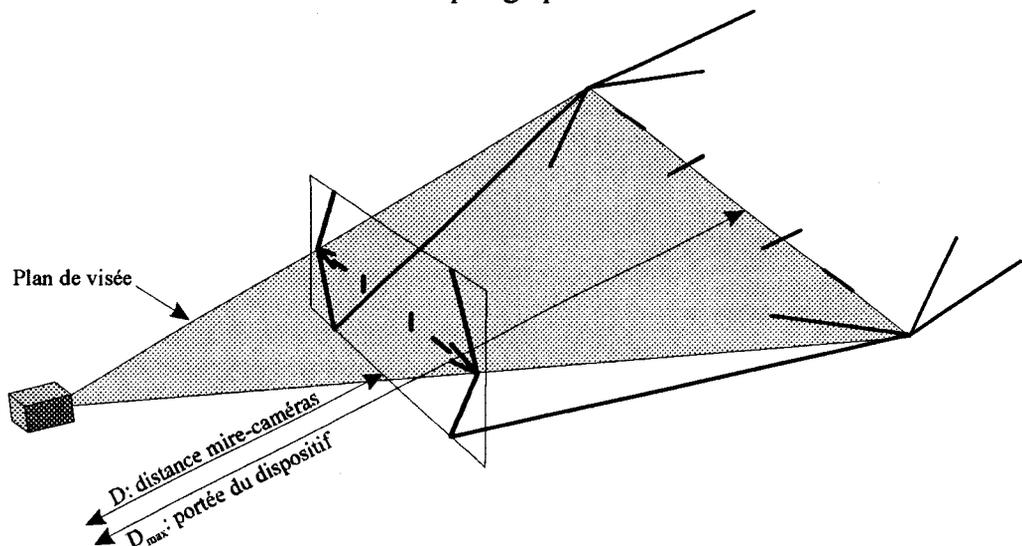


Figure 3.3 : La mire de calibrage telle qu'elle devrait être (horizontale)...  
et telle qu'elle est en pratique (verticale).

L'éclairage de la mire est assuré par deux projecteurs halogène de 500 à 1000 W, à champ large (flood), placés symétriquement de part et d'autre de la mire, qui suffisent à assurer un éclairage d'une intensité et d'une uniformité acceptables. Notons que, si les photos de ce Chapitre montrent des projecteurs de type "mandarine", destinés au cinéma et à la télévision, n'importe quelles lampes halogène feraient aussi bien l'affaire. Le placement idéal de ces projecteurs est celui adopté en général pour la reproduction photographique de documents: chaque projecteur forme un angle de  $45^\circ$  avec l'axe de la mire afin d'éviter toute brillance parasite, la distance entre les projecteurs et la mire devant être déterminée expérimentalement pour assurer un éclairage uniforme (cf. Figure 3.4).

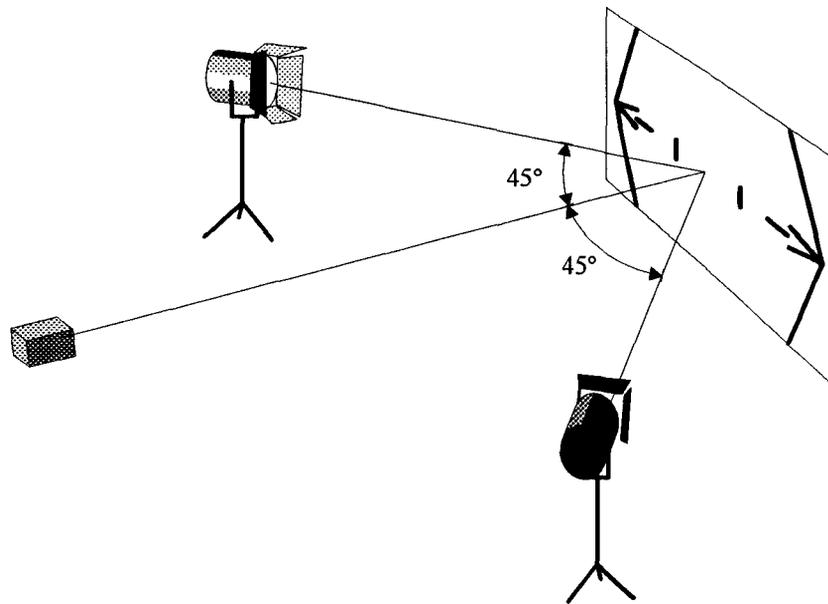


Figure 3.4 : Eclairage de la mire de calibration.

Pour des raisons pratiques et économiques évidentes, l'électronique et l'informatique utilisées pour le calibration sont celles embarquées dans le véhicule pendant l'exploitation normale du dispositif de détection d'obstacles: nous y ajoutons simplement un clavier et un moniteur permettant l'affichage en niveaux de gris des images acquises par les caméras. Ce moniteur, dont la résolution horizontale est inférieure à celle des caméras (typiquement 320 à 800 lignes), peut fonctionner selon deux modes.

- En mode "plein cadre", un pixel est affiché pour  $n$  pixels de la caméra, de façon à visualiser l'ensemble de l'image .
- En mode "gros plan", tous les pixels de la caméra sont affichés, afin de visualiser le maximum de détails sur une portion sélectionnée de l'image sur l'écran. Dans les deux modes d'affichage, un réticule peut être incrusté sur l'écran, afin d'aider l'opérateur pour certaines phases du réglage.

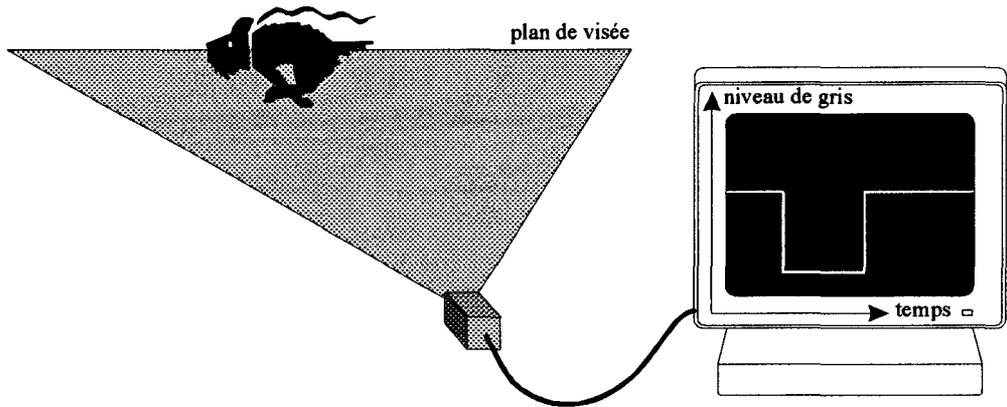


Figure 3.5 : Affichage de type "oscilloscope".

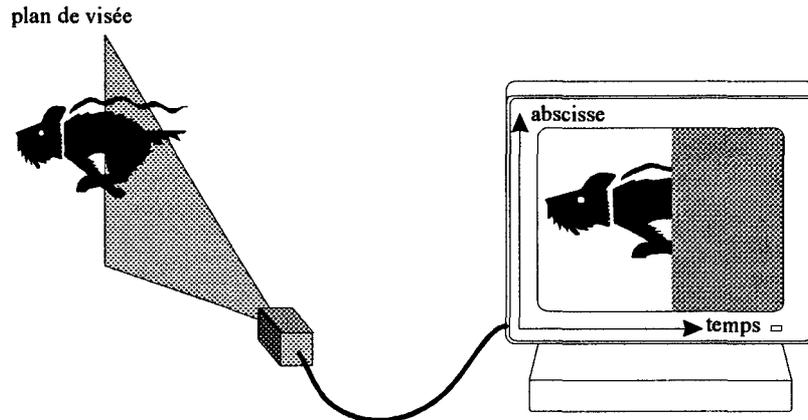


Figure 3.6 : Affichage en balayage horizontal.

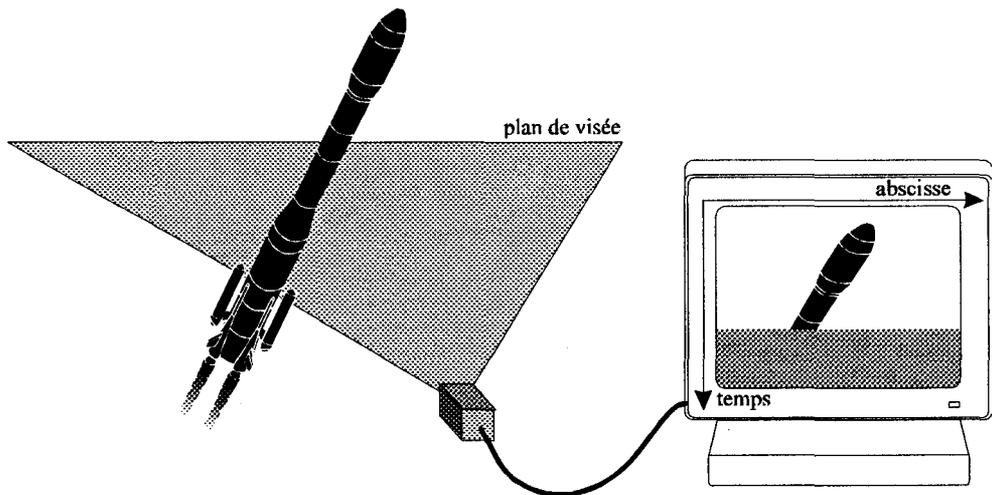


Figure 3.7 : Affichage en balayage vertical.

L'affichage d'une image linéaire (1D) sur un écran 2D peut se faire de différentes manières suivant l'application envisagée: comme sur un oscilloscope, pour avoir des informations très précises sur le contenu de l'image (cf. Figure 3.5), en balayage horizontal, mieux adapté à un plan de visée vertical (cf. Figure 3.6), ou en balayage vertical, mieux adapté à un plan de visée horizontal (cf. Figure 3.7). Le plan de visée de nos caméras étant horizontal, nous avons choisi le balayage vertical, les images linéaires étant affichées comme des lignes horizontales qui viennent s'empiler les unes en dessous des autres (cf. Figure 3.7). Lorsque le bas de l'écran est atteint, le balayage recommence en haut (mode de balayage "roll", en bon français). Ce type de balayage, rappelant le sens de l'écriture, est supposé plus instinctif pour l'opérateur. On visualise ainsi à l'écran l'évolution de l'image dans le temps (représenté par l'axe vertical), ce qui sera utile lors du calibrage.

Le véhicule portant le stéréoscope linéaire est placé devant la mire, d'une manière qui sera précisée au § 3.3.2.

### 3.3. Mire de calibrage

La mire de calibrage est plane, réalisée en noir mat sur un fond blanc diffusant destiné à mieux "accrocher" la lumière. Elle est représentée sur la Figure 3.8.

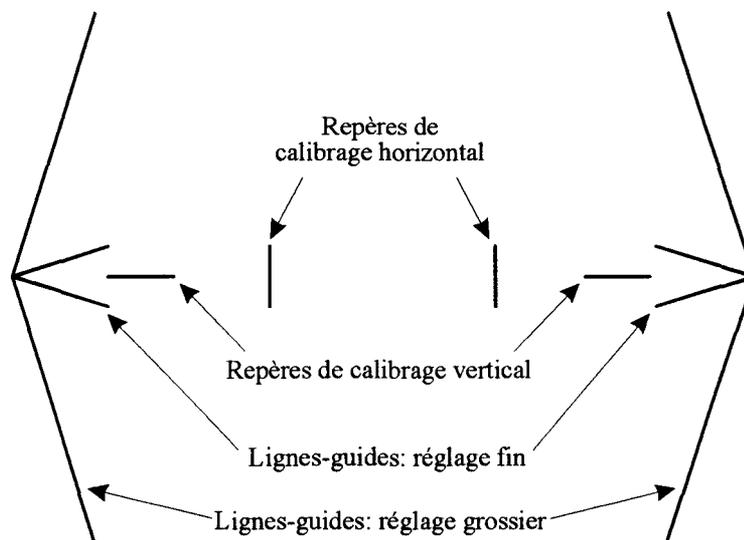


Figure 3.8 : Constitution de la mire de calibrage.

#### 3.3.1. Constitution de la mire

##### Repères de calibrage vertical

Deux segments de droite horizontaux, placés dans le prolongement l'un de l'autre, constituent les *repères de calibrage vertical*. Le plan de visée d'une caméra est correctement ajusté lorsque cette dernière "voit" ces deux repères simultanément. Les deux plans de visée sont coplanaires lorsqu'ils sont tous les deux parfaitement ajustés.

## Repères de calibrage horizontal

Deux segments de droite verticaux, symétriques par rapport au centre de la mire et placés de façon à couper la droite qui porte les repères de calibrage vertical, constituent les *repères de calibrage horizontal*. Ils sont écartés d'une distance égale à la distance qui sépare les centres optiques des objectifs gauche et droit du stéréoscope linéaire. Les axes optiques des deux caméras sont parallèles lorsque l'image du repère gauche est au centre du capteur de la caméra gauche, et que l'image du repère droit est au centre du capteur de la caméra droite, pourvu que les centres optiques des deux caméras sont eux-mêmes sur une lignes parallèle à celle joignant les repères de calibrage vertical. Nous reviendrons sur cette condition au § 3.4.1.

## Lignes-guides

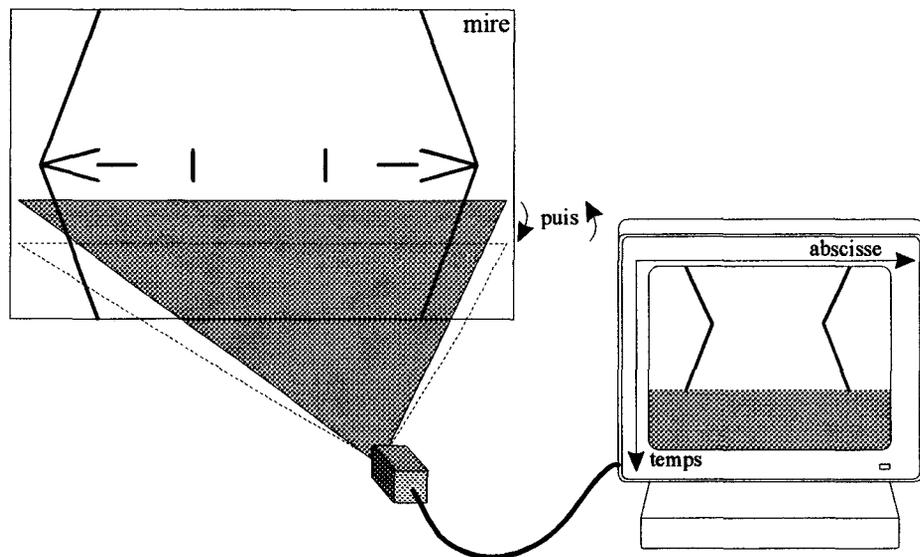


Figure 3.9 : Utilisation des lignes-guides pour régler l'angle de site.

Comme nous l'avons vu dans le chapitre 2, les caméras linéaires ont un champ vertical réduit à une ligne, qui ne permet pas, si les repères de calibrage ne sont pas visibles sur l'image, de dire si le champ de visée se trouve au-dessus ou en dessous de ces derniers. La mire doit donc fournir cette information. C'est le rôle de lignes obliques, situées de part et d'autre des repères de calibrage vertical, et appelées *lignes-guides de calibrage vertical*. Le principe de ces lignes est d'indiquer, par l'évolution de leur image lors de l'ajustage du plan de visée, si la direction choisie par l'opérateur pour cet ajustage est la bonne. Sur l'écran, les images des lignes-guides doivent s'écarter l'une de l'autre pour indiquer que l'ajustement va dans le bon sens (cf. Figure 3.9).

Deux jeux de lignes-guides sont prévus. Des lignes quasi-verticales (pente  $\pm 3$  environ), appelées *lignes-guides de réglage grossier*, permettent un réglage grossier au début du travail de calibrage. Des lignes quasi-horizontales (pente  $\pm 1/3$  environ) permettent le réglage fin lorsque le plan de visée est proche des repères de calibrage vertical. Elles sont appelées *lignes-guides de réglage fin*. Les lignes quasi-verticales et quasi-horizontales se confondent dans le prolongement des repères de calibrage vertical.

### 3.3.2. Dimensions de la mire

Les dimensions de la mire ne sont pas arbitraires. Elles doivent être fixées en fonction des caractéristiques de la caméra et des dimensions de la scène à observer (cf. Figure 3.10) :

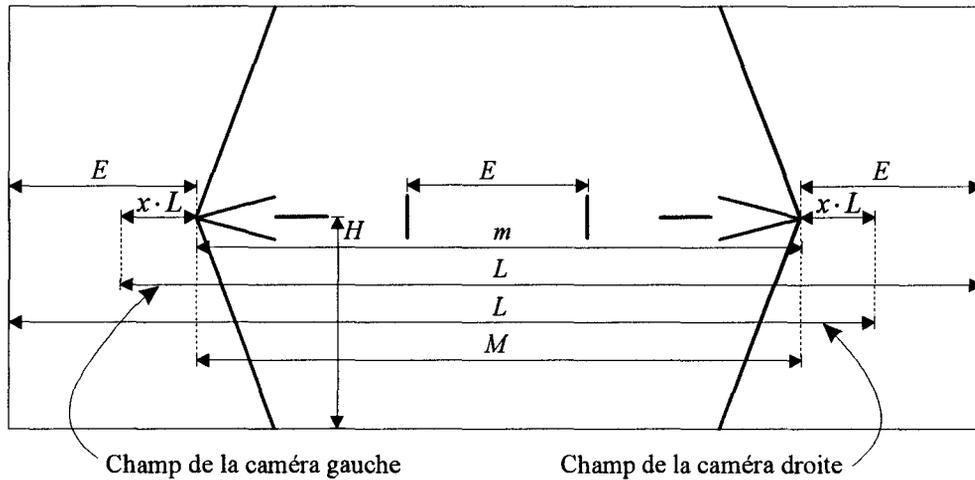


Figure 3.10 : Dimensions de la mire de calibrage.

- $l$  : longueur du capteur (fournie par le fabricant).
- $p$  : largeur du pixel sur le capteur (fournie par le fabricant).
- $f$  : distance focale des objectifs (choisie en fonction de  $l$  et du champ horizontal souhaité).
- $E$  : écartement des caméras (choisi en fonction de la précision, du champ horizontal et de la distance minimale de perception souhaités).
- $h$  : hauteur des caméras par rapport au sol (imposée par l'emplacement du stéréoscope linéaire dans le véhicule).
- $D_{max}$  : distance maximale de détection (à choisir en fonction de l'application).
- $D$  : distance caméras-mire (choisie en fonction de la place disponible dans le local où est fait le calibrage; la précision est d'autant meilleure que  $D$  est plus proche de  $D_{max}$ ).
- $x \cdot L$  : marge de sécurité sur le cadrage.  $x \cdot L$  est la distance qui sépare le bord du secteur de stéréovision du bord de la mire. Si  $x \cdot L$  est nulle, la mire est la plus large possible, ce qui améliore la précision du réglage, mais tout dérèglement de l'angle d'azimut  $\psi$  empêche les caméras de voir l'ensemble de la mire. Si  $x \cdot L$  est grande, la mire est étroite, ce qui permet de la voir en entier même si  $\psi$  est très mal ajusté, mais la précision de réglage est moins bonne. Une valeur de  $x$  comprise entre 0,05 et 0,1 semble un bon compromis.

Des considérations géométriques simples donnent :

$$L = \frac{l \cdot D}{f} : \quad \text{largeur du champ de visée sur la mire.}$$

$$m = L \cdot (1 - 2x) - E : \quad \text{largeur effective de la mire.}$$

$$M = L + E : \quad \text{largeur totale de la mire.}$$

$$H = h \cdot \left( 1 - \frac{D}{D_{\max}} \right) : \quad \text{hauteur au sol des repères de calibrage vertical.}$$

$$P = \frac{p \cdot D}{f} : \quad \text{largeur du "pixel objet" sur la mire.}$$

L'épaisseur des lignes-guides et des repères tracés sur la mire doit être suffisante pour être clairement visible sur l'image. Pour autant, elles ne doivent pas être trop épaisses, afin de permettre la focalisation des objectifs et une bonne précision de repérage. Les essais ont montré qu'une largeur correspondant à 3 pixels environ, soit  $3P$ , constituait une bonne valeur.

## 3.4. Procédure de calibrage

### 3.4.1. Procédure pratique de calibrage

La procédure de calibrage d'un stéréoscope linéaire est plus difficile à expliquer qu'à pratiquer.

Nous présenterons cette procédure sous la forme d'un guide pratique utilisable par l'opérateur. C'est d'ailleurs l'expérience de séances de calibrage qui nous a permis de la mettre au point. Les justifications théoriques nécessaires seront données au § 3.4.2.

#### Phase 1: mise en place

Installer la mire verticalement en s'assurant que la ligne portant les repères de calibration verticale est parfaitement horizontale (le mieux est de vérifier, à l'aide d'un mètre à ruban, qu'elle est parallèle au sol dont on aura vérifié la planéité). Au besoin, placer des cales sous la mire pour respecter cette condition.

Mettre le stéréoscope en position devant la mire, en s'assurant — à l'aide d'un mètre à ruban ou d'un réglet — que son axe est parfaitement horizontal et parallèle à celui de la mire, et que sa hauteur et sa distance par rapport à la mire sont bien celles choisies pour déterminer les dimensions de cette dernière (cf. § 3.3.2). Au besoin, retoucher les hauteurs des vis (A), (B) et (C) des supports orientables pour respecter ces conditions.

Allumer les projecteurs et s'assurer que la mire est uniformément éclairée. Au besoin, ne pas hésiter à déplacer les projecteurs.

Mettre les objectifs sur la position de mise au point à l'infini ( $\infty$ ). Mettre en marche le stéréoscope linéaire et le moniteur de visualisation. Mettre ce dernier en mode "plein cadre" pour s'assurer que les lignes-guides sont bien visibles. Si l'écran est trop sombre, ouvrir le diaphragme de l'objectif sur la caméra jusqu'à obtenir une image bien contrastée, blanche sans saturation. Si l'image est trop claire, fermer le diaphragme. Passer en mode "gros plan" et visualiser l'image d'une des lignes-guides sur l'une des caméras. Régler la bague de mise au point de cette caméra jusqu'à obtenir une image nette de cette ligne-guide (cf. Figure 3.11). Effectuer ensuite les mêmes réglages sur l'autre caméra.

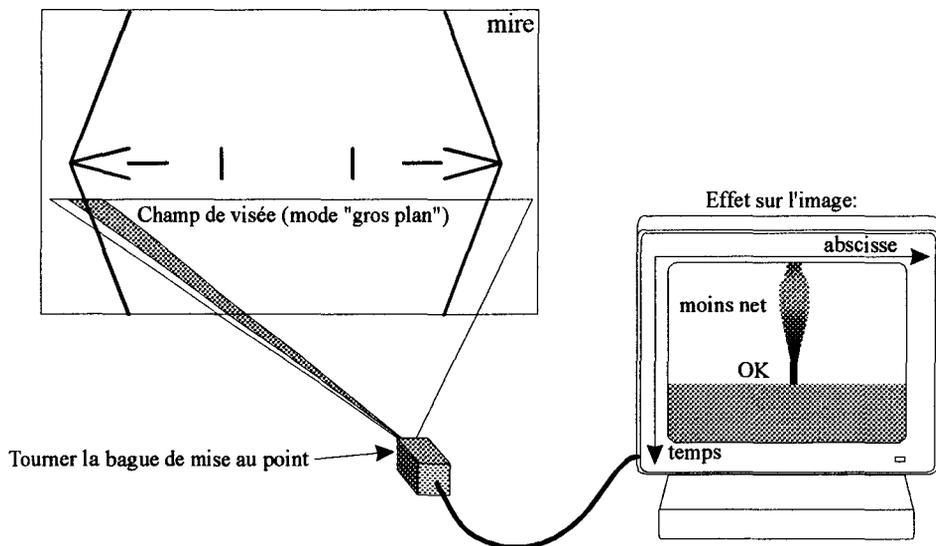


Figure 3.11 : Réglage du point.

Les phases suivantes doivent être faites d'abord sur une caméra, puis sur l'autre (par exemple, faire les phases 2 à 7 sur la caméra de gauche, puis recommencer à partir de la phase 2 sur la caméra de droite). Si l'image de l'une ou l'autre des lignes-guides devait sortir de l'écran en mode "plein cadre", ajuster l'angle d'azimut à l'aide des vis (D) et (E) (cf. Figure 3.15).

## Phase 2: réglage grossier de l'angle de site

Passer en mode d'affichage "plein cadre". Les deux lignes-guides de réglage grossier doivent être visibles (cf. Figure 3.12a). Desserrer les vis tirant (A), (B) et (C) du support pour permettre le réglage, puis tourner la vis poussant de site (C) dans le sens que l'on désire, tout en observant l'écran. Si les images des lignes-guides s'écartent l'une de l'autre, continuer dans ce sens, sinon tourner dans l'autre sens. Les images des lignes-guides de réglage fin deviennent visibles, indiquant que l'on approche du but (cf. Figure 3.12b). Arrêter de tourner la vis lorsque les images des lignes-guides de réglage grossier et fin se confondent sur l'écran, soit d'un côté soit de l'autre (cf. Figure 3.12c). On est alors ajusté à peu près correctement en site, mais pas en roulis.

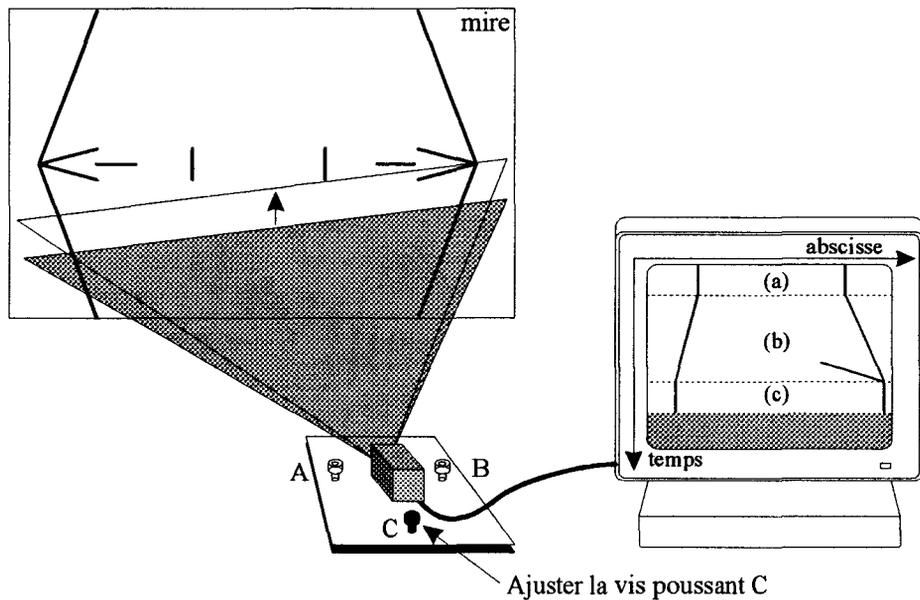


Figure 3.12 : Phase 2 - réglage grossier de l'angle de site.

### Phase 3: réglage fin des angles de site et de roulis

Tourner la vis poussant (A) ou (B) située du côté opposé aux lignes-guides dont les images sont confondues sur l'écran à la fin de la phase 2, jusqu'à ce que les images des lignes-guides situées de l'autre côté soient à leur tour confondues (cf. Figure 3.13a). On remarquera que cette opération modifie également l'angle de site, ce qui fait que — sauf si l'angle de roulis était très proche du bon réglage au départ — les images des lignes-guides qui étaient confondues ne le sont plus (cf. Figure 3.13b): c'est parfaitement normal. Tourner maintenant l'autre vis pour confondre à nouveau ces images. Recommencer l'opération en tournant alternativement les vis poussant (A) et (B) jusqu'à ce que les images des lignes-guides soient confondues des deux côtés. La justification de cette procédure sera donnée au § 3.4.2. Pendant toute cette phase, ne pas tenir compte des images des repères de calibrage vertical et horizontal qui apparaissent à l'écran (cf. Figure 3.13c).

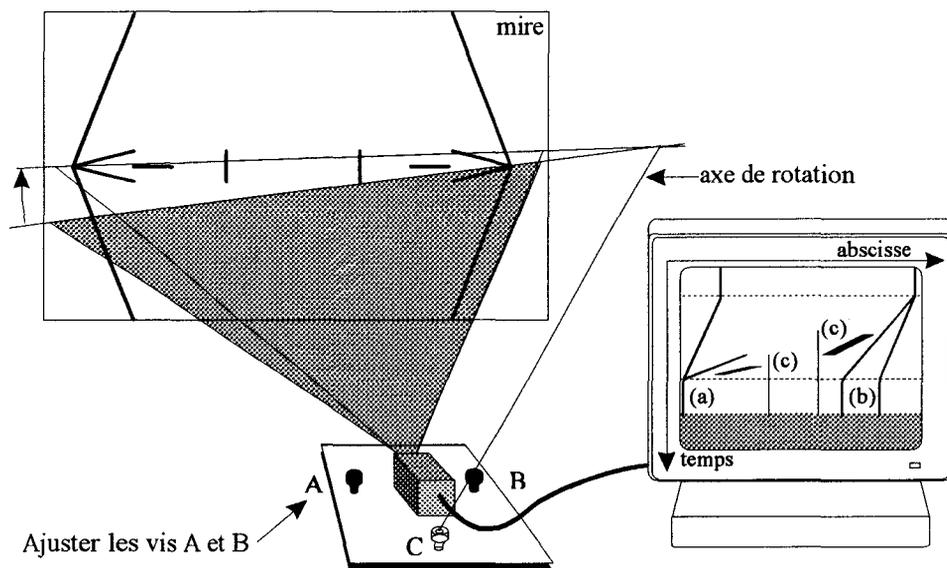


Figure 3.13 : Phase 3 - réglage fin des angles de site et de roulis.

### Phase 4: ajustement final du calibrage vertical

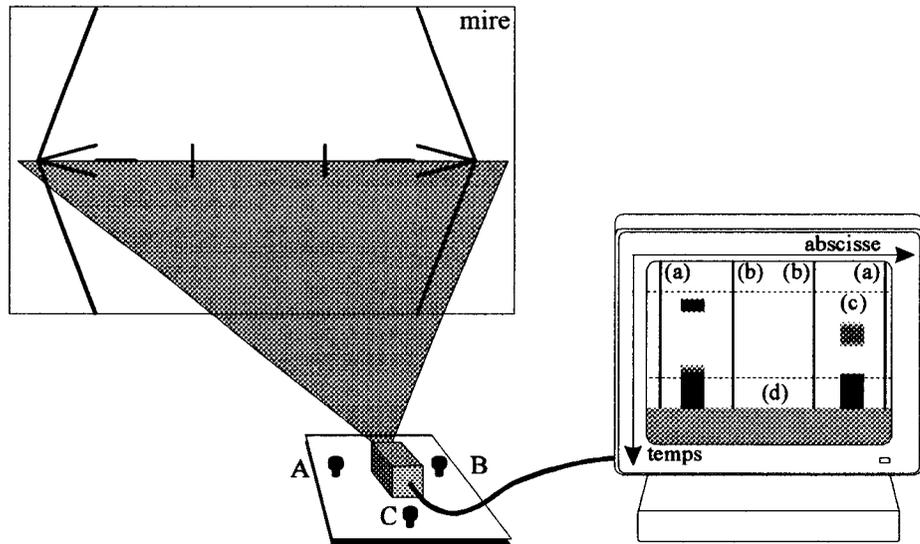


Figure 3.14 : Phase 4 - ajustement final du calibrage vertical.

Lorsque les images des lignes-guides sont confondues de chaque côté de la mire (cf. Figure 3.14a), il ne reste plus qu'à rendre visible les deux repères de calibrage vertical. Pour ce faire, on reprendra le réglage de la vis poussant (C) et, au besoin, celui de la vis (A) ou (B), mais avec beaucoup de doigté: on est maintenant très près du réglage optimal, et une très légère retouche suffit pour y être tout à fait (cf. Figure 3.14d)... ou pour ne plus y être du tout! Pendant cette phase, ne pas tenir compte des repères de calibrage horizontal qui apparaissent à l'écran (cf. Figure 3.14b). Enfin, faire attention au contraste des images des repères de calibrage vertical (cf. Figure 3.14c): elles doivent être aussi sombres que possible.

On pourra peaufiner le réglage en manoeuvrant **LEGEREMENT** la vis (C): les images des deux repères de calibrage vertical doivent apparaître et disparaître simultanément. Dans le cas contraire, l'angle de roulis n'est pas parfaitement ajusté, et une action (prudente!) sur les vis (A) et (B) est nécessaire.

### Phase 5: blocage du calibrage vertical

Il faut maintenant serrer les vis tirant (A), (B) et (C) pour bloquer le calibrage vertical. Mais attention: du fait de sa grande sensibilité, le réglage risque d'être remis en cause par les flexions des platines si l'on se contente de serrer les vis sans précautions. C'est pourquoi nous donnons ici LA méthode qui nous a permis de terminer nos séances de calibrage sans tout gâcher!

- Commencer par serrer TRES LEGEREMENT les trois vis tirant, simplement pour les mettre au contact de la platine supérieure afin d'éviter tout "décollage" de celle-ci par la suite.
- Serrer légèrement l'une des vis tirant en observant sur le moniteur l'image de la mire. Si un décalibrage devient visible, serrer légèrement la vis poussant correspondante pour le faire disparaître. On passe ensuite à la vis suivante, puis à la troisième.
- Continuer à faire le tour, en serrant à chaque fois un peu plus. Il est normal d'avoir à reprendre le réglage à chaque fois, celui-ci étant extrêmement sensible.
- Cette phase ne prend fin que lorsque les trois vis sont serrées à bloc: elles ne devront pas bouger lorsque le stéréoscope sera en opération sur route!

### Phase 6: réglage de l'angle d'azimut

Passer en mode d'affichage "gros plan" et afficher le réticule au milieu de l'image (cf. Figure 3.15). Desserrer les vis de blocage (D) et de réglage (E) de l'angle d'azimut, puis tourner DELICATEMENT la caméra jusqu'à ce que l'image du repère de calibrage horizontal correspondant à la caméra soit confondu avec le réticule. Resserrer les vis (D) et (E), à bloc mais en procédant progressivement comme dans la phase 5, afin de ne pas compromettre les réglages. Si tel était le cas (c'est à dire si les repères de calibrage horizontal disparaissaient de l'image) il faudrait alors reprendre les réglages à partir de la phase 2.

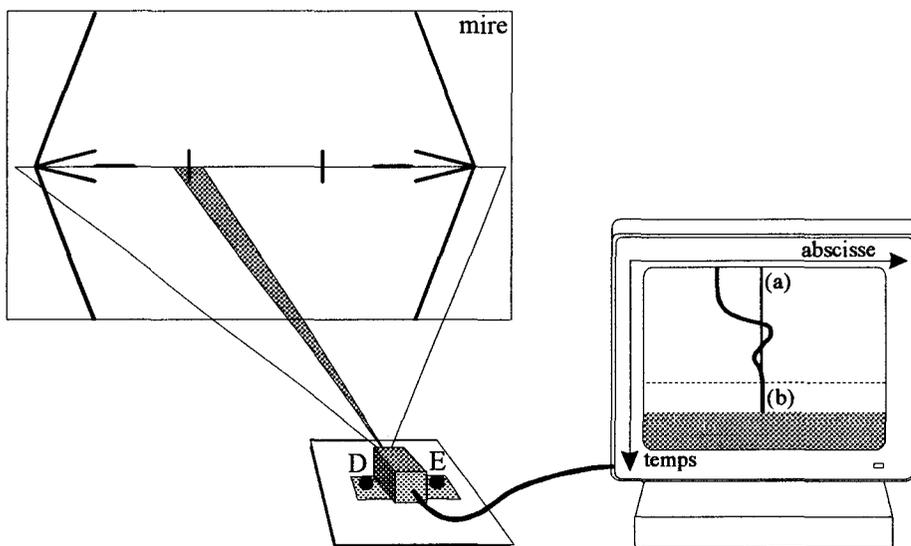


Figure 3.15 : Réglage de l'angle d'azimut.

### 3.4.2. Justification théorique de la procédure

La méthode itérative utilisée pour l'ajustement fin des angles de site et de roulis, dans la phase 3 de la procédure de calibrage, peut sembler déroutante. Elle est pourtant parfaitement adaptée à un réglage très précis et simultané de ces deux angles, comme le raisonnement ci-dessous le montre.

Définissons d'abord l'axe horizontal comme étant confondu avec l'axe qui joint les repères de calibrage vertical de la mire (cf. Figure 3.16). L'axe vertical est normal à l'axe horizontal, orienté vers le haut.

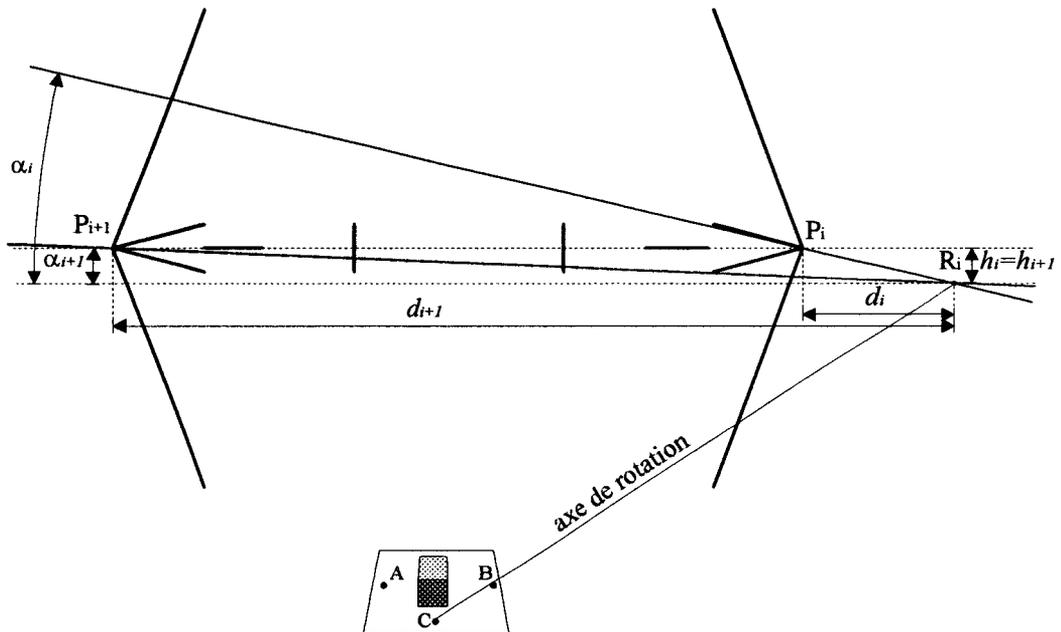


Figure 3.16 : Justification de la procédure de calibrage.

On suppose que l'on a ajusté les angles de site et de roulis de façon à viser l'intersection des lignes-guides de calibrage vertical d'un côté de la mire. La Figure 3.16 représente cette situation, le point visé étant noté  $P_i$ . La ligne correspondant au champ de visée sur la mire à ce moment fait un angle  $\alpha_i$  avec l'horizontale.

La méthode de réglage impose maintenant d'ajuster la vis (A) ou (B) située du côté **opposé** aux lignes-guides dont l'intersection est atteinte, afin de viser l'intersection des lignes-guides situées du côté opposé. On vise alors le point  $P_{i+1}$ . La ligne qui correspond au nouveau champ de visée de la caméra sur la mire fait avec l'horizontale l'angle  $\alpha_{i+1}$  (cf. Figure 3.16).

On doit démontrer que  $\alpha_{i+1} < \alpha_i, \forall i$ , ce qui signifie que l'angle de roulis de la caméra est mieux toujours mieux ajusté après réglage qu'avant. Ceci implique qu'un nombre suffisant d'itérations de cette procédure permet d'arriver au réglage correct de l'angle de roulis — et de l'angle de site, puisqu'alors le champ de visée inclut les deux repères de calibrage vertical.

Rappelons (cf. Figure 3.16) qu'en laissant fixes deux vis du système mécanique de réglage (ici, A et C, ou B et C) on définit un axe de rotation passant par ces deux vis, et qui forme avec l'axe optique du stéréoscope un angle de  $30^\circ$  (cf. Figure 2.34). L'intersection de cet axe avec la mire est un point  $R_i$ , d'ordonnée  $h_i$  dans le repère défini ci-dessus, et qui reste invariant par rotation autour de l'axe  $\{(B),(C)\}$ . Remarquons également que, du fait de la procédure de réglage,  $R_i$  est situé du même côté de la mire que le point  $P_i$  (cf. Figure 3.16).

Avant réglage, la distance qui sépare le point  $R_i$  du point  $P_i$  est notée  $d_i$  selon l'axe horizontal et  $h_i$  selon l'axe vertical.

Après réglage de l'angle de roulis pour faire passer le champ de visée par le point  $P_{i+1}$ , la distance verticale  $h_{i+1}$  est égale à  $h_i$ , puisque la hauteur de  $P_{i+1}$  est égale à celle de  $P_i$ . En revanche, la distance horizontale  $d_{i+1}$  est supérieure à  $d_i$ , puisque  $P_{i+1}$  est situé du côté de la mire opposé à l'axe de rotation, alors que  $P_i$  est situé du même côté.

On a donc les relations suivantes:

- avant réglage,  $\tan \alpha_i = \frac{h_i}{d_i}$
- après réglage,  $\tan \alpha_{i+1} = \frac{h_{i+1}}{d_{i+1}}$ , avec  $\begin{cases} h_{i+1} = h_i \\ d_{i+1} > d_i \end{cases}$

d'où :  $\tan \alpha_{i+1} < \tan \alpha_i$

et donc :  $\alpha_{i+1} < \alpha_i, \quad \forall i$

## 3.5. Résultats expérimentaux

### 3.5.1. Premiers essais de calibrage

#### Réalisation de la mire

Pour les essais, nous avons réalisé la mire de calibrage sur plusieurs plaques de *cadapack*, sorte de mousse plastique rigide recouverte sur ses deux faces de papier glacé mat de très haute qualité. Ce type de support permet de réaliser des figures planes de grandes dimensions tout en conservant une excellente rigidité, une bonne stabilité dimensionnelle et un poids très faible. Le stockage, le transport et la mise en place en sont facilitées. Les traits noirs ont été réalisés avec du ruban de dessinateur noir mat, de largeur calibrée, facile à couper au scalpel.

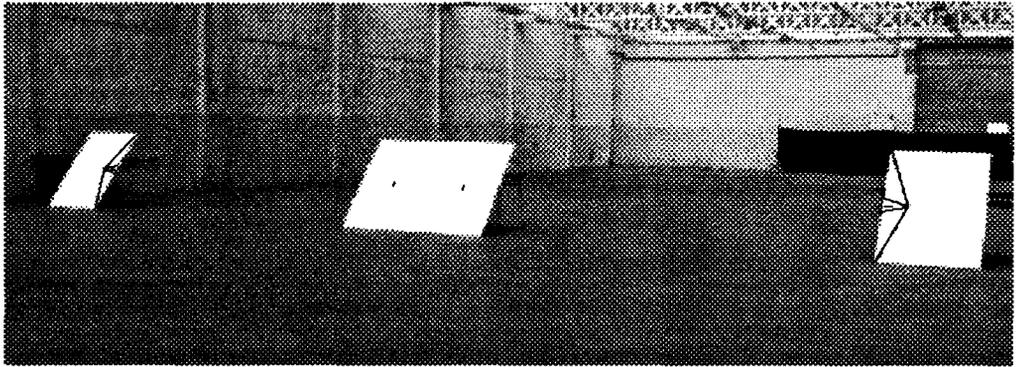


Figure 3.17 : Disposition de la mire pour un champ large (objectifs grand-angle).

Sur notre prototype, deux plaques (à gauche et à droite du champ) portent les repères de calibrage vertical et les lignes-guides, tandis qu'une troisième plaque (au centre du champ) porte les repères de calibrage horizontal. Cette disposition de la mire en trois parties indépendantes permet d'en modifier facilement les dimensions selon les paramètres de prise de vue et la distance mire-stéréoscope (cf. § 3.3.2), étant donné que la mire peut comporter des "trous" (cf. Figure 3.17) sans poser à l'opérateur de problème d'interprétation de l'image, comme les essais l'ont démontré.

### Réalisation du support des caméras

Pour les premiers essais de calibrage, il était évidemment peu souhaitable de monter d'entrée de jeu les caméras sur un véhicule, ce qui aurait imposé la conception d'un système mécanique complexe avant même d'avoir pu valider cette conception par l'expérience. De plus, l'accessibilité des réglages, qui est un facteur très important durant la période de mise au point d'un prototype, aurait posé des problèmes.

Nous avons donc utilisé une maquette à l'échelle, dont les caractéristiques géométriques devaient pouvoir être ajustées afin d'estimer la qualité des résultats obtenus dans différentes configurations. Les paramètres à ajuster étaient l'écartement des deux axes optiques, la hauteur des caméras par rapport à la chaussée et l'angle de site.

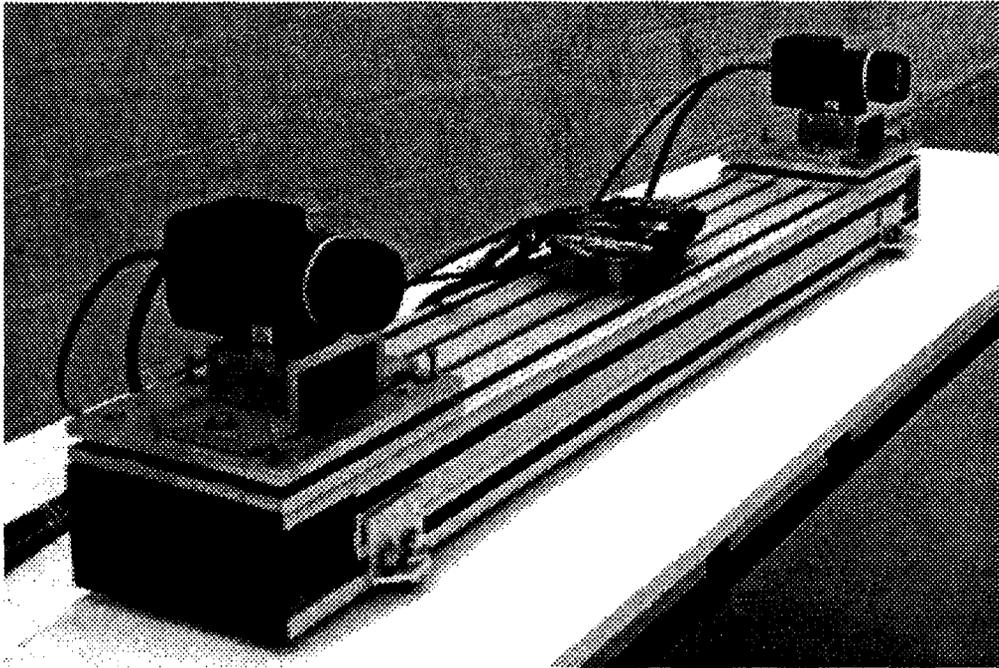


Figure 3.18 : Monture utilisée pour les essais du stéréoscope.

Les caméras, sur leurs montures orientables respectives, étaient portées par une poutre commune très rigide, en profilé d'aluminium (cf. Figure 3.18), qui modélisait le châssis du véhicule, et dont l'assiette par rapport au sol pouvait être ajustée par 4 vis calantes. Il était ainsi possible de rendre l'axe joignant les caméras parallèle à l'axe de la mire. L'ensemble était simplement posé sur une table rigide, dont la stabilité était améliorée par la masse importante du stéréoscope (25 kg).

### **Méthodes de mesure des positions et des orientations**

Les dimensions et distances inférieures à 3 mètres ont été relevées à l'aide d'un simple mètre à ruban métallique, qui permet une précision de l'ordre du millimètre. La distance mire-stéréoscope a été mesurée au moyen d'un décamètre à ruban, avec une précision de l'ordre du centimètre. Les moyens utilisés pour ces mesures étaient volontairement limités, afin de permettre une mise en oeuvre facile (cf. § 3.1.2).

### **Méthode d'étalonnage du stéréoscope après calibrage**

Après calibrage, le stéréoscope a été soumis à un étalonnage destiné à mesurer la précision de localisation des objets. Dans ce but, nous avons utilisé un dispositif adapté, décrit ci-après.

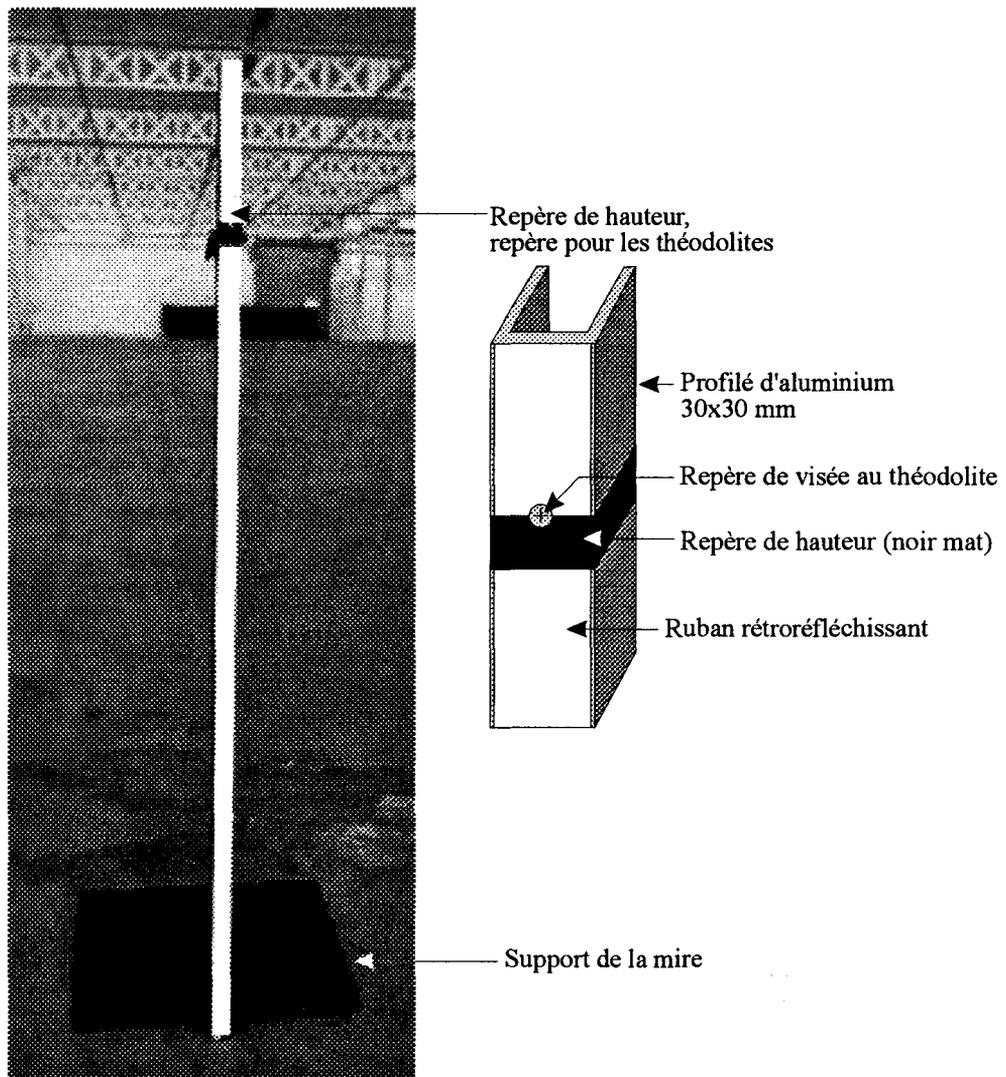


Figure 3.19 : Mire utilisée pour l'étalonnage.

La mire de calibration est remplacée par une mire particulière, comportant un simple trait vertical rétro réfléchissant (cf. Figure 3.19). Cette mire est éclairée par un projecteur halogène de 800W placé très près de l'axe du stéréoscope, afin que le plus possible d'énergie lumineuse soit renvoyé par la mire vers les caméras (cf. Figure 3.20).

Le programme de gestion du stéréoscope détecte le point le plus lumineux sur chaque image, qui correspond par hypothèse à la mire, et calcule sa position dans l'espace dans le repère  $(X, Y, Z)$  défini au chapitre 2, en partant des paramètres suivants, entrés à la main par l'opérateur: hauteur  $h$ , portée  $D_{max}$  et entraxe  $E$  du stéréoscope (fixés par l'opérateur en utilisant un mètre à ruban, comme vu plus haut), et distance focale des objectifs (donnée par le fabricant de ces derniers, avec une tolérance de  $\pm 5\%$ , cf. Chapitre 2). Cette position peut alors être comparée à la position réelle de la mire, qui doit être déterminée par un moyen de mesure très précis.

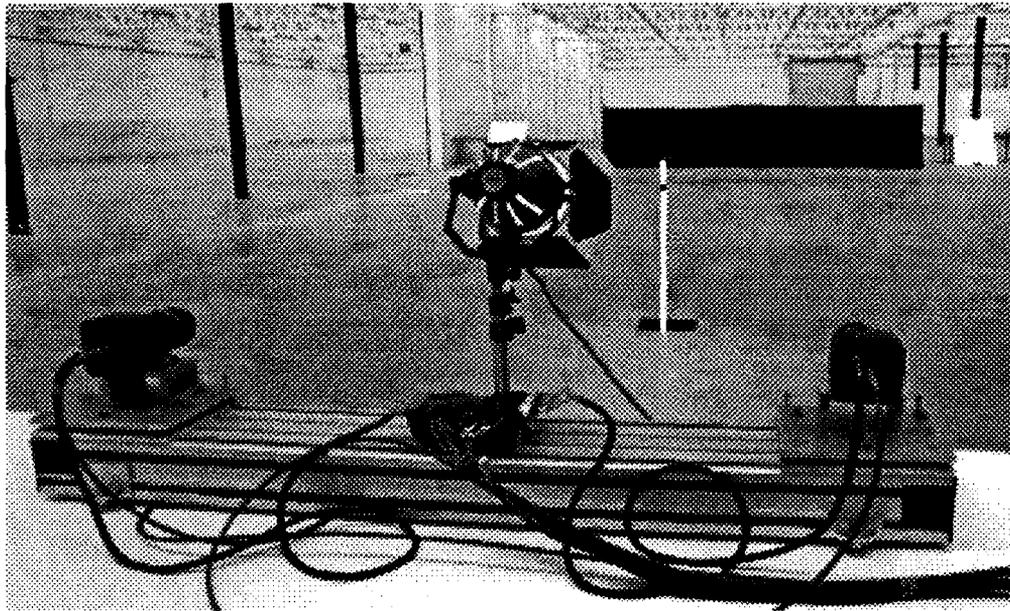


Figure 3.20 : Disposition de la mire et de l'éclairage pour l'étalonnage.

Nous avons utilisé deux théodolites de chantier pour déterminer par triangulation la position réelle de notre mire, avec une précision théorique de  $20 \mu\text{m}$  par mètre de distance, bien meilleure que celle que l'on peut espérer du stéréoscope, en raison de l'échantillonnage spatial opéré par les capteurs des caméras, dont l'influence sur les mesures se traduit par une incertitude ( $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$ ):

$$\begin{cases} \Delta X = \frac{Ep}{\delta} \\ \Delta Y = X \frac{p'}{2f} \\ \Delta Z = \frac{Ef}{\delta} - \frac{Ef}{\delta - p} \end{cases}$$

où:

- $E$  est l'écartement des axes optiques des deux caméras,
- $f$  est la distance focale des objectifs,
- $p$  est la largeur des photosites du capteur,
- $p'$  est la hauteur de ces mêmes photosites,
- $\delta$  est la disparité.

Pour mesurer la hauteur  $Y$  du point visé, nous avons placé un repère noir mat sur la surface brillante de la mire d'étalonnage. En faisant glisser ce repère verticalement le long de la mire rétro réfléchissante, on peut enregistrer la disparition de cette dernière sur les images fournies par le stéréoscope, lorsqu'il arrive dans le plan de visée des caméras. La position du repère (cf. Figure 3.19) est alors relevée à l'aide des théodolites.

### 3.5.2. Premiers essais du stéréoscope: précision de localisation

Les figures qui suivent montrent les résultats de l'étalonnage du stéréoscope, muni d'objectifs de 35 mm (cf. Figure 3.21), puis de 20 mm (cf. Figure 3.22). L'axe vertical représente la profondeur (distance le long de l'axe optique), l'axe horizontal représente l'abscisse (distance perpendiculairement à l'axe optique).

Chaque point muni d'un trait représente une mesure, le gros point correspondant à la position réelle de la mire (mesurée aux théodolites), tandis que l'extrémité du trait représente la position mesurée au stéréoscope. Les rectangles d'incertitude sur la position correspondant au seul échantillonnage du capteur sont également tracés, les distorsions et tout autre effet étant ignorés. Ceci permet de vérifier l'influence réelle de ces effets.

Les rectangles d'incertitude, correspondant à l'échantillonnage spatial par les capteurs CCD du stéréoscope, ont les dimensions suivantes:

$$\begin{cases} \Delta X = \frac{Ep}{\delta} \\ \Delta Z = \frac{Ef}{\delta} - \frac{Ef}{\delta - p} \end{cases}$$

Notons enfin que les contraintes liées aux salles dans lesquelles nous avons opéré nous ont empêché de réaliser certaines mesures:

- La salle dans laquelle nous avons réalisé les essais avec les objectifs de 35 mm n'était pas assez large pour étalonner tout le champ. Il reste donc des lacunes sur les bords.
- La salle dans laquelle nous avons réalisé les essais avec les objectifs de 20 mm n'était longue que de 20 m. La profondeur est donc limitée.

On constate que les erreurs sur la localisation dépassent notablement la seule imprécision de l'échantillonnage spatial, en particulier en bord de champ. Les écarts entre la distance réelle et la distance mesurée atteignent, avec une focale de 20 mm, jusqu'à 5% à l'extrême droite du champ, et 15% à l'extrême gauche (cf. Figure 3.22).

Des erreurs de mesure aussi importantes ne semblent pas explicables par la seule distorsion des objectifs (cf. § 2.3.1.): [Chasseur 92] donne 0,4% de distorsion pour les Nikkor AF 2,8/20 mm et AIS 2/35 mm.

Par ailleurs, l'asymétrie constatée entre les deux côtés du champ semble indiquer un désalignement de l'axe optique de l'une au moins des deux caméras.

Cependant, la précision atteinte étant suffisante pour le but fixé, nous n'avons pas poussé plus loin nos investigations sur ce sujet.

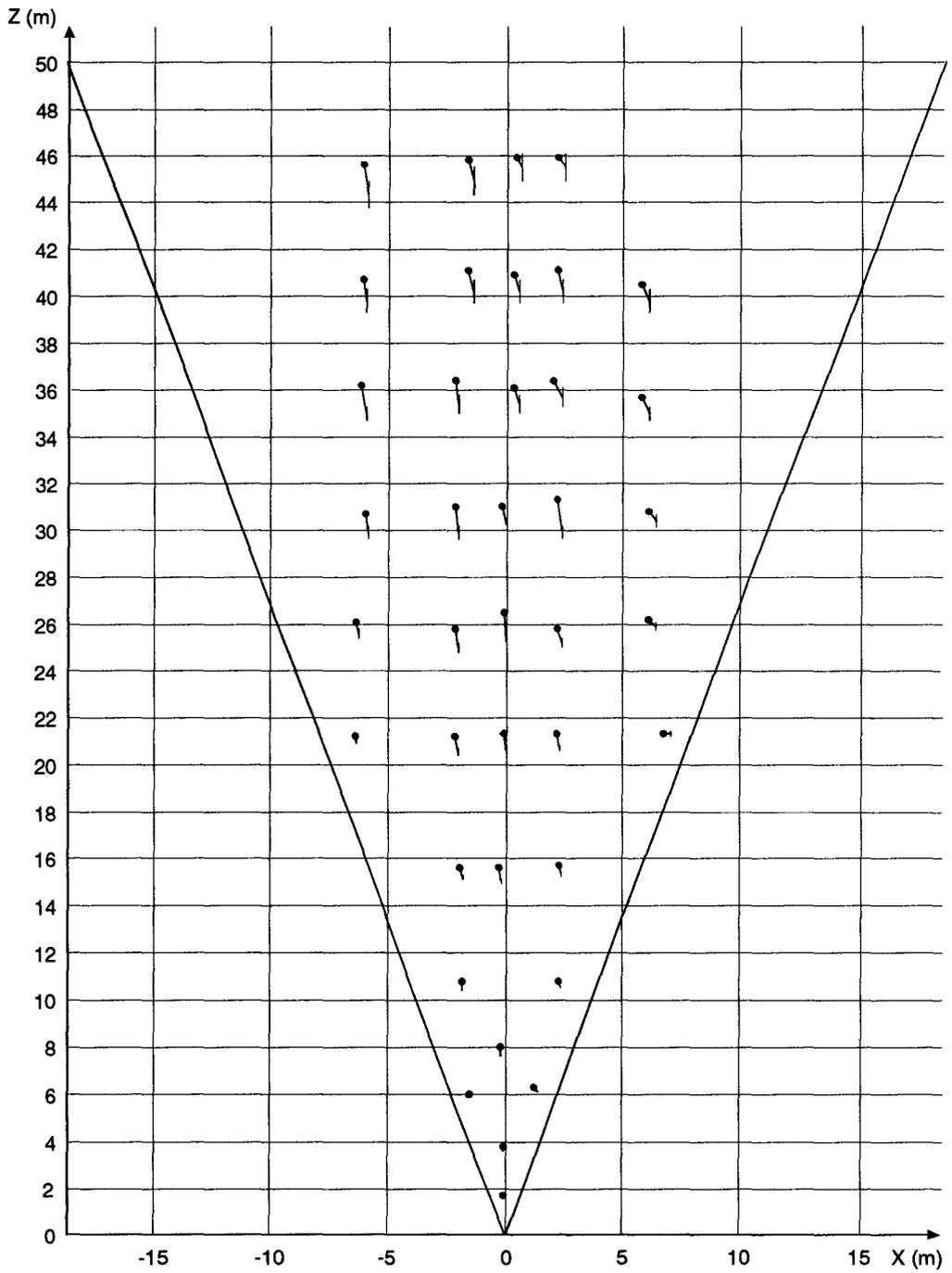


Figure 3.21 : Résultat de l'étalonnage avec des objectifs Nikon AIS 2/35mm

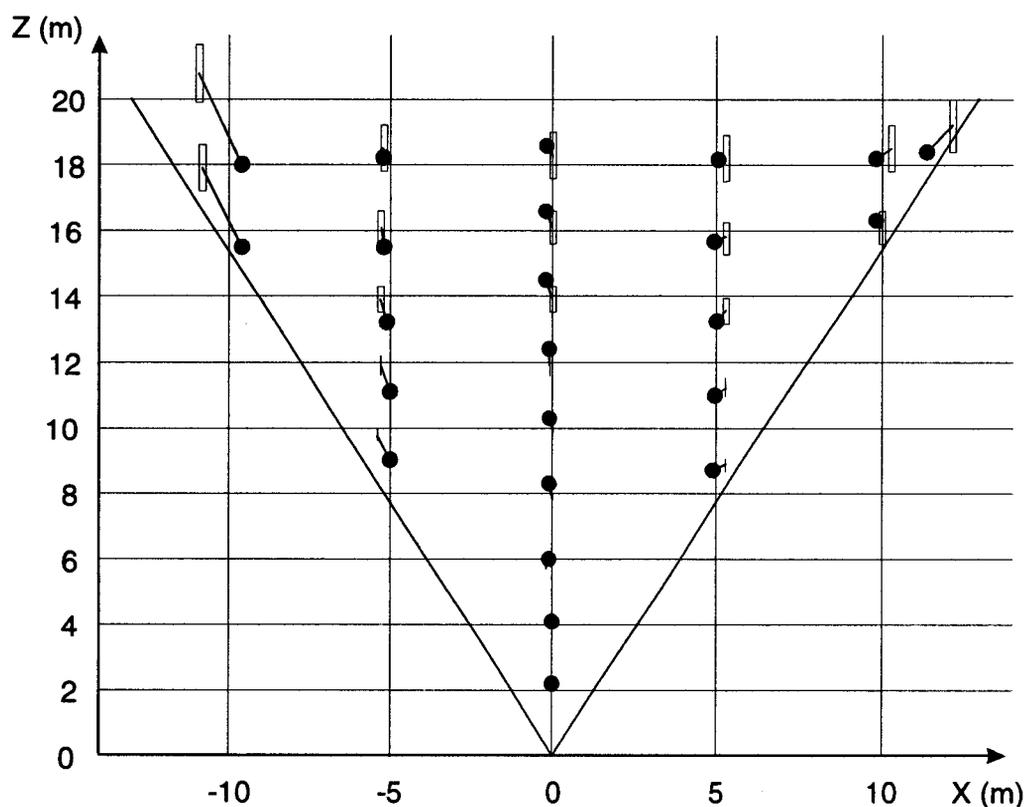


Figure 3.22 : Résultat de l'étalonnage avec des objectifs Nikkor AF 2,8/20mm

### 3.6. Remarques pratiques

Pour les essais, nous avons volontairement simplifié les manipulations en plaçant les caméras sur un support rigide au lieu de les mettre sur un véhicule. En effet, ce dernier possède une suspension, dont l'effet - volontaire - est de modifier son assiette lorsqu'une force extérieure lui est appliquée. Malheureusement, cet effet se fait également sentir lorsque l'on manipule les vis de la monture des caméras, avec pour conséquence que le réglage du stéréoscope devient impossible. Pour éviter tout mouvement provoqué par les réactions de la suspension du véhicule durant les manipulations, ce dernier doit être placé sur cales pour la durée du calibrage. Ceci entraîne inévitablement des changements de son assiette et de sa hauteur au sol, qui doivent être mesurées (il suffit de mesurer la hauteur des quatre points d'ancrage avec un mètre à ruban) afin de les prendre en compte dans le calcul de la hauteur  $H$  des repères de calibrage vertical de la mire (cf. § 3.3.2).

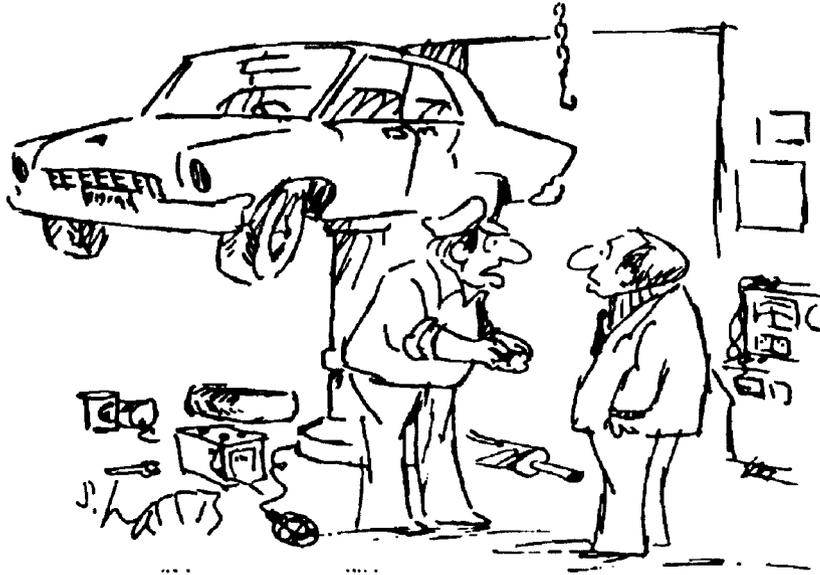
---

## 3.7. Conclusion

La procédure de calibrage du stéréoscope linéaire est originale par le fait qu'elle met le stéréoscope dans une configuration déterminée à l'avance par l'opérateur en fonction des besoins de l'application. Elle permet de plus la mise en coplanéité des champs de visée des deux caméras, condition indispensable pour pouvoir mettre leurs images en correspondance.

Les étalonnages réalisés avec une distance focale de 35mm indiquent une précision de localisation meilleure que 4% sur la distance  $Z$  des objets détectés (cf. Figure 3.21). Avec une distance focale de 20mm, l'erreur maximale monte à 15%, mais ce chiffre assez médiocre n'est obtenu qu'à l'extrême gauche du champ des objectifs, la précision étant bien meilleure hors de cette zone (cf. Figure 3.22). Nous pourrions facilement l'améliorer en prenant en compte l'étalonnage dans les calculs de reconstruction 3D, mais nous nous en abstenons pour éviter la perte de temps occasionnée par les calculs nécessaires. La précision déjà obtenue est, de toutes manières, très suffisante pour nos besoins. Des méthodes permettant d'améliorer la précision du stéréoscope linéaire sont présentées dans l'Annexe B.

La description du stéréoscope linéaire lui-même est maintenant terminée. Le prochain Chapitre présente les méthodes de traitement d'image qui utilisent les images qui en sont issues pour détecter et localiser les objets présents dans son secteur de stéréovision.



«C'est l'ordinateur de bord, il va falloir que j'appelle un analyste-système.»

## Chapitre 4

# TRAITEMENT DES IMAGES STEREOSCOPIQUES LINEAIRES

---

## 4.1. Introduction

Nous disposons maintenant de tous les outils nécessaires pour capter des images avec le stéréoscope linéaire (cf. Chapitre 2) et localiser dans la scène les objets que nous y aurons mis en évidence (cf. Chapitre 3). Ce chapitre présente les méthodes de traitement qui vont nous permettre de reconnaître dans les images les objets présents dans la scène, de les localiser et de déterminer s'ils représentent ou non des obstacles.

Les traitements stéréoscopiques se conforment à la structure générale décrite dans le Chapitre 1: extraction de primitives (§ 4.2), mise en correspondance, (§ 4.3), et triangulation géométrique pour la reconstruction de la scène tridimensionnelle (§ 4.4).

Comme les images linéaires contiennent relativement peu d'information sur la scène, quelques erreurs et/ou ambiguïtés de mise en correspondance sont possibles. C'est pourquoi la procédure de mise en correspondance doit s'assurer de la cohérence de ses résultats pour éliminer ces erreurs éventuelles. Il s'agit ici d'analyser l'évolution temporelle de la scène tridimensionnelle reconstruite afin d'y rechercher d'éventuelles incohérences et, le cas échéant, d'y remédier.

Comme on l'a vu dans l'introduction de ce mémoire ( cf. § 1.5), les traitements présentés ici ne sont qu'une préfiguration des traitements définitifs, qui font l'objet d'une deuxième thèse sur le sujet. Dans le cadre de cette première thèse, plus centrée sur le dispositif de prise de vue, la procédure de reconstruction de la scène tridimensionnelle et l'analyse de l'évolution dynamique de la scène, qui ne sont pas nécessaires pour vérifier le bon fonctionnement du stéréoscope linéaire et tester sa précision, ne sont pas encore développées.

---

## 4.2. Extraction des primitives

### 4.2.1. Nature des primitives

Comme on l'a vu au § 1.4, on peut utiliser de nombreux types de primitives pour analyser des images matricielles. Cependant, la vision linéaire réduit notablement les possibilités de choix.

- En effet, la notion de contour n'existe pas réellement en une seule dimension (cf. Figure 4.2): les contours, qui sont des courbes en vision 2D, se réduisent sur des images linéaires à leur intersection avec le plan de visée, et donc à des points, naturellement baptisés *points de contour* (cf. Figure 4.2).

Notons que cette notion de point de contour est différente de la notion de point de contour en vision matricielle: dans ce dernier cas, un point de contour est défini comme un point particulier de l'ensemble d'un contour perceptible au moins en partie. Il est courant d'attribuer au point de contour des propriétés géométriques semi-locales, telles que le vecteur direction du contour en ce point (cf. Figure 4.1), qui apportent des informations utiles à

la mise en correspondance [Horaud 93-2]. En vision linéaire, de telles propriétés ne peuvent pas être définies (cf. Figure 4.2d), et un point de contour doit être considéré comme un point unique, et non comme partie d'une courbe.

- De même, du fait de l'extrême pauvreté des arguments géométriques en vision linéaire, les points d'intérêt se limitent en pratique aux seuls points de contour.
- Le choix des éléments significatifs se limite donc aux points de contour et aux régions homogènes, lesquelles se réduisent ici à des segments de droite, parties de l'image linéaire sur lesquelles le niveau de gris est uniforme. Or, comme on l'a vu au § 1.4.2, les images des régions homogènes sont trop sensibles aux effets de perspective pour être réellement utilisables. Par ailleurs, le seul intérêt des régions pour la mise en correspondance est leur forme, et celle-ci ne peut être connue par des images linéaires (cf. Figure 4.2f). Les primitives de type région sont donc sans intérêt en vision linéaire.

Nous choisirons donc d'utiliser les points de contour, qui se définissent comme les *fronts* présents dans l'image linéaire, c'est à dire les variations brusques et d'amplitude importante du niveau de gris de l'image (cf. Figure 4.3).

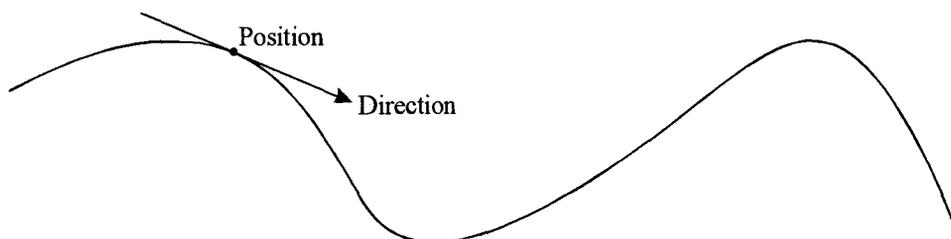


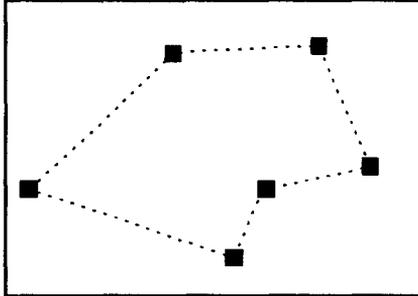
Figure 4.1: Attributs d'un point de contour en vision matricielle.

## 4.2.2. Méthode d'extraction des primitives

La méthode d'extraction des primitives d'une image se compose de deux étapes: on commence par calculer le gradient de l'image, puis on cherche les extréma de ce gradient, qui correspondent aux endroits où les variations du niveau de gris sont les plus importantes, et donc aux points de contour (cf. Figure 4.3).

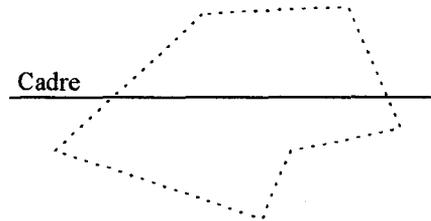
L'utilisation du gradient de l'image (qui, en une dimension, est simplement la dérivée du signal  $g(x)$ , où  $g$  est le niveau de gris et où  $x$  est l'abscisse du point objet sur le capteur) est une méthode de détection des contours très efficace. Son seul défaut est sa grande sensibilité aux bruits de toutes sortes, que ces bruits soient présents dans la scène (poussières, texture), dus au capteur (bruit de pair-impair), aux circuits vidéo analogiques (souffle) ou à l'échantillonneur (bruit de conversion). Dans la pratique, l'utilisation sans précautions de la dérivée du signal ferait sans doute apparaître plus de "faux" points de contour, dus au bruit, que de vrais.

Limites du cadre



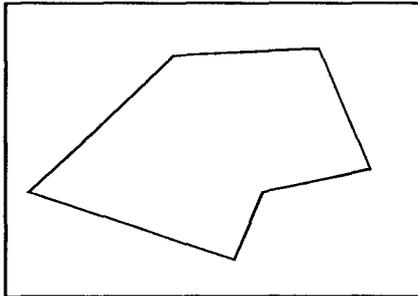
(a) points d'intérêt (ici les angles) en vision matricielle

Cadre



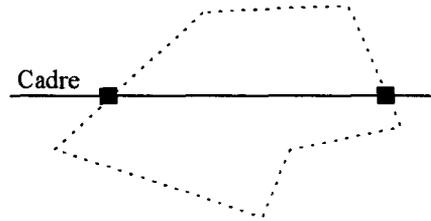
(b) Points d'intérêt en vision linéaire:  
La caméra n'en voit aucun.  
D'ailleurs comment les définir?

Limites du cadre



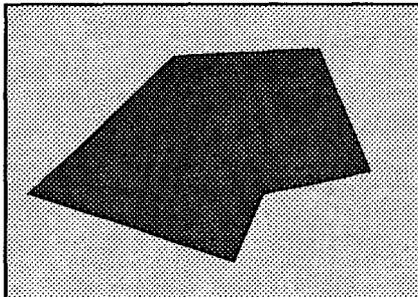
(c) Contours en vision matricielle

Cadre



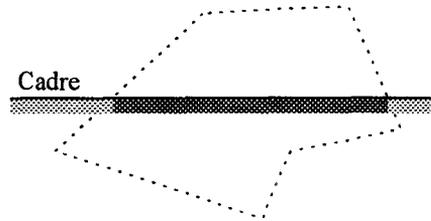
(d) Contours en vision linéaire  
(points de contour)

Limites du cadre



(e) Régions en vision matricielle  
(cette image contient 2 régions)

Cadre



(f) Régions en vision linéaire  
(cette image contient 3 régions)

Figure 4.2 : Comparaisons des primitives en vision matricielle et en vision linéaire.

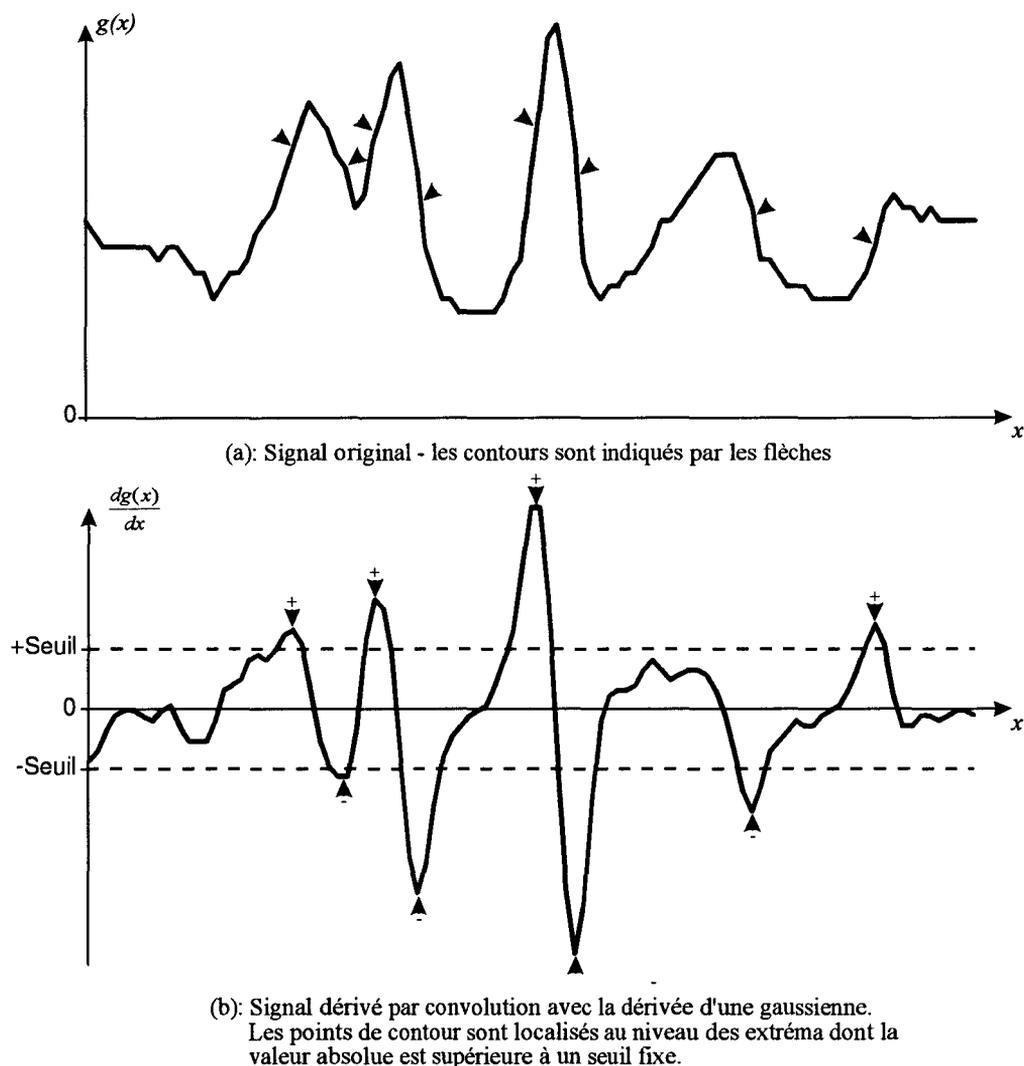


Figure 4.3 : Principe de la détection des points de contour dans une image linéaire.

Pour réduire l'influence du bruit sur la localisation des contours, on peut utiliser deux moyens, souvent associés:

- **Lissage du signal.** Il s'agit de convoluer le signal  $g(x)$  par un filtre moyenneur (ou tout autre filtre passe-bas) avant d'en calculer la dérivée. Ce lissage du signal diminue l'influence du bruit, mais aussi la résolution, puisque le filtre passe-bas estompe les fronts. On peut remédier partiellement à cet inconvénient en utilisant, au lieu du filtre moyenneur, un filtre médian [Tukey 74, Gallagher 81].
- **Seuillage du gradient.** Au cours de la recherche des extrêmes du gradient, on peut ignorer les "pics" dont l'amplitude est trop faible, comme sur la Figure 4.3. Un simple seuillage élimine ceux qui sont dus au bruit de fond, tout en conservant ceux qui correspondent à des contours significatifs. Le problème est évidemment de choisir correctement la valeur du seuil, de manière à éliminer l'ensemble du bruit de fond tout en conservant tous les "vrais" contours. Le choix du

seuil est donc important, et ce dernier doit souvent être ajusté en fonction des conditions de prise de vue. Nous conserverons cependant un seuil fixe choisi par l'opérateur, pour des raisons de simplicité et de rapidité d'exécution des traitements.

La Figure 4.4 représente l'arbre programmatique qui correspond à l'algorithme général de recherche des points de contour. Les "feuilles" de cet arbre seront développées dans les paragraphes suivants. Le type PrimMono ("primitive monoscopique") sera également décrit en temps utile (cf. § 4.2.3). Les notations sont celles du langage C.

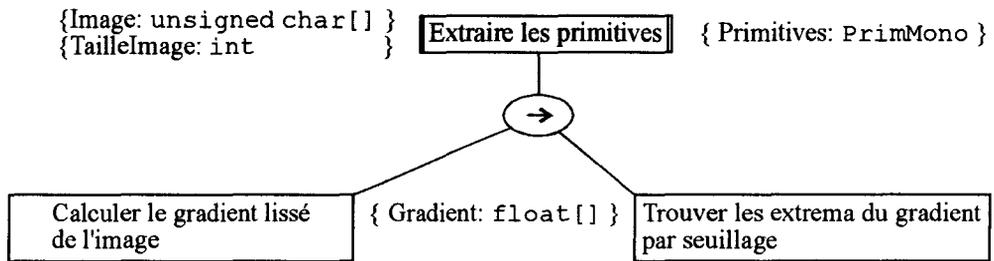


Figure 4.4 : Algorithme d'extraction des points de contour (vue générale).

### Calcul du gradient de l'image

Pour calculer la dérivée du signal  $g(x)$ , nous avons choisi la méthode de Canny [Canny 86], qui donne généralement de bons résultats, même avec des images complexes. De plus, elle est très bien adaptée aux images linéaires, son implantation unidimensionnelle étant très facile et plutôt rapide. Elle peut même devenir très rapide si on utilise la version récursive qui en est proposée par [Deriche 87].

Le signal est convolué avec la dérivée d'une fonction gaussienne (cf. Figure 4.5). Le résultat est une approximation du gradient de l'image linéaire, dont les maxima correspondent aux points de contour de l'image [Canny 86].

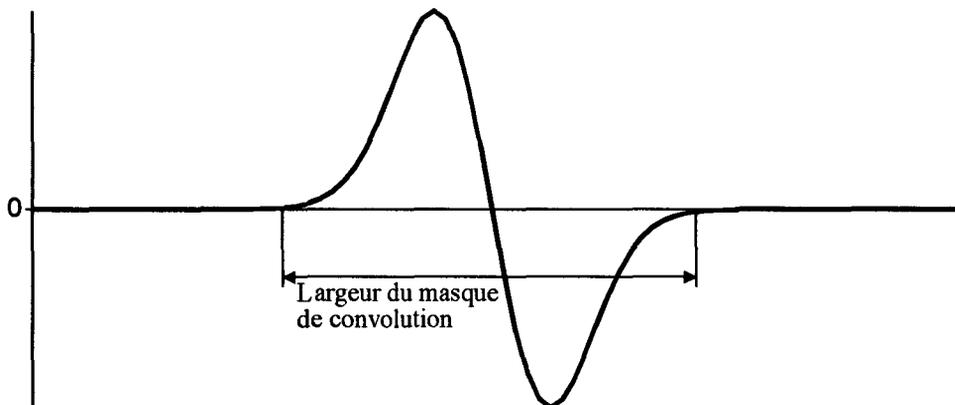


Figure 4.5 : Fonction dérivée d'une gaussienne, convoluée avec l'image pour la détection de contours.

On remarquera que la convolution avec une fonction gaussienne réalise implicitement une moyenne pondérée, qui correspond au lissage défini plus haut.

La largeur du masque de convolution, liée à l'écart-type de la fonction gaussienne, influe sur la résolution: un masque étroit fait apparaître de nombreux contours locaux, mais cache la structure générale. Un masque large laisse apparaître les contours les plus représentatifs, mais élimine les détails fins, nécessaires à la compréhension de la scène.

Des méthodes pyramidales, combinant des masques de différentes largeurs (cf. Chapitre 1), permettraient de combiner le respect des détails fins de la scène et la perception de ses structures essentielles. Nous avons cependant préféré nous en tenir à un masque de largeur fixe, qui réduit notablement le temps de calcul, ce qui est essentiel pour notre application. Après de nombreux essais sur des images réelles (cf. Chapitre 5), la largeur a été fixée à 5 pixels.

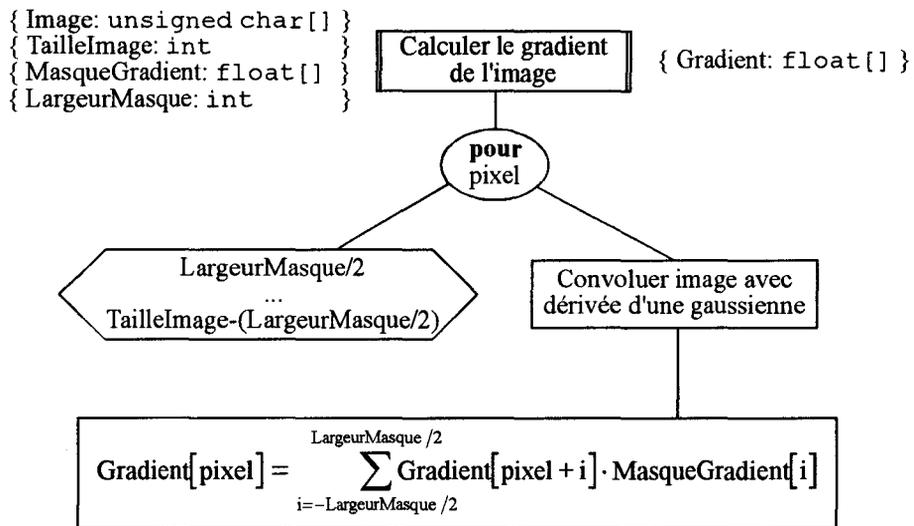


Figure 4.6 : Calcul du gradient de l'image linéaire selon Canny

### Extraction des points de contour

Sont considérés comme points de contour tous les points de l'image où la valeur du gradient présente un extremum local. Pour éviter que des maxima de trop faible amplitude, associés *a priori* au bruit de fond, soient pris pour des points de contour, une valeur absolue minimale du gradient est imposée.

Signalons une petite difficulté, due à la quantification du niveau de gris par la chaîne d'acquisition. Si, en continu, un extremum local peut être défini simplement:

$$\exists E_1 < 0, \exists E_2 > 0 \quad \text{tq} \quad |g(x)| > |g(x + \varepsilon)|, \forall \varepsilon \in ]E_1, E_2[$$

sur un signal échantillonné, les extrema se traduisent souvent par l'apparition de plateaux (cf. Figure 4.7), que l'on trouve par exemple sur la Figure 4.3. Dans ce cas, le point de contour est par convention localisé au centre du plateau. D'où la nouvelle définition:

$$\exists k \in \mathbb{N}^* \begin{cases} |g_n| = |g_{n+i}|, \forall i \in ]-k, +k[ \\ |g_n| > |g_{n+k}| \\ |g_n| > |g_{n-k}| \end{cases}$$

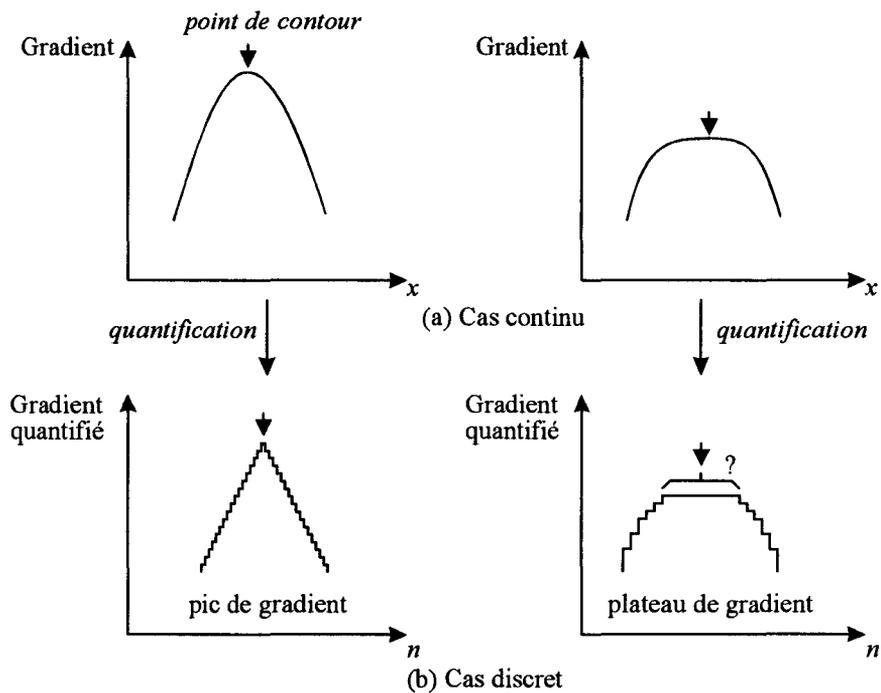


Figure 4.7 : Comparaison des points de contours dans les cas continu et discret.

### Attributs des points de contour

Chaque primitive de type point de contour possède des arguments qui seront utiles lors de la mise en correspondance, pour calculer le critère de ressemblance d'un couple de primitives gauche/droite et pour déterminer si les contraintes d'appariement sont satisfaites ou non. Ces attributs sont de types géométriques et photométriques.

Les attributs géométriques sont limités à la seule position dans l'image, c'est à dire l'abscisse du point image sur le capteur. Comme on l'a déjà dit, la vision linéaire ne permet pas de définir d'autres attributs géométriques.

Les attributs photométriques sont plus riches:

- Le critère de ressemblance utilise les niveaux de gris sur un voisinage autour des primitives (cf. § 4.3.1). Ces niveaux de gris sont un attribut de chaque primitive. Notons cependant qu'il est plus simple et plus rapide, du point de vue de l'exécution du programme, de transmettre simplement l'image entière à la procédure d'appariement. C'est pourquoi nous ne retrouverons pas cet argument dans les structures de données liées aux primitives.
- Une contrainte de mise en correspondance (cf. p. 105) utilise le signe du gradient local, indiquant si un point de contour correspond à un front montant ou descendant du niveau de gris (cf. Figure 4.8). Cet attribut prend

la forme d'une variable binaire, indiquant si le front est montant ou descendant.

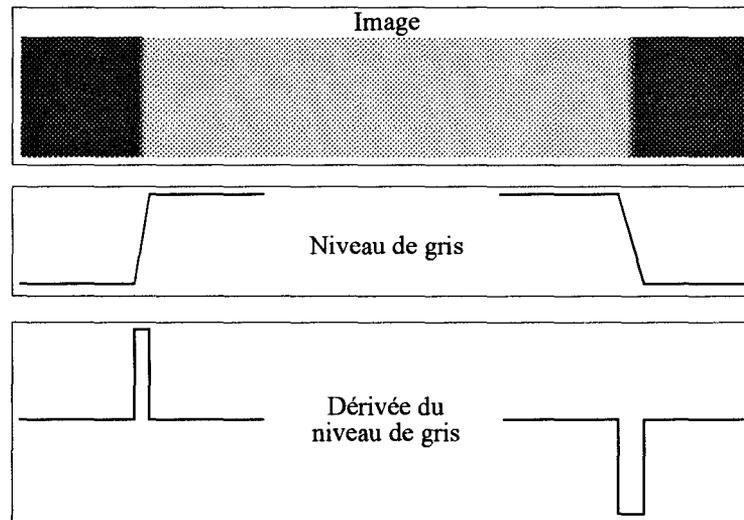


Figure 4.8 : Distinction entre fronts montants et fronts descendants.

### Algorithme de détection des points de contour

L'algorithme de détection des points de contours, qui est appliqué séparément à l'image gauche et à l'image droite, est représenté sur la Figure 4.9. Il tient compte des remarques précédentes de la manière suivante:

1. On unifie les notions de *pic de gradient* et de *plateau de gradient* en définissant un pic comme un plateau de largeur égale à 1 pixel (cf. Figure 4.7). Dans ce cas, le point de contour est à la fois le bord gauche et le bord droit du plateau.
2. Chaque pixel est considéré successivement, en balayant l'image de gauche à droite.
3. On détermine d'abord si le pixel considéré est le bord gauche d'un plateau de gradient, tel que défini au point 1. C'est le cas si le gradient en ce point est strictement supérieur au gradient au point situé immédiatement à sa gauche, et supérieur ou égal au gradient au point situé immédiatement à sa droite. On notera que la prise en compte du signe du gradient complique légèrement ce critère (cf. Figure 4.9). Si un point est le bord gauche d'un plateau de gradient, il est marqué comme tel en plaçant sa position dans une variable appelée *PosGauche*.
4. On détermine ensuite si ce même pixel est le bord droit d'un plateau de gradient. C'est le cas si le gradient en ce point est supérieur ou égal au gradient au point situé immédiatement à gauche, et strictement supérieur au gradient au point situé immédiatement à droite. Comme dans le cas précédent, le critère est légèrement compliqué par le fait que le gradient peut être négatif.
5. Si un point est le bord droit d'un plateau de gradient, une primitive de type *PrimMono* est alors créée, dont la position est le centre du plateau, soit la

moyenne arithmétique de la position du point courant et de PosGauche.  
 Son type est déterminé par le signe du gradient au point courant.

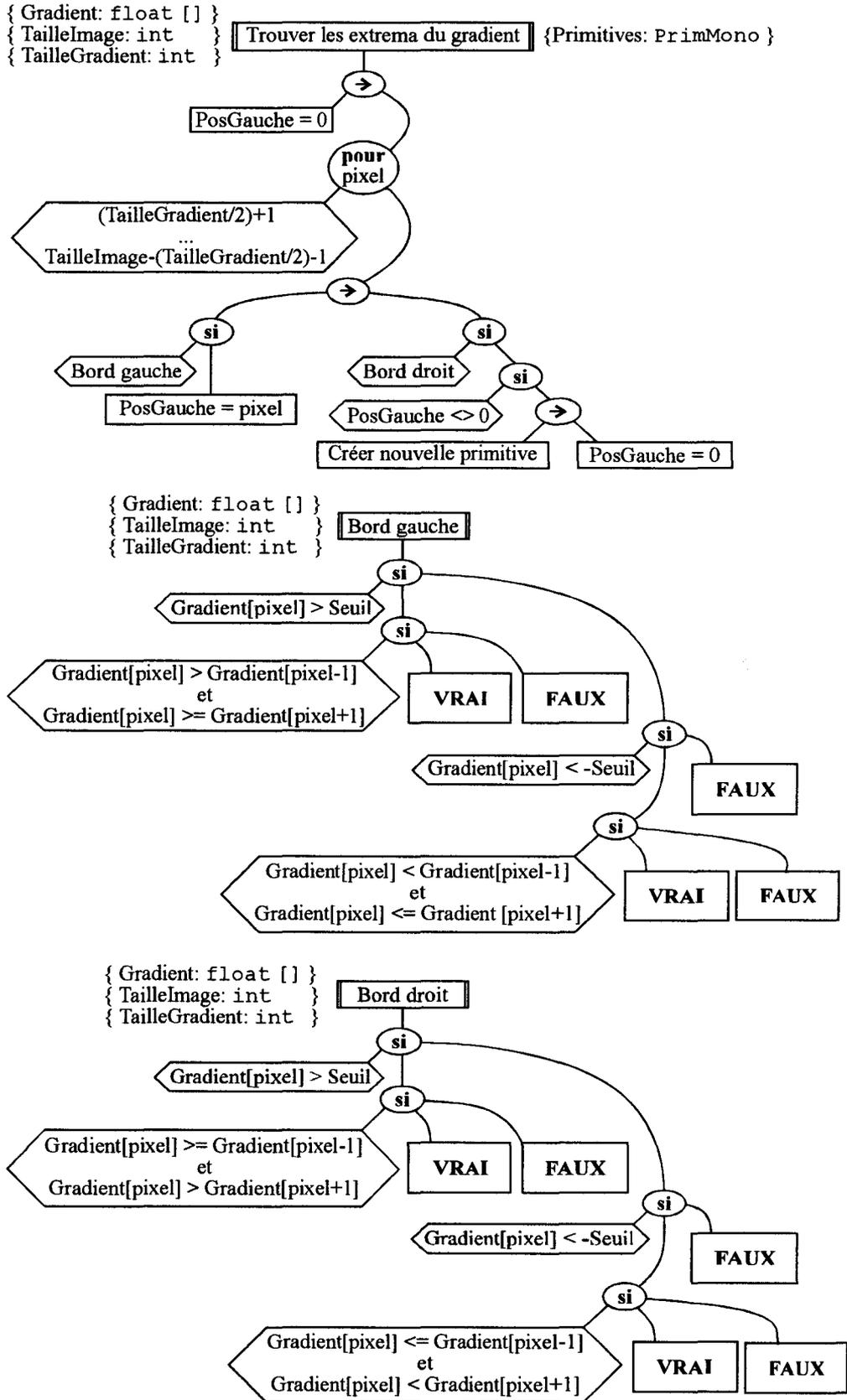


Figure 4.9 : Recherche des points de contour à partir du gradient de l'image.

Quelques précisions permettront de mieux comprendre le fonctionnement de l'algorithme:

1. Comme on l'a vu, un point peut être à la fois un bord gauche et un bord droit. Dans ce cas, il correspond à un pic de gradient (le gradient en ce point est supérieur aux gradients aux deux points voisins). Une primitive est donc créée, dont la position est le point courant.
2. Au départ de l'exécution de l'algorithme (c'est à dire à l'extrême gauche de l'image), `PosGauche` est marquée comme indéfinie en lui donnant une valeur égale à 0. Si le premier bord de plateau détecté est un bord droit, `PosGauche` étant encore indéfinie, alors on ne sait pas si on a affaire à un plateau ou à un épaulement de gradient (c'est à dire une zone où le gradient évolue en  $\nearrow$  ou en  $\searrow$ , marquant une simple pause dans une évolution uniformément croissante ou décroissante). Dans ce cas, on n'a pas affaire à un point de contour et aucune primitive n'est créée.
3. Dans la suite de l'exploration de l'image-ligne, il faut également éviter que des épaulements de gradient donnent lieu à la création de primitives. Dans ce but, `PosGauche` est réinitialisé après la création de chaque nouvelle primitive, indiquant que le point courant n'est pas sur un plateau. La détection d'un bord droit ne donne alors lieu à aucune action.

### 4.2.3. Représentation des primitives en mémoire

Nous pourrions mémoriser les primitives extraites de l'image dans un tableau de taille fixe, ce qui est la méthode la plus simple à programmer. Cependant, cette solution présente un grave inconvénient: comme nous ne connaissons pas à l'avance le nombre de primitives que nous trouverons dans une image, ce tableau devrait toujours avoir la taille qui correspond au nombre maximal de primitives qui pourraient être trouvées dans une image. Or, ce nombre est égal au nombre de pixels de l'image. Un tel gaspillage, qui devrait de plus être appliqué à toutes les structures de données qui dépendent des primitives, entraînerait vite une saturation de la mémoire disponible.

Nous utilisons donc une structure de *liste chaînée*. Une telle structure est créée dynamiquement au fur et à mesure des besoins, et peut être facilement détruite lorsqu'elle n'est plus utile. On n'utilise donc que la mémoire strictement nécessaire pour stocker les éléments de la liste.

Un élément d'une liste chaînée contient, outre les champs classiques d'une structure de type "enregistrement", deux champs supplémentaires qui sont deux pointeurs, indiquant respectivement l'adresse de l'élément suivant et de l'élément précédent dans la liste. Le premier élément de la liste est repéré par le fait que son pointeur sur l'élément précédent est égal à la valeur NULL (pointeur vide). De même, le dernier élément est repéré par un pointeur NULL sur l'élément suivant.

A l'exécution, l'élément courant de la liste est connu du programme par un pointeur. On passe d'un élément de la liste au suivant en assignant à ce pointeur la valeur du champ suivant.

Les primitives sont donc représentées par la structure de données suivante (déclarée ici en langage C):

```
#define UP                1
#define DOWN             -1

typedef struct PRIMMONO
{
    unsigned short  x;           /* position sur la ligne      */
    char            slope;      /* UP si gradient +, DOWN si - */
    struct PRIMMONO *next;      /* pointeur vers pt. suivant   */
    struct PRIMMONO *prev;      /* pointeur vers pt. précédent */
} PrimMono;
```

## 4.3. Mise en correspondance

En vision stéréoscopique matricielle, il existe de nombreuses méthodes de mise en correspondance, utilisant souvent des arguments purement géométriques. En vision linéaire, le manque d'informations géométriques amène à utiliser exclusivement les niveaux de gris de l'image, selon des méthodes basées sur la corrélation des images gauche et droite au voisinage des points de contour. On trouvera des exemples de méthodes de ce type, appliquées à la stéréovision linéaire, dans [Colle 90], [Inigo 85] et [Inigo 87].

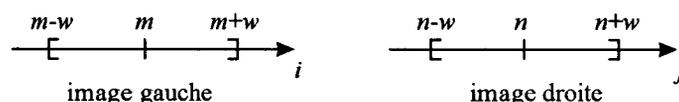
### 4.3.1. Critère de mise en correspondance

Le principe de mise en correspondance consiste à examiner les niveaux de gris sur deux fenêtres de même largeur, centrées sur les points de contour que l'on cherche à mettre en correspondance. Afin d'obtenir une valeur chiffrée de la ressemblance du contenu de ces fenêtres, on calcule le coefficient de corrélation:

$$\rho = \frac{\sum_{i=-w}^w \sum_{j=-w}^w (g_{g_{m+i}} - \bar{g}_g) \cdot (g_{d_{n+j}} - \bar{g}_d)}{\sqrt{\sum_{i=-w}^w (g_{g_{m+i}} - \bar{g}_g)^2 \cdot \sum_{j=-w}^w (g_{d_{n+j}} - \bar{g}_d)^2}}$$

où

- $g_g$  et  $g_d$  sont les niveaux de gris des images gauche et droite.
- $\bar{g}_g$  et  $\bar{g}_d$  sont les valeurs moyennes des niveaux de gris des images gauche et droite sur la largeur de la fenêtre de corrélation.
- $i$  et  $j$  sont les indices des pixels, respectivement sur l'image gauche et sur l'image droite.
- La fenêtre de corrélation est centrée sur le pixel  $m$  de l'image gauche et sur le pixel  $n$  de l'image droite. Sa largeur (nécessairement identique sur les deux images) est  $2w+1$ :



Le coefficient de corrélation est un nombre réel compris entre -1 et 1:

- Une valeur égale à 1 indique une ressemblance parfaite: les variations du niveau de gris dans l'image droite sont exactement identiques à celles de l'image gauche.
- Une valeur négative indique que les niveaux de gris dans les deux images évoluent en sens inverse: l'appariement est alors impossible. Dans ce cas, le coefficient de corrélation ne sera pas calculé puisque, comme nous l'avons vu au § 4.2.2, on mémorise le signe du gradient, autrement dit le sens de variation du niveau de gris au voisinage des points de contour, afin d'éviter de prendre de tels candidats en considération.
- Une valeur nulle indique que l'une des images présente un niveau de gris parfaitement uniforme (variation nulle). Ceci n'arrive jamais lors d'une mise en correspondance, puisque le coefficient de corrélation n'est calculé qu'au voisinage des points de contour, c'est à dire des points autour desquels le niveau de gris varie au contraire de manière très importante.

### 4.3.2. Contraintes de mise en correspondance

Le seul critère de mise en correspondance, basé sur des propriétés locales des différentes images, est insuffisant pour déterminer si deux primitives, de deux images du même couple stéréoscopique, sont engendrées par le même point de la scène. Il existe en effet:

- des relations géométriques que doivent respecter les positions de primitives correspondant aux différentes images d'un même point objet. Il s'agit notamment de la contrainte épipolaire;
- des relations que doivent respecter deux points localisés dans la scène pour être compatibles. Il s'agit notamment des contraintes d'ordre et d'unicité.

De manière générale, le processus de mise en correspondance procède comme suit:

1. Vérification de la contrainte épipolaire, qui permet d'économiser beaucoup de calculs (moins en vision linéaire qu'en vision matricielle): on arrête la recherche dès que les couples de primitives ne respectent pas la contrainte épipolaire.
2. Calcul du critère de mise en correspondance. On ne garde que les couples de primitives pour lesquels ce critère prend une valeur jugée suffisante. On garde souvent plusieurs possibilités concurrentes, parmi lesquelles la suite des traitements permettra de n'en conserver qu'une.
3. Vérification de la compatibilité des appariements possibles avec les appariements déjà acquis dans l'image stéréoscopique. Ceci permet d'éliminer les hypothèses d'appariement qui sont incompatibles pour des raisons de vraisemblance géométrique, et ainsi de résoudre les appariements multiples dans la plupart des cas.

Les contraintes qui interviennent dans la mise en correspondance sont en général:

- La **contrainte épipolaire**: elle exprime la relation qui existe entre les positions des différentes images d'un même point objet sur les différents capteurs (cf. Figure 4.10). La prise en compte de cette contrainte offre un gain de temps appréciable, en éliminant *a priori* les couples de primitives incompatibles, et en évitant de calculer le critère de mise en correspondance sur ces couples.

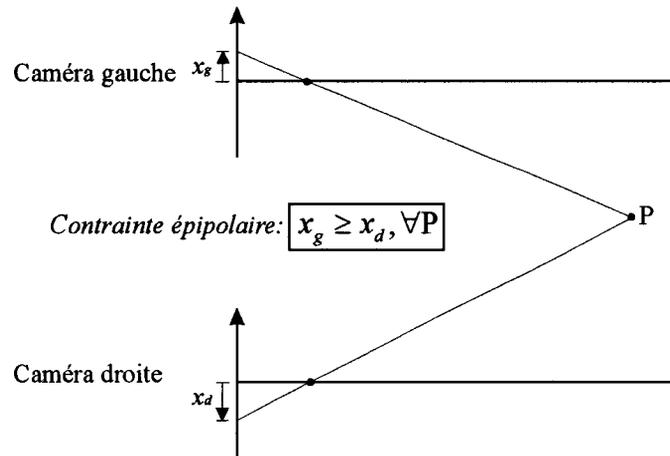


Figure 4.10: La contrainte épipolaire dans le cas de caméras à axes optiques parallèles.

- La **contrainte d'ordre**. Elle exprime le fait qu'un point objet A situé à droite d'un autre point objet B donne naissance par projection sur les deux capteurs à deux points images situés à droite des images créées par le point B. Le raisonnement qui amène à cette conclusion est en tout point semblable à celui qui nous amènera à la contrainte de signe du gradient qui en découle (cf. Figure 4.11). On verra dans ce raisonnement que la seule exception à la contrainte d'ordre est celle d'un objet transparent fortement incliné par rapport à l'axe du stéréoscope, cas très rare dans les scènes naturelles. La contrainte d'ordre est prise en compte implicitement dans toutes les méthodes de mise en correspondance basées sur la programmation dynamique (par exemple [Ohta 85]).
- En vision matricielle, une autre contrainte très utile est la **contrainte d'orientation** [Horaud 93]. Cette contrainte géométrique impose une relation entre les orientations des segments de droite, ou des vecteurs direction des points de contour, des deux primitives gauche et droite que l'on souhaite apparier. Si leurs directions ne respectent pas la contrainte d'orientation, alors elles ne peuvent pas être les deux images d'un même objet. La contrainte d'orientation n'existe évidemment pas en vision linéaire.
- Nous utilisons par contre une contrainte photométrique, introduite dans le cadre de ce travail, que nous appelons la **contrainte de signe du gradient**. Elle utilise l'attribut de signe du gradient des points de contour défini au § 4.2.2, et postule que les gradients locaux au niveau des primitives gauche et droite engendrées par le même point objet sont de même signe.

Un gradient positif indique un front montant (le niveau de gris évoluant du sombre vers le clair lorsque l'abscisse augmente), alors qu'un gradient négatif

indique un front descendant (du clair vers le sombre). Or, un front montant ne peut pas être le correspondant d'un front descendant. Pour s'en convaincre, il suffit d'examiner la Figure 4.11:

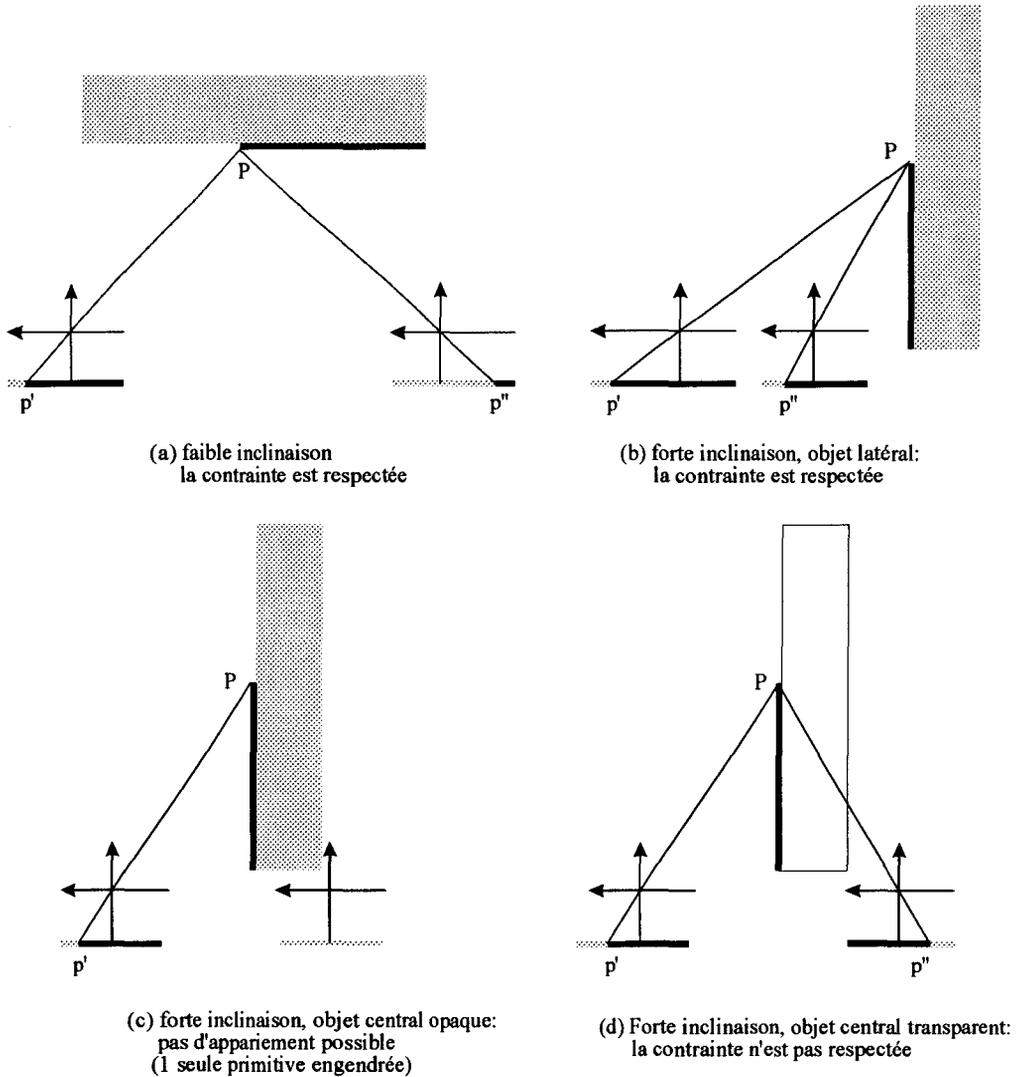


Figure 4.11: Démonstration graphique de la contrainte de signe du gradient.

- Soit un objet faiblement incliné par rapport au plan des capteurs normal à l'axe optique (a). Dans ce cas, les images de cet objet présentent des fronts de même signe, et la contrainte est respectée.
- Soit un objet fortement incliné par rapport au plan des capteurs (b). Dans ce cas encore, les gradients du niveau de gris au niveau du front sont de même signe. La contrainte est encore respectée.
- Soit enfin un objet fortement incliné, mais cette fois placé entre les deux caméras (c). Deux cas peuvent se présenter:
  - Si l'objet est opaque, on a affaire à un cas d'occlusion: le point de contour P n'est vu que par une seule caméra. Une seule primitive lui correspond, et la question de l'appariement ne se pose donc pas.

- Si l'objet est transparent, le point de contour P placé à sa surface est vu par les deux caméras, et les images gauche et droite du point de contour P ont des gradients de signes contraires. La contrainte est donc violée, mais un tel cas est extrêmement rare: nous ne l'avons jamais rencontré au cours de nos essais.

Afin de gagner du temps lors de la mise en correspondance, on évite donc de tenter d'apparier des primitives gauche et droite dont les gradients sont de signes contraires.

Nous ne parlerons pas ici des contraintes mises en oeuvre dans la troisième partie du processus de mise en correspondance, qui permettent de rejeter les appariements incompatibles. Elles seront traitées dans la thèse qui fera suite à celle-ci. Les lecteurs intéressés par ce sujet peuvent se référer au chapitre 6 de [Hraud 93], qui passe en revue plusieurs de ces contraintes.

### 4.3.3. Appariements multiples

La mise en correspondance des points de contours n'est pas toujours facile. En effet, dans une scène stéréoscopique linéaire, pauvre en information, beaucoup d'objets différents peuvent se ressembler, et le meilleur coefficient de corrélation ne correspond pas toujours aux différentes images du *même* point objet. La Figure 4.12 donne un exemple de ce phénomène: sur la largeur de la fenêtre de corrélation, le point de contour n° 8 de l'image gauche ressemble autant au point de contour A qu'au point de contour H de l'image droite.

Dans la grande majorité des cas de ce genre, la valeur du coefficient de corrélation est légèrement plus élevée pour la bonne mise en correspondance. Cependant, lorsque plusieurs objets très semblables se trouvent dans la scène — par exemple, les bandes blanches de la chaussée — de légères altérations des images (reflets spéculaires vus par une seule caméra, réponses photométriques différentes des capteurs CCD, etc.) peuvent suffire à perturber le processus de mise en correspondance: le coefficient de corrélation maximal ne correspond alors plus au couple de points de contour correspondants. C'est le cas sur la Figure 4.12, où le correspondant réel du point 8 (le point A) est rejeté au profit d'un usurpateur (le point H) pour une différence de 0,001 sur le coefficient de corrélation.

Notons qu'il ne s'agit pas ici d'une simple spéculation: de telles erreurs de mise en correspondance se produisent effectivement dans les images réelles qui seront présentées dans le Chapitre 5.

C'est pourquoi la procédure de mise en correspondance ne fournit pas seulement l'appariement le plus vraisemblable, mais tous les appariements possibles: lorsque plusieurs couples de points de contour ont des coefficients de corrélation très proches, ils sont tous placés dans la liste des mises en correspondance. Par exemple, dans la Figure 4.12, les couples {8,A} et {8,H} sont placés dans la liste des mises en correspondance. Le choix du "bon" couple — {8,A} dans notre exemple — sera effectué par la procédure de vérification de la cohérence spatio-temporelle de la scène reconstituée.

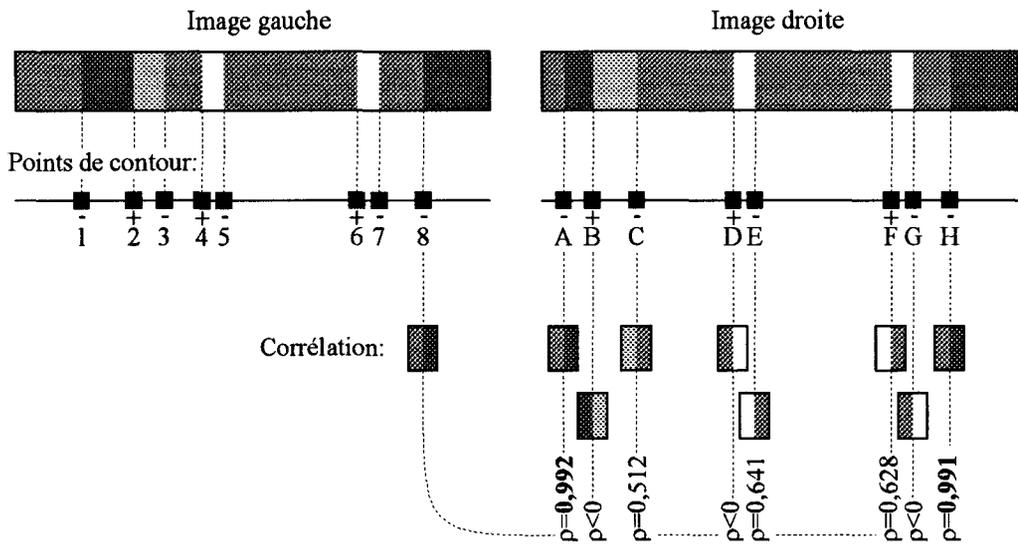


Figure 4.12 : Utilisation de la corrélation pour la mise en correspondance.

#### 4.3.4. Primitives non appariées

Un autre cas doit être pris en considération pendant le travail de mise en correspondance: celui où une primitive donnée ne trouve pas de correspondant dans l'autre image. Ceci arrive en particulier dans le cas de réflexions spéculaires vues par une seule caméra, ou dans le cas d'occlusions, un objet venant en masquer un autre à l'une des caméras (cf. Figure 4.13).

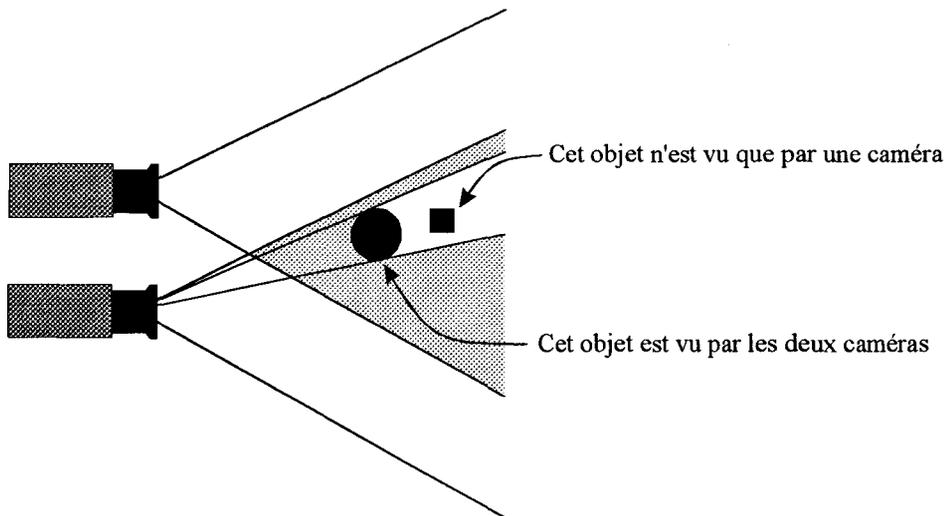


Figure 4.13: Occlusion d'un objet par un autre.

Un autre cas courant est celui où un objet se trouve dans le champ de l'une des caméras, mais hors du champ de vision stéréoscopique (cf. Figure 4.14).

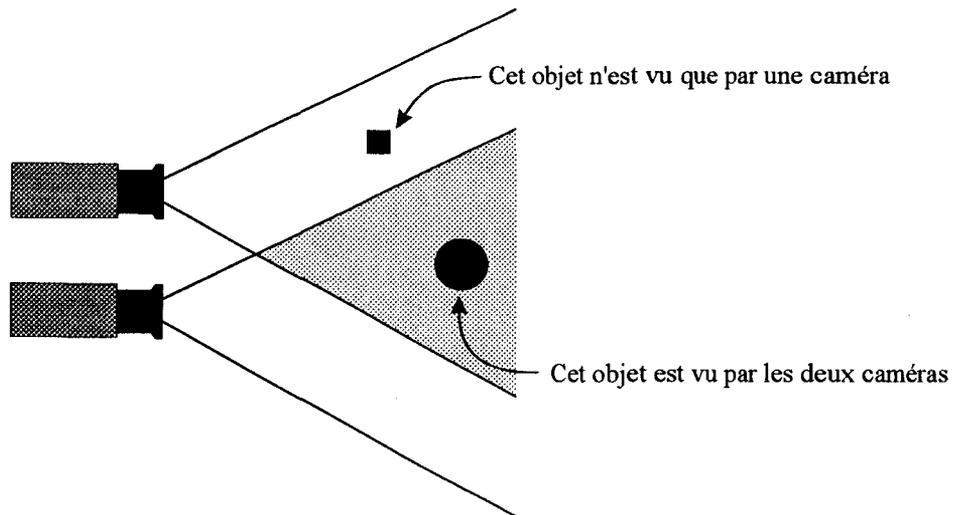


Figure 4.14: Objets placés dans et hors du secteur de stéréovision.

Il est également intéressant de conserver ces primitives non appariées, car elles appartiennent souvent à des objets qui ont été identifiés et localisés dans la scène à partir d'images précédentes, auquel cas la connaissance d'une des deux images de ces objets permet d'estimer leur évolution d'après leur vitesse (déjà connue) et l'évolution de leur seule image monoscopique.

### 4.3.5. Représentation en mémoire

La structure de donnée suivante est utilisée pour mémoriser:

- la liste des appariements trouvés, y compris les appariements multiples;
- les primitives monoscopiques non appariées, de l'image gauche ou droite.

```
typedef struct PRIMSTEREO
{
    unsigned short    xLeft;      /* position sur ligne G      */
    unsigned short    xRight;     /* position sur ligne D      */
    float             corr;       /* coef. de corrélation G/D  */
    float             distance;   /* distance (axe Oz)         */
    float             position;   /* position (axe Ox)         */
    struct PRIMSTEREO *next;     /* pointeur sur pt. suivant  */
    struct PRIMSTEREO *prev;     /* pointeur sur pt. précédent*/
} PrimStereo;
```

On notera de nouveau la structure de liste chaînée, et la présence des deux champs `distance` et `position`, qui recevront une valeur lors de l'exécution de la procédure de triangulation géométrique (cf. § 4.4).

Une primitive non appariée appartenant à l'image gauche est repérée par `xRight = 0`, tandis qu'une primitive non appariée appartenant à l'image droite est repérée par `xLeft = 0`. Dans ces deux cas, les champs `corr`, `distance` et `position` sont indéfinis.

Pour un couple stéréoscopique d'images linéaire, la liste de type `PrimStereo` contient d'abord les couples de points de contour mis en correspondance, classés par ordre croissant d'abscisse sur l'image gauche. Dans le cas d'appariements multiples, les appariements possibles sont classés par ordre

croissant de l'abscisse sur l'image droite. A la fin de la liste se trouvent les points non appariés de l'image gauche, suivis des points non appariés de l'image droite.

Pour l'exemple de la Figure 4.12 (64 couples au total), la liste des appariements `PrimStereo` contient, dans l'ordre, les 12 appariements suivants:

{1,A}{2,B} {3,C} {4,D} {5,E} {6,D} {6,F} {7,E} {7,G} {8,A} {8,E} {8,H}

- Les 28 couples {1,B} {1,C} {1,D} {1,E} {1,F} {1,G} {1,H} {2,C} {2,D} {2,E} {2,F} {2,G} {2,H} {3,D} {3,E} {3,F} {3,G} {3,H} {4,E} {4,F} {4,G} {4,H}, {5,F} {5,G} {5,H} {6,G} {6,H} {7,H} sont éliminés par application de la contrainte épipolaire (cf. § 4.3.2): leurs coefficients de corrélation ne sont pas calculés.
- Les 15 couples {2,A} {3,B} {4,A} {4,C} {5,B} {5,D} {6,A} {6,C} {6,E} {7,B} {7,D} {7,F} {8,B} {8,D} {8,G} sont éliminés d'office par application de la règle sur le sens de variation des niveaux de gris (cf. § 4.3.2): leurs coefficients de corrélation ne sont pas calculés.
- Les 9 couples {3,A} {4,B} {5,A} {5,C} {6,B} {7,A} {7,C} {8,C} {8,F} ont un coefficient de corrélation inférieur au seuil.

## 4.4. Triangulation géométrique

La procédure de triangulation reçoit la liste de primitives mises en correspondance définies au § 4.3, et calcule la distance  $Z$  et la position  $X$  de chaque primitive appariée, en fonction:

- Des positions des deux points objets sur les capteurs gauche et droit.
- Des paramètres de calibration (cf. § 1.4.4), déterminés lors de la phase de calibrage du stéréoscope linéaire (cf. Chapitre 3).

Les équations sont les suivantes (cf. § 2.3.3):

$$\begin{cases} Z = \frac{f \cdot E \cdot N}{l \cdot ((x_g - o_g) - (x_d - o_d))} \\ X = \frac{l \cdot Z \cdot (x_g - o_g)}{f \cdot N} - \frac{E}{2} \end{cases}$$

où:

- $X$  est la position du point objet le long de l'axe  $Ox$  (cf. § 2.3.3). Elle est mémorisée dans le champ `position` de la structure `PrimStereo`.
- $Z$  est la distance du point objet le long de l'axe  $Oz$ . Elle est mémorisée dans le champ `distance`.
- $n_g$  et  $n_d$  sont la position du point image sur le capteur gauche ( $n_g$ ) et sur le capteur droit ( $n_d$ ), mémorisées respectivement dans les champs `xLeft` et `xRight`.
- $o_g$  et  $o_d$  sont la position du pixel central (c'est à dire le pixel sur lequel se projette un point situé le long de l'axe optique) sur le capteur gauche

( $o_g$ ) et sur le capteur droit ( $o_d$ ). Ces paramètres de calibration sont déterminés par la méthode de calibrage décrite dans le Chapitre 3.

- $f$  est la distance focale des objectifs.
- $E$  est l'entraxe du stéréoscope.
- $N$  est la résolution du capteur (le nombre de pixels dans l'image).
- $l$  est la largeur du capteur.

On remarquera que les résultats sont exprimés dans le repère lié au stéréoscope, et non dans celui lié à la scène (cf. § 2.3.4). Ce choix est justifié par le fait que beaucoup de points de la liste `PrimStereo` seront éliminés après vérification de la cohérence spatio-temporelle, que celle-ci peut parfaitement s'effectuer dans le repère lié au stéréoscope, et qu'il est même avantageux de travailler de cette manière car les calculs impliqués par le changement de repère prennent un temps non négligeable. On aura donc avantage à les effectuer sur un nombre de points aussi réduit que possible. Le passage au repère lié à la scène, comme défini au § 2.3.4, est par conséquent renvoyé à la fin du travail de mise en correspondance.

L'algorithme de triangulation géométrique est schématisé sur la Figure 4.15:

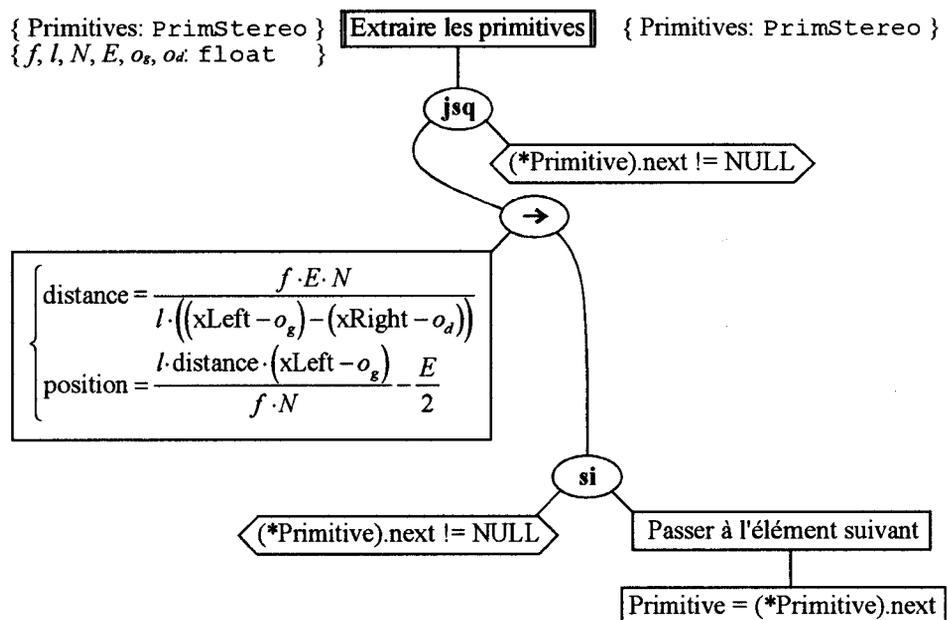


Figure 4.15 : Algorithme de triangulation géométrique.

La hauteur mesurée  $Y$  n'étant pas utile à ce niveau, il n'existe pas de champ correspondant dans la structure `PrimStereo`. Il serait cependant facile de la calculer:

$$Y = Y_0 \cdot \left(1 - \frac{Z}{Z_{\max}}\right)$$

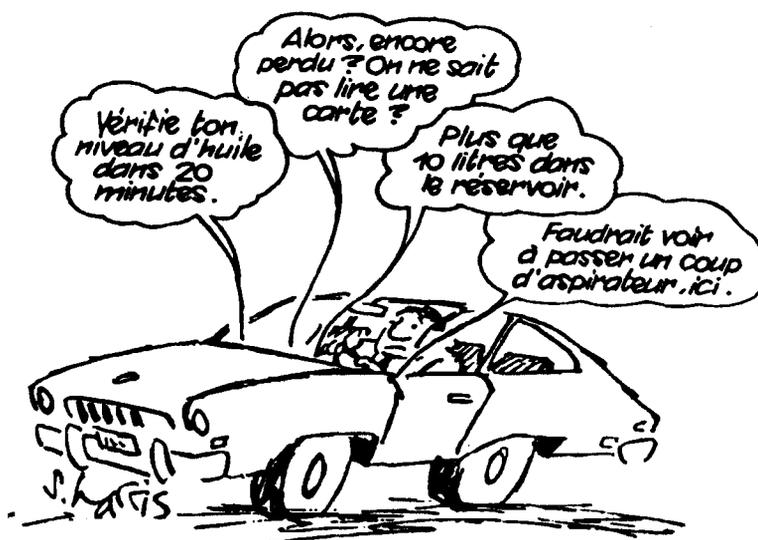
où:

- $Y_0$  est la hauteur des centres optiques des caméras.
- $Z_{\max}$  est la portée du stéréoscope (c'est à dire la distance à laquelle le plan de visée rencontre la route supposée plane).

## **4.5. Conclusion**

Les traitements qui sont présentés dans ce chapitre sont simplement destinés à exploiter les images fournies par le stéréoscope linéaire dans les cas les plus courants. Ils représentent cependant une base pour le développement des traitements définitifs, qui seront implantés dans le dispositif de détection d'obstacle. Ces traitements incluront de plus la reconstruction de la scène tridimensionnelle et une procédure de vérification de la cohérence spatio-temporelle des résultats de la mise en correspondance. De plus, la robustesse des procédures existantes (notamment l'extraction de contours) sera améliorée.

Cependant, les résultats obtenus par la version actuelle avec des scènes enregistrées en conditions réelles sont déjà suffisamment fiables pour assurer une détection d'obstacle dans la plupart des cas, comme nous le verrons au chapitre suivant.



## Chapitre 5

# INTEGRATION SUR UN VEHICULE, ESSAIS SUR ROUTE

---

## 5.1. Introduction

Après la description du stéréoscope linéaire (cf. Chapitre 2), de ses méthodes de calibrage (cf. Chapitre 3) et de traitement d'image associées (cf. Chapitre 4), nous avons maintenant terminé la description de notre dispositif de prise de vue.

Il nous reste cependant quelques détails à résoudre avant de pouvoir le valider pour l'application qui nous intéresse (cf. Chapitre 1).

Tout d'abord, nous devons pouvoir enregistrer les séquences d'images recueillies par le stéréoscope linéaire, afin d'obtenir des données de référence: c'est l'objet du § 5.2.

Nous devons également résoudre un problème spécifique à la vision linéaire: les images linéaires sont illisibles, ce qui rend les séquences enregistrées difficiles à interpréter pour juger de la validité des traitements. Nous avons conçu un dispositif qui permet d'enregistrer, en synchronisme avec la séquence acquise par le stéréoscope linéaire, une séquence vidéo standard. Cette séquence sert ensuite de référence pour l'interprétation. Ce dispositif est décrit au § 5.3.

Enfin, des exemples de séquences d'images, représentatives des conditions réelles d'utilisation du stéréoscope linéaire, sont présentées au § 5.4.

---

## 5.2. Enregistrement des séquences acquises par le stéréoscopique linéaire

Pour les essais du stéréoscope linéaire et la validation des méthodes de traitement d'images associées, il est nécessaire d'enregistrer les images acquises par le stéréoscope linéaire, et ceci pour de nombreuses raisons:

- Au stade des essais, les algorithmes ne fonctionnent généralement pas en temps réel. Dans le contexte de la détection d'obstacles sur route, le temps de réaction admissible ne dépasse pas quelques dixièmes de seconde, alors qu'un algorithme en cours de mise au point, non optimisé, met généralement plusieurs secondes pour traiter une seule image. On doit donc utiliser des images enregistrées, que l'on fournit au ralenti aux algorithmes qui peuvent ainsi traiter les séquences d'images à leur rythme.
- Il arrive souvent que l'on désire comparer les performances de plusieurs méthodes, ou de plusieurs variantes d'une même méthode, dans une situation donnée. Bien entendu, une telle comparaison n'a de sens que si chaque programme testé travaille sur les mêmes données, ce qui implique l'utilisation d'enregistrements représentatifs des situations à traiter.
- Enfin, si nous n'avons pas d'enregistrement de nos images, nous ne pourrions pas montrer de résultats à la fin de ce chapitre!

Pour réaliser les enregistrements, nous utilisons le disque dur du PC comme un magnétoscope, et le stéréoscope comme une caméra vidéo dont le viseur serait le moniteur VGA du PC hôte (cf. Figure 5.1).

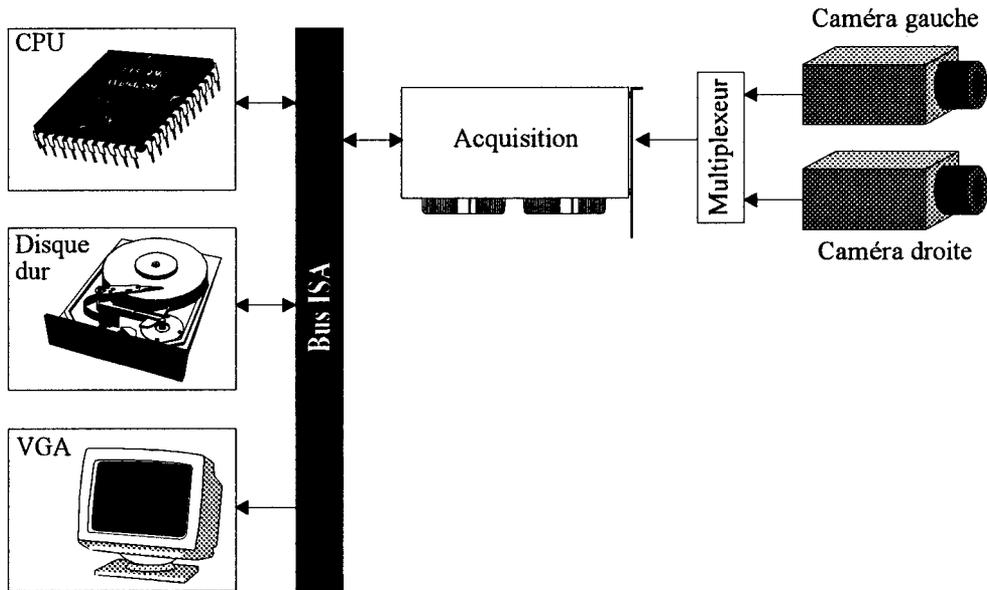


Figure 5.1 : Configuration matérielle pour l'enregistrement de séquences linéaires stéréoscopiques.

### 5.2.1. Caractéristiques du disque dur

Les caractéristiques du disque dur doivent être choisies en fonction des exigences particulières de l'enregistrement de séquences :

- Le temps d'accès doit être court: il faut pouvoir écrire les données au rythme où elles arrivent. Dans notre cas, il faut enregistrer 2 images de 2592 pixels toutes les 40 ms, à raison d'un octet par pixel, ce qui représente un flot de 129600 octets/seconde. Notons que la gestion du stéréoscope, également assurée par le PC, doit être suffisamment rapide pour ne pas freiner cette cadence.

Au sujet du temps d'accès, signalons également que *tous* les modèles de disque dur doivent régulièrement recalibrer leurs têtes pour compenser la dilatation des plateaux due aux variations de température. Ceci se traduit par une indisponibilité momentanée du disque. Certains modèles limitent par une mémoire cache cet inconvénient, parfois jusqu'à le rendre indécélable. D'autres refusent tout service pendant parfois plusieurs secondes [Holzbaur 93]. On voit donc que le choix du modèle de disque dur est loin d'être indifférent.

- La capacité disponible sur le disque, après installation du système d'exploitation et du logiciel de gestion du stéréoscope, doit permettre l'enregistrement de séquences suffisamment longues pour être représentatives de l'évolution dynamique d'une situation réelle sur route. En une minute, les images acquises par le stéréoscope linéaire représentent 7,4 Mo de données, soit 450 Mo pour une heure d'enregistrement. Notons également que toutes les méthodes de compression de données sans pertes que nous avons testées pour augmenter la durée d'enregistrement (notamment LZW [Ziv 77, Welsch 84]) se sont avérées trop lentes pour permettre un enregistrement "en direct" sur le disque dur.

## 5.2.2. Contrôle de prise de vue sur le moniteur de l'ordinateur hôte

L'écran VGA ne permet évidemment pas d'afficher une image stéréoscopique, et les systèmes d'affichage stéréoscopique existant sur le marché, outre leur prix très élevé, sont trop encombrants et trop fragiles pour pouvoir être utilisés à bord d'une voiture. Nous avons donc opté pour le compromis suivant: les images provenant des deux caméras sont affichées simultanément sur l'écran (cf. Figure 5.2). Le balayage de l'écran par les images linéaires, du haut vers le bas, qui avait donné satisfaction pendant le calibrage (cf. Chapitre 3), a été conservé pour contrôler la qualité des prises de vue.

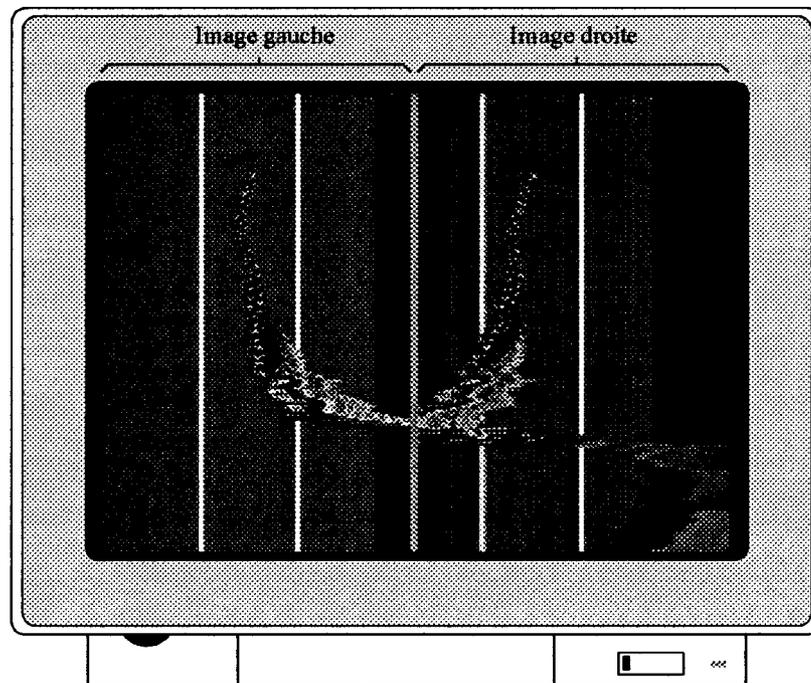


Figure 5.2 : Affichage des images stéréoscopiques linéaires pour la visée.

## 5.3. Enregistrement synchronisé de séquences d'images linéaires et vidéo

### 5.3.1. Nécessité d'un enregistrement vidéo synchronisé avec les images linéaires

La Figure 5.3 représente l'image gauche d'une séquence filmée par le stéréoscope linéaire. On aura vite compris qu'il est inutile de chercher à savoir ce qu'elle représente. Cette absence totale de lisibilité est particulièrement gênante lorsque l'on cherche à valider les méthodes de traitement de ces images: on ne sait pas, même grossièrement, ce qu'elles doivent mettre en évidence!

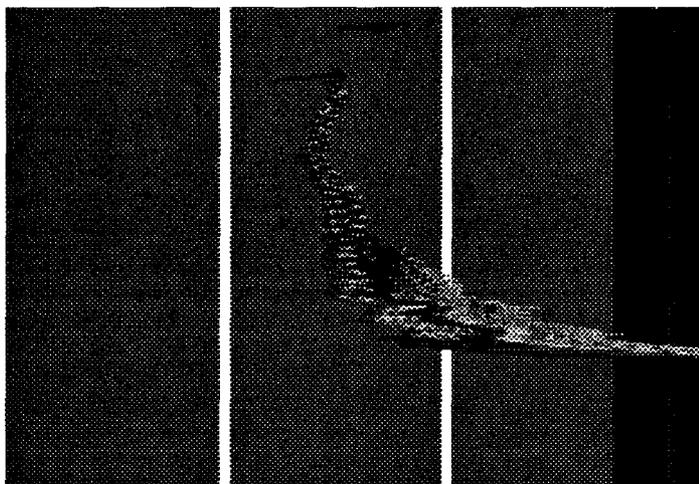


Figure 5.3 : Exemple de séquence enregistrée par le stéréoscope linéaire.

Nous réalisons donc un enregistrement simultané de la scène par une caméra vidéo placée au centre du stéréoscope. A la lecture, cet enregistrement vidéo sert de référence pour interpréter les images linéaires. A l'usage cependant, le manque de lisibilité des images linéaires entraîne de nombreuses erreurs d'interprétation, ce qui provoque inévitablement des erreurs de jugement sur les résultats des traitements.

D'où l'idée de dater les deux enregistrements: en lisant sur l'écran du PC l'instant auquel une image donnée a été prise, nous pouvons rechercher ce même instant sur la bande vidéo, et ainsi retrouver facilement l'image de référence dont nous avons besoin pour interpréter la séquence.

La synchronisation de plusieurs enregistrements, quelque soient leurs supports physiques — et même s'ils sont différents — existe depuis déjà longtemps: elle repose sur l'utilisation du *code temporel de commande*.

### 5.3.2. Le code temporel de commande

Les besoins de la production audiovisuelle ont entraîné le développement, à la fin des années 60, du *code temporel de commande*, plus souvent appelé *time code*, qui permet de synchroniser plusieurs magnétoscopes, magnétophones et/ou tout autre dispositif d'enregistrement ou de reproduction. Les matériels pouvant être synchronisés par ce moyen sont très nombreux, et un dispositif informatique relativement simple peut asservir par code temporel (on dit généralement *time-coder*) tout appareil d'enregistrement ou de lecture.

Le code temporel de commande existe sous deux formes, normalisées par l'Union Européenne de Radiodiffusion [UER 82]:

- Le *code temporel longitudinal*, ou *LTC*, qui permet l'enregistrement sur une piste audio longitudinale. Il peut donc être utilisé sur n'importe quel dispositif d'enregistrement ou de transmission du son;
- Le *code temporel de trame*, ou *VITC*, qui prend place dans la suppression de trame du signal vidéo, et est donc réservé aux magnétoscopes. La norme UER interdit son utilisation pour les échanges de programmes, nous n'en parlerons donc pas davantage.

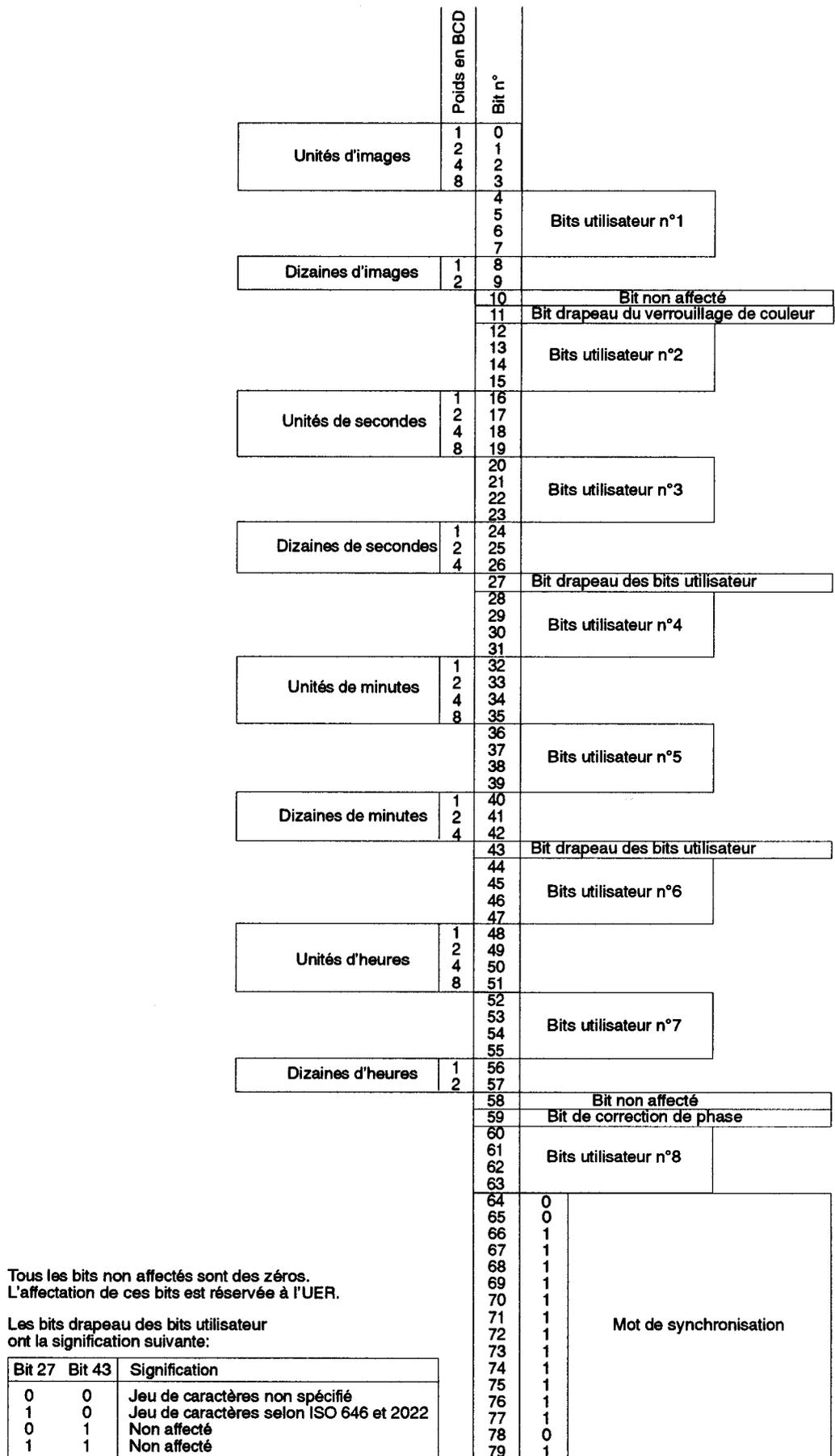


Figure 5.4 : Composition du code temporel de commande de l'UER (code longitudinal).

## Constitution du code temporel de commande

Le principe du code temporel de commande est simple: on enregistre selon des moyens appropriés (piste latérale dédiée, intervalle de suppression trame sur les magnétoscopes, structure de type enregistrement dans les ordinateurs...) un signal numérique qui contient les informations suivantes (cf. Figure 5.4):

- *Heure, minute, seconde, image* (= 1/25 s) de l'image associée. Ces données sont codées en BCD sous forme de 8 groupes de 4 bits.
- Les *bits utilisateur*: 8 groupes de 4 bits dont l'usage est laissé à la discrétion de l'utilisateur.
- Des *bits de service*, dont la signification varie selon le support d'enregistrement.

## Transmission et enregistrement du code temporel de commande

Le code LTC est transmis et enregistré sous forme binaire, en codage biphasé marqué [UER 82] qui conserve une valeur moyenne nulle et intègre l'horloge (2 kHz) au signal. Sa transmission se fait, comme pour un signal audio ordinaire, sur liaison symétrique 600 ohms (niveau 0dBm), ou parfois sur câble coaxial 75 ohms. Bien entendu, il est également possible de le transmettre par téléphone ou par une liaison radio, en tenant compte de l'éventuel délai dû à la transmission.

## Enregistrement synchronisé par le code temporel de commande

L'avantage essentiel du code temporel est d'attribuer un numéro individuel à chaque image, qui peut dès lors être retrouvée par son numéro. Cette possibilité a révolutionné les techniques audiovisuelles, en permettant l'apparition du *montage virtuel* [Beauviala 81], qui a d'abord été mis en oeuvre dans le montage cinéma et vidéo [Villain 91], mais dont l'application la plus spectaculaire est l'interactivité (bornes d'information vidéo, jeux vidéo, enseignement assisté par ordinateur, etc.) [Boursier 93]. Mais le code temporel de commande a d'abord été développé pour permettre la synchronisation de plusieurs enregistrements: en reportant le code temporel issu d'un même générateur sur plusieurs enregistreurs, il est possible de retrouver les différents enregistrements pris au même instant [Beauviala 78]. C'est cette dernière application qui nous intéresse.

Supposons, par exemple, que nous voulions enregistrer l'image sur un magnétoscope, tout en enregistrant le son synchrone sur un magnétophone numérique. C'est possible si l'on enregistre, simultanément sur le magnétoscope et sur le magnétophone, un code temporel issu de la même horloge. A la diffusion, il suffit d'utiliser un dispositif informatique spécialisé pour commander automatiquement, au même instant, le démarrage des deux bandes à partir du même point (cf. Figure 5.5). De plus, la machine vérifie en permanence la coïncidence des deux codes temporels, et corrige toute désynchronisation avant même qu'elle devienne perceptible [Rumsey 94].

Remplaçons maintenant le magnétophone numérique par un stéréoscope linéaire enregistreur... et notre problème est résolu.

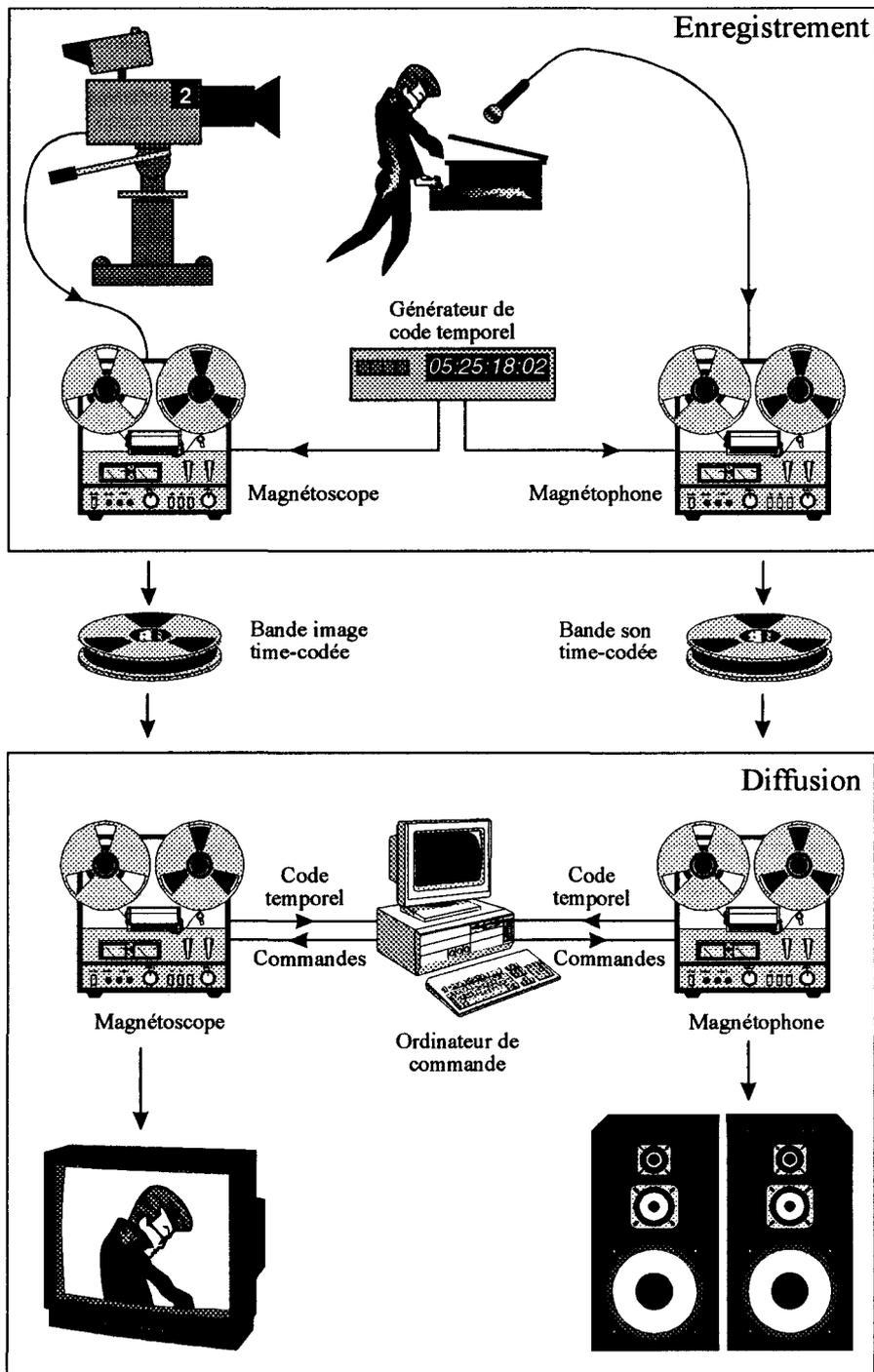


Figure 5.5 : Synchronisation d'un magnéscope et d'un magnéphone grâce au code temporel.

### 5.3.3. Utilisation du code temporel de commande pour l'enregistrement synchronisé de séquences linéaires et vidéo

Bien entendu, les fabricants de semiconducteurs fabriquent des circuits intégrés générateurs/décodeurs de code temporel de commande (chez Sony, référence CX7912 pour le LTC et CX7913 pour le VITC). Cependant, ces circuits ne sont pas disponibles à l'unité et, de plus, aucun matériel n'existe sur le marché pour exploiter le code temporel sur un ordinateur. Nous avons donc utilisé quelques astuces pour enregistrer ce code avec les images du stéréoscope linéaire.

A l'époque déjà ancienne où tous les bancs de montage vidéo n'étaient pas équipés en série du code temporel, certains fabricants proposaient des appareils spécifiques pour permettre aux monteurs d'exploiter les possibilités décrites plus haut. Nous avons pu acheter l'un des derniers exemplaires de l'un de ces appareils: le TCR 3300, fabriqué par For-A, qui dispose d'une interface informatique spécifique à la marque.

Pour réaliser un enregistrement synchronisé, nous disposons d'un magnétope équipé d'un générateur de code temporel (Sony VO 8800 avec carte BKU 706), qui enregistre le code temporel LTC sur la cassette en même temps qu'il le délivre sur une liaison de type audio. Le TCR 3300 le décode, puis le transmet au PC via une carte d'interface spécialement développée à cet effet. Aux instants des transitions du code temporel, correspondant à une nouvelle image vidéo [UER 82], le PC déclenche une acquisition du stéréoscope linéaire. Le couple d'images linéaires est alors enregistré sur le disque dur avec le code temporel correspondant (cf. Figure 5.6).

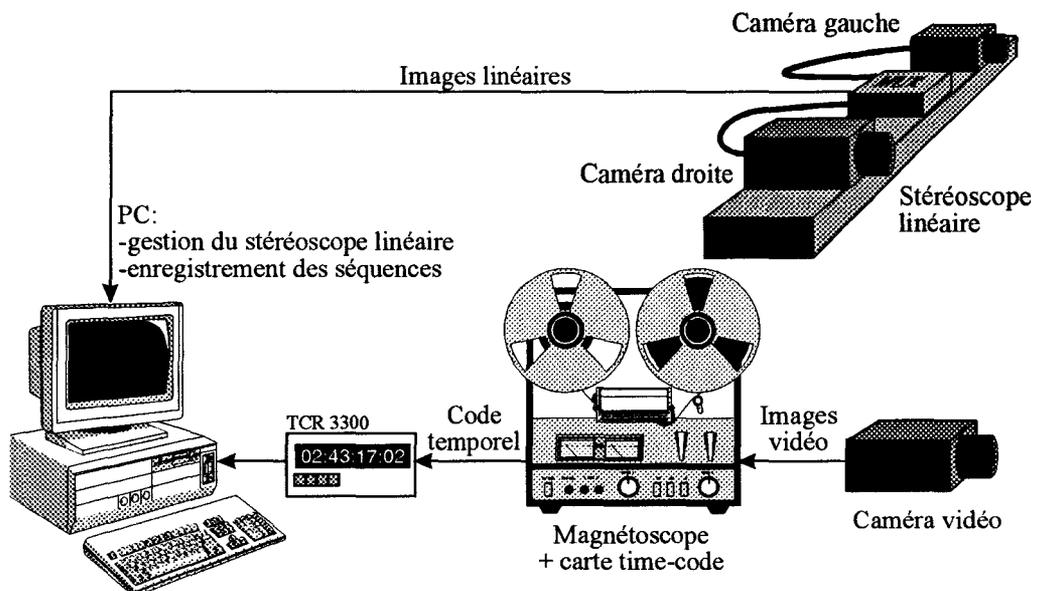


Figure 5.6 : Constitution de la chaîne d'acquisition pour l'enregistrement synchronisé de séquences vidéo et linéaires stéréoscopiques.

## 5.4. Exemples de séquences enregistrées

Les séquences qui suivent ont été prises par beau temps, en ville, sur route de campagne et sur autoroute.

Nous donnons les images gauche et droite de chaque stéréogramme, avec l'interprétation de l'action filmée. Cette interprétation a été retrouvée grâce à l'enregistrement vidéo time-codé. Dans les rares cas où celui-ci n'est pas disponible, l'interprétation provient des notes de l'opérateur.

### 5.4.1. Conditions expérimentales des prises de vue

Pour les essais présentés ici, le stéréoscope linéaire, réalisé comme il a été décrit au Chapitre 2, a été placé sur le toit d'une voiture de tourisme aimablement prêtée par P.S.A. (cf. Figure 5.7). La fixation sur la galerie de toit est assurée par 4 brides fortement serrées pour éviter tout glissement.

Notons que cette disposition ne préjuge pas de l'emplacement définitif qui serait choisi pour un éventuel montage en série. L'emplacement sur le toit n'a été choisi que parce qu'il était le plus pratique pour un essai rapide, et qu'il facilite l'accès aux caméras pour le démontage et pour régler les diaphragmes des objectifs lorsque les conditions d'éclairage varient.

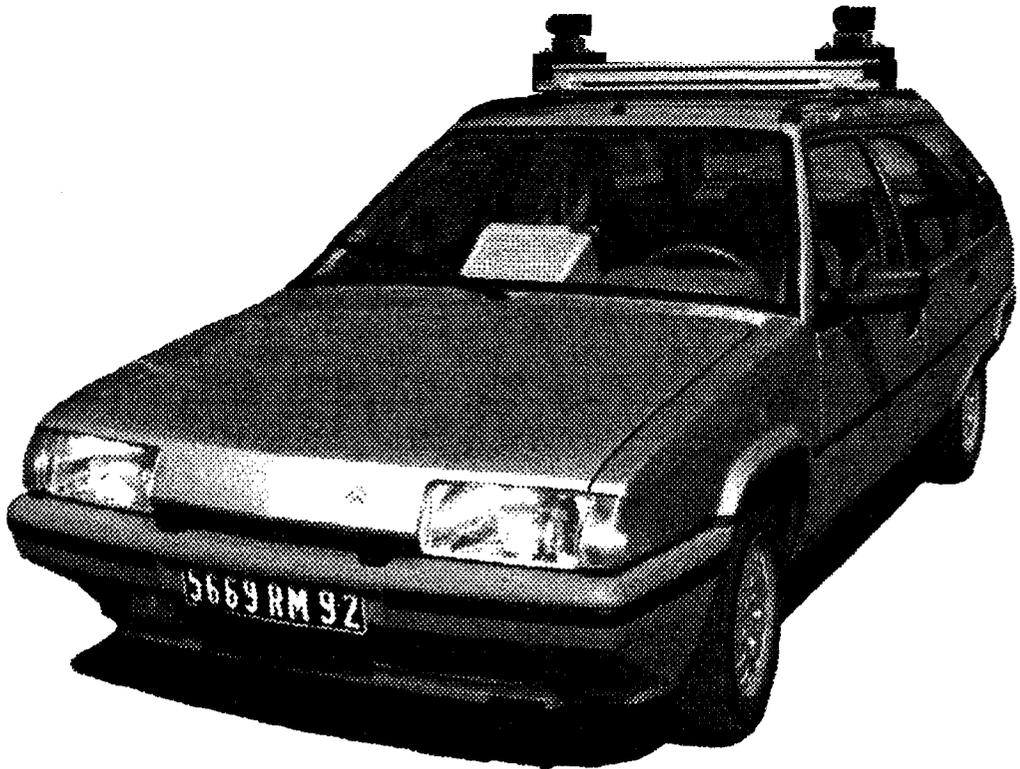


Figure 5.7: La voiture utilisée pour les essais de prise de vue. Noter le stéréoscope linéaire sur le toit, et le moniteur VGA à la place du passager avant.

### 5.4.2. Lecture des séquences stéréoscopiques linéaires

Les séquences se lisent comme suit:

- L'axe horizontal correspond au plan de visée du stéréoscope.
- L'axe vertical est l'axe du temps, dirigé vers le bas de l'image. Pour faciliter la lecture, toutes les séquences imprimées ont la même longueur de 200 lignes. Afin d'enregistrer l'intégralité d'une action (virage, dépassement, etc.) sur une séquence, la durée totale de cette dernière est fixée à 20 secondes. L'intervalle entre deux prises de vue est donc de 100 ms.

La lecture des stéréogrammes n'est en principe possible qu'avec un *stéréoscope*, appareil constitué de deux loupes placées chacune devant un oeil de l'opérateur, de telle sorte que celui-ci regarde l'image gauche avec l'oeil gauche, et l'image droite avec l'oeil droit (cf. Figure 5.8). Ce principe a été exploité pendant très longtemps pour l'interprétation de photographies aériennes; il est également utilisé depuis la fin du 19<sup>e</sup> siècle dans bon nombre de dispositifs de photographie en relief.

Ne pouvant pas fournir un stéréoscope avec ce mémoire, nous avons préféré utiliser un principe connu sous le nom d'*observation en vision croisée*, fatigant pour les yeux, mais ne nécessitant aucun matériel. Nous reproduisons ici la description qui en est donnée dans [Gaillard 89]:



Figure 5.8: Un stéréoscope. Chaque oeil regarde l'image correspondante (image gauche ou droite) à travers une loupe qui renvoie l'image à l'infini. L'observateur a ainsi l'illusion de regarder la scène 3D originale.

*«La méthode d'observation du relief en vision croisée repose sur une constatation très simple: en louchant légèrement, vous vous affranchissez du réflexe de convergence. Vos deux yeux regardent deux endroits différents. Or, c'est là très exactement ce que nous cherchions. On peut objecter qu'en louchant on y voit double, et flou de surcroît. C'est tout à fait exact. Vous voyez double car vos yeux ne convergeant sur aucun endroit précis ne peuvent pas former une image unique. Vous voyez flou parce que votre cerveau n'a pas d'image cohérente sur laquelle faire une "mise au point". Cette mise au point est un autre réflexe oculaire inconscient.»*

*«C'est ce réflexe qui nous permet de voir toujours net, quelque soit la distance qui nous sépare de l'endroit que nous regardons. Mais, supposez maintenant que l'on CROISE les vues du couple stéréoscopique (vue l'oeil gauche placée à droite et vice versa), et que vous les regardiez en louchant légèrement (donc en CROISANT les yeux) tout en vous plaçant à une distance double de votre distance normale de lecture — la largeur d'un couple est double de celle d'une image normale — que va t-il se passer?»*

*«Vous louchez, donc chaque vue du couple va se dédoubler en deux images floues. En louchant ni trop ni trop peu, les deux doubles vont exactement se superposer entre les deux vues initiales du couple. A ce moment précis, chacun de vos yeux ne garde que la vue qui lui est destinée, et celle-ci uniquement. Vous verrez donc trois images floues, celle du milieu étant virtuelle.»*

*«Observez attentivement l'image centrale floue sans vous approcher (vous devez impérativement vous placer à une distance double de la normale) et attendez quelques instants. Cette image centrale virtuelle que vous observez est cohérente: elle est la synthèse de deux vues stéréoscopiques exactement comme une vision naturelle. Votre cerveau, trompé, va donc tout naturellement déclencher le réflexe d'accommodation pour vous donner, au bout de quelques secondes au début puis, immédiatement, dès que vous aurez assimilé cette petite gymnastique oculaire, une image nette, (...) en relief, que vous pourrez même détailler à loisir! En effet, les yeux restent couplés même en vision croisée.»*

*«La majorité des personnes assimilent la technique de la vision croisée en moins de 6 à 7 essais (...).»*

Nous avons inversé les images gauche et droite de chaque stéréogramme pour permettre au lecteur d'essayer cette technique. Nous lui conseillons de se placer à plus d'un mètre de distance pour les premiers essais.

Pour pouvoir augmenter leur taille, et donc les rendre plus lisibles, les images stéréoscopiques sont imprimées verticalement. Pour les regarder, il suffit de tourner ce rapport de 90°, de façon à avoir le côté droit vers soi.

Les séquences présentées sont accompagnées d'un texte de description et d'une figure d'explication, afin d'aider le lecteur à les interpréter. Nous lui suggérons de s'essayer à déchiffrer lui-même les séquences avant de lire les explications: cette petite gymnastique intellectuelle permet de mieux comprendre ce qui se passe au cours du temps, et facilite l'interprétation visuelle des résultats des traitements.

### 5.4.3. Détection et localisation d'un piéton

La séquence représentée sur la Figure 5.9 est la première qui ait jamais été enregistrée par le stéréoscope linéaire; elle a surtout été utilisée pour apprécier ses qualités dans des conditions de prise de vue simplifiées par rapport aux cas réels.

Les caméras sont montées sur un support fixe, et visent la surface de la chaussée à une distance de 50 mètres. Les lignes blanches, clairement visibles, ont une largeur de 14 centimètres. Une voiture située hors du champ projette une ombre, visible sur la droite des images comme une zone entièrement noire.

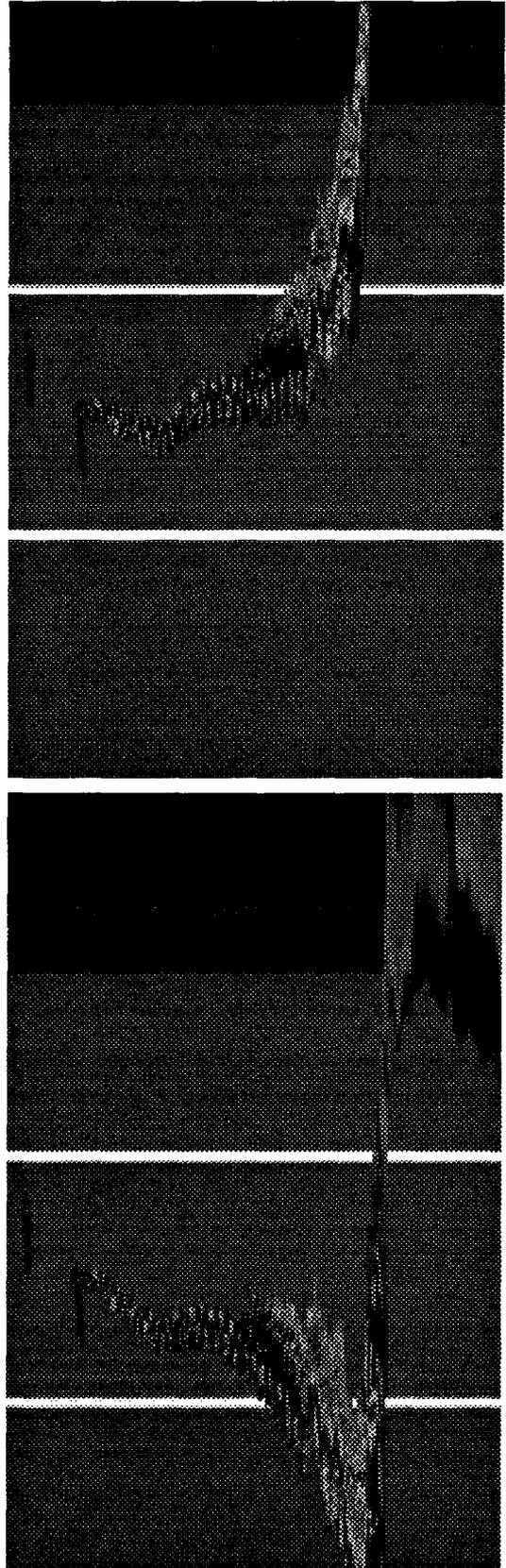


Figure 5.9: Un piéton, vu par le stéréoscope linéaire.

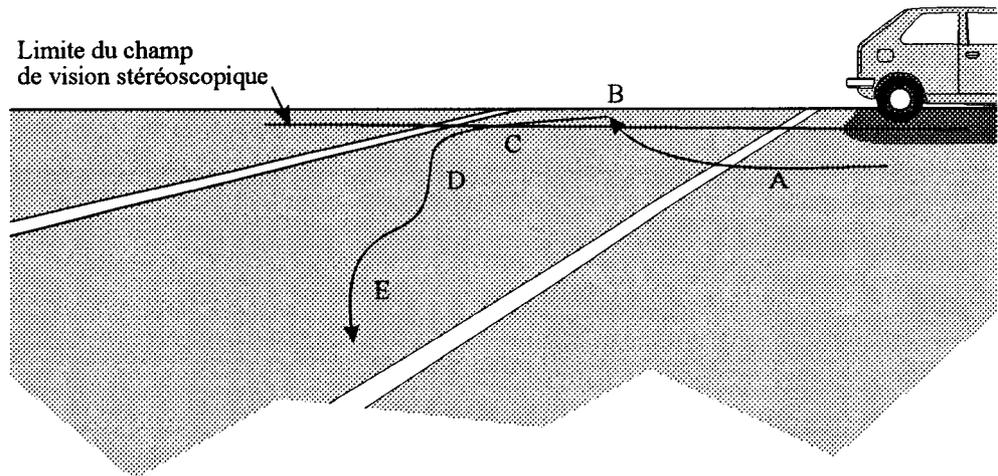


Figure 5.10: Les déplacements du piéton durant la séquence de la Figure 5.9.

La Figure 5.10 décrit schématiquement les déplacements du piéton au cours du temps. On le voit d'abord se diriger vers son repère de départ, situé un peu au-delà de l'intersection du plan de visée et de la chaussée (A). Ce faisant, il passe hors de portée du stéréoscope linéaire et, en conséquence, disparaît de l'image (B). 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 champ de vision stéréoscopique (E).

La Figure 5.11 montre le résultat de la première étape des traitements stéréoscopiques: la détection des primitives, définies comme les points de contour, dans les images issues des deux caméras. L'algorithme distingue deux types de points de contour: les fronts montants et les fronts descendants (cf. § 4.1.2). Cependant, pour conserver une bonne lisibilité, la Figure 5.11 montre les deux types sans les distinguer.

On constate une bonne détection des points de contour de l'image, avec cependant quelques omissions sur des contours à contraste faible (ombre d'une voiture sur une chaussée sombre). Les bords des lignes blanches sont parfaitement détectés, ainsi que le piéton, pourtant vêtu de couleurs plutôt sombres (chaussures marron et blue-jean). Certains contours très peu marqués correspondent à des irrégularités de la chaussée. Ils n'apparaissent pas clairement à l'oeil, et sont détectés de manière erratique, notamment à l'extrême droite de l'image et au voisinage des bandes blanches. Ces contours peuvent poser des problèmes de mise en correspondance, car ils ne sont souvent détectés que sur l'une des deux images à un instant donné. Cette remarque justifie les précautions décrites au Chapitre 4.

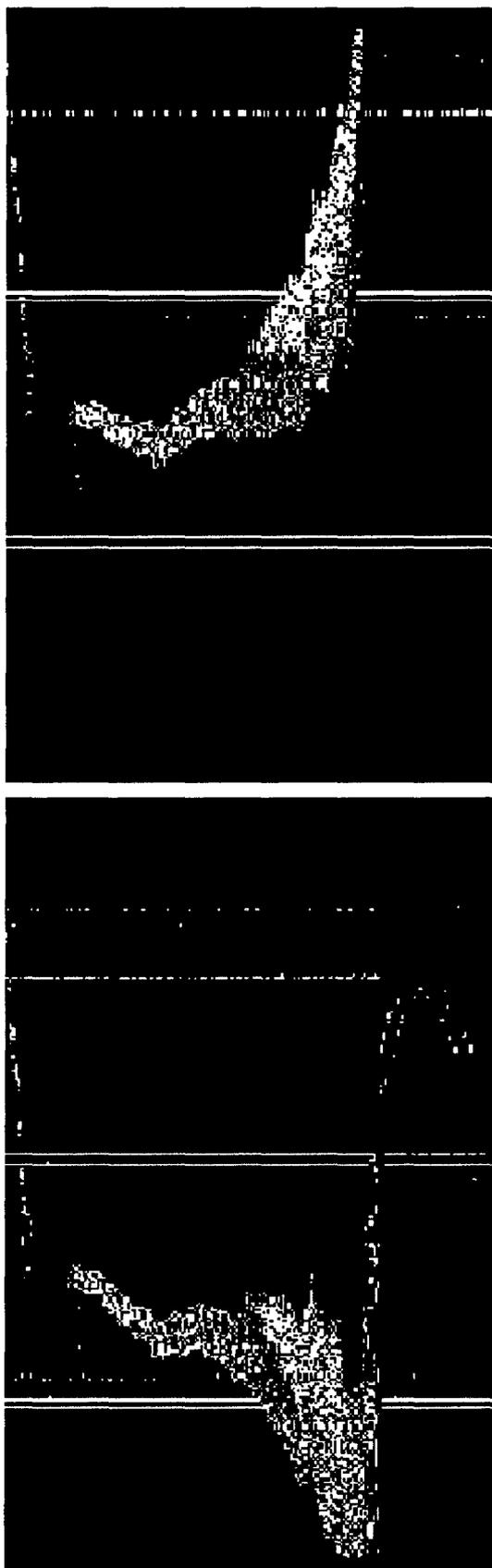


Figure 5.11: Les points de contours extraits de la séquence précédente.

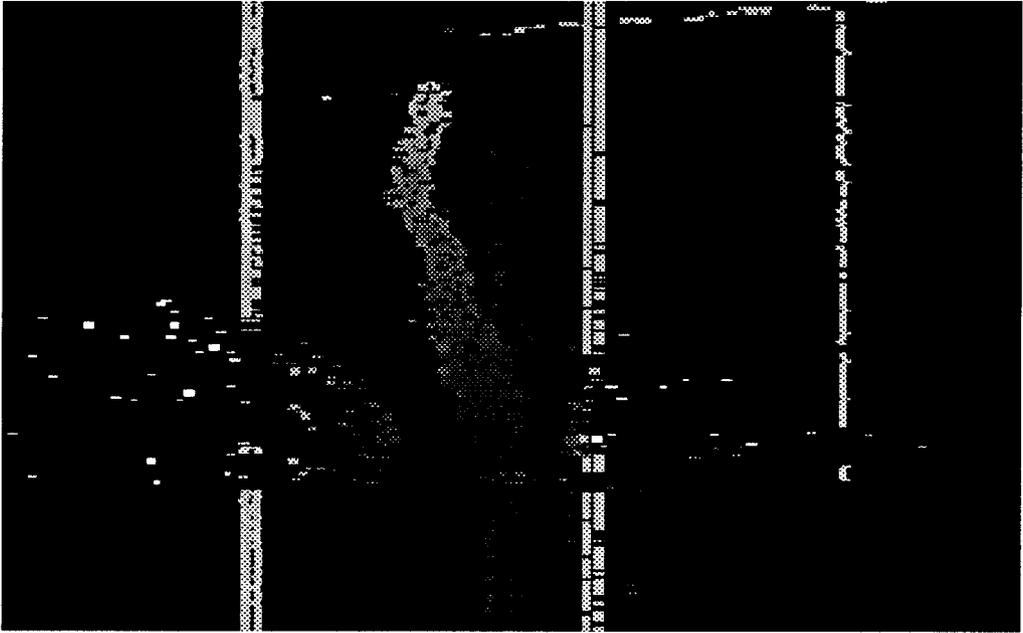


Figure 5.12: Scène reconstituée par les traitements.

La Figure 5.12 montre les résultats de la mise en correspondance des primitives, y compris les appariements multiples (cf. Chapitre 4) qui sont généralement vus comme des points très proches des caméras, situés presque dans l'axe du stéréoscope linéaires. L'ensemble de ces points donne naissance aux lignes très sombres visibles au centre de l'image.

Les axes sont orientés comme suit:

- L'axe horizontal représente l'abscisse dans la scène. Un point situé sur le bord gauche de l'image est à 5 m à gauche de l'axe optique du stéréoscope linéaire. Un point situé sur le bord droit de l'image est à 5 m à droite de ce même axe.
- L'axe vertical représente le temps. Le haut de l'image représente les premières prises de vue, tandis que les dernières sont en bas.
- La distance (mesurée le long de l'axe optique du stéréoscope linéaire) est représentée par le niveau de gris, d'autant plus clair que la distance par rapport au stéréoscope est plus grande, selon la formule:

$$g(d) = g_{\max} \left( 1 - \left( d_{\max} - d \right) \right)$$

où:

- $g_{\max}$  est le niveau de gris maximal, correspondant au blanc pur.
- $d_{\max}$  est la distance maximale représentable. Tout point dont la distance  $d$  est supérieure ou égale à  $d_{\max}$  est affiché en blanc.
- Pour cette séquence,  $d_{\max}$  est égale à 70 m.

Les résultats sont très bons tant qu'il n'y a pas d'occlusion dans l'image. Dans ce cas, la mise en correspondance se fait sans erreur, avec peu de mises en correspondances multiples ou de primitives non appariées. Lorsque des telles erreurs apparaissent, elles ne se reproduiront pas sur les couples stéréoscopiques suivants. Elles pourront donc facilement être éliminées par analyse de la cohérence temporelle des résultats de la mise en correspondance.

Lorsque des occlusions se produisent, c'est à dire que l'un des deux points images correspondant à un point objet disparaît, des appariements erronés peuvent apparaître entre le point objet restant visible et de nombreux autres points objets.

Ceci arrive dans le bas de l'image, lorsque le piéton masque l'une des bandes blanches (ou les deux) à l'une des caméras. Dans ce cas, des appariements erronés apparaissent entre la bande blanche et le piéton.

Ces appariements erronés ne sont en réalité que des propositions apparaissant dans le cadre d'appariements multiples. N'oublions pas, en effet, que l'image de la Figure 5.12 représente l'ensemble de la liste des appariements définie au Chapitre 4, et que cette liste contient tous les appariements *vraisemblables* au sens du critère d'appariement (cf. § 4.2.1) et des contraintes (cf. § 4.2.2). Chacun de ces appariements est vraisemblable au sens du critère de mise en correspondance, et la procédure de mise en correspondance, encore incomplète, ne peut pas déterminer quel appariement correspond aux deux images d'un même point objet.

Les appariements multiples seront résolus lors de l'analyse de la cohérence temporelle (car les "mauvaises" solutions n'apparaissent souvent qu'une seule fois au cours du temps) ou lors de la reconstitution de la scène tridimensionnelle, au cours de laquelle seront détectées des incohérences spatiales. Si aucun des appariements proposés n'est valable, en particulier en cas d'occlusion, ils seront tous éliminés.

#### 5.4.4. Route bien dégagée - aucun obstacle

Cette séquence est représentative d'une chaussée sans le moindre obstacle à détecter.

Nous passons une sortie d'autoroute: le triangle gris que l'on voit sur la droite est le marquage au sol qui délimite cette sortie, tandis que le ruban gris est la voie qui quitte l'autoroute. Le contraste de l'image est important en raison de la présence de lignes blanches, mais aucun relief n'apparaît: tout ce que l'on voit sur cette séquence est situé dans le plan de la chaussée.

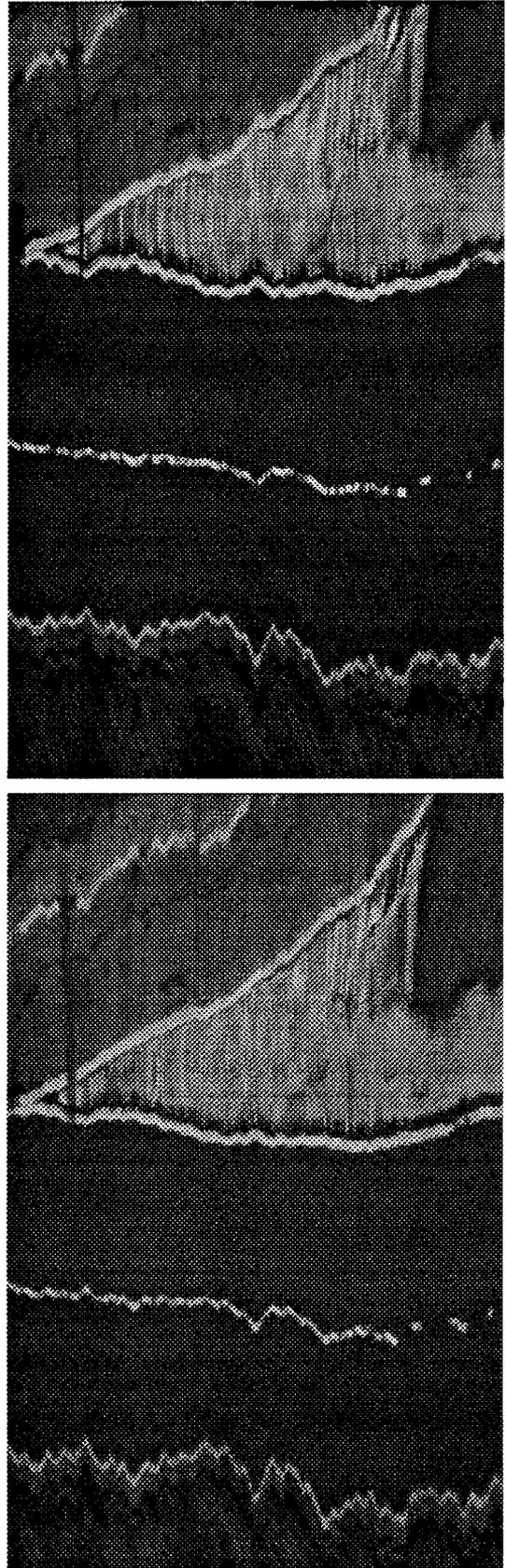


Figure 5.13: Passage d'une sortie d'autoroute.



Figure 5.14: Scène reconstituée à partir de la séquence de la Figure 5.13.

La scène reconstituée (cf. Figure 5.14) montre nettement les bandes blanches, qui sont les seuls éléments visibles dans la séquence de la Figure 5.13. On peut faire plusieurs remarques sur ces résultats:

- Des bandes verticales très sombres sont visibles au centre de l'image, qui correspondraient à des objets très proches des caméras. Ces bandes correspondent en réalité à des appariements multiples, comme on l'a vu au § 5.4.3. Ces appariements multiples seront résolus dans l'avenir, lorsque la procédure prévue à cet effet sera développée. Actuellement, ils sont tous représentés dans la scène reconstituée.
- La scène reconstituée présente une bonne stabilité dans le temps: les objets existants sont détectés en permanence (sans "trous") et leur position calculée reste stable. Ceci est essentiel pour permettre la vérification de la cohérence spatio-temporelle (cf. Chapitre 4). De plus, on constate peu de fausses détections, qui se présentent comme des points isolés, répartis au hasard, et qui seront donc éliminés par cette même procédure.
- La cohérence visuelle entre la scène reconstituée (cf. Figure 5.14) et la scène originale (cf. Figure 5.13) est très bonne. Ceci peut être considéré comme une validation empirique de la pertinence des primitives extraites de la scène, et de la qualité de leur mise en correspondance.

### 5.4.5. Influence du tangage sur la détection

Nous roulons sur une autoroute parfaitement plane, quand nous passons sur un raccord de bitume, surépaisseur de la chaussée qui provoque un cahot. Ce cahot provoque un mouvement de tangage de grande amplitude, visible par la brutale variation de la largeur de l'image de la chaussée, qui indique un changement de la portée du stéréoscope linéaire. Des voitures, situées plus loin que la portée théorique du dispositif, apparaissent brièvement.

L'aspect de l'image lors du cahot est expliqué par la Figure 5.16, la Figure 5.17 et la Figure 5.18: lorsque les roues avant se soulèvent en passant sur une bosse (b), les caméras ont un mouvement vers le haut qui augmente momentanément la portée du stéréoscope linéaire. Des objets situés à la limite de la portée peuvent alors être détectés. L'image de la chaussée sur les capteurs est également moins large qu'avant le cahot (a), comme on le voit sur la Figure 5.17 et sur la Figure 5.18).

Quelques instants plus tard, du fait du mouvement de notre voiture, ce sont les roues arrière qui se soulèvent en passant sur la bosse (c), et les caméras penchent vers le bas, diminuant ainsi la portée. Les objets qui étaient rendus détectables un bref instant redeviennent invisibles au stéréoscope linéaire. La Figure 5.17 et la Figure 5.18 montrent que l'on obtient alors sur les capteurs une image de la chaussée plus large qu'avant le cahot.

Enfin, notre voiture retrouve une chaussée plane (d), et le stéréoscope linéaire reprend sa portée nominale, comme avant le cahot (a). L'image redevient stable.

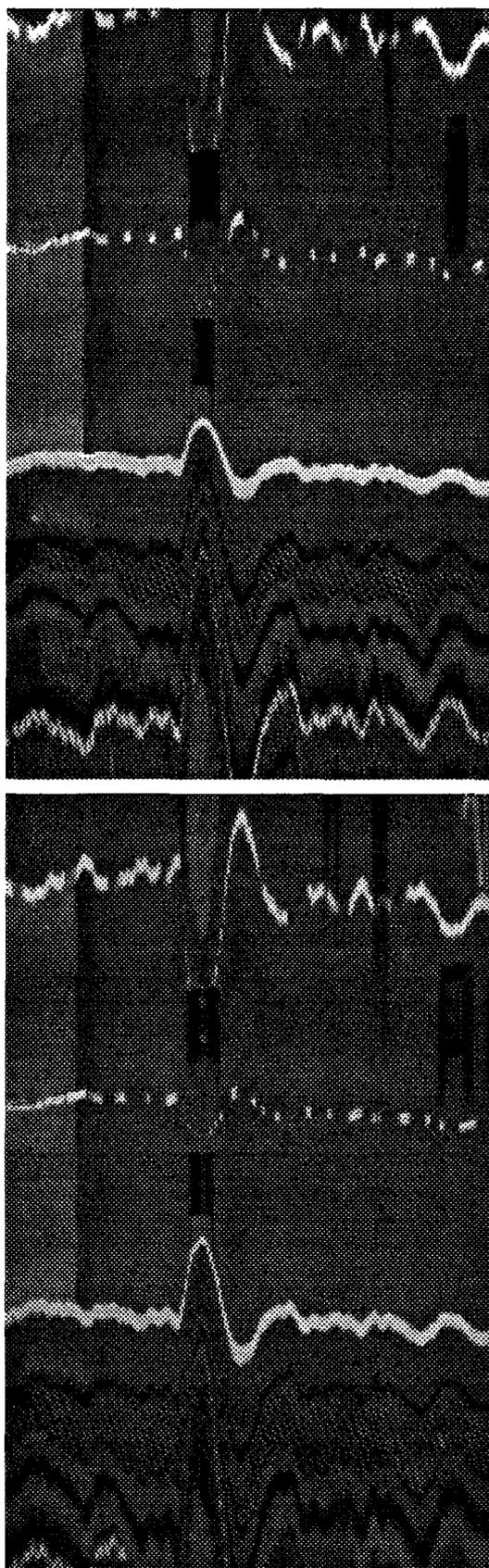


Figure 5.15: Evolution de l'image d'une autoroute durant un cahot.

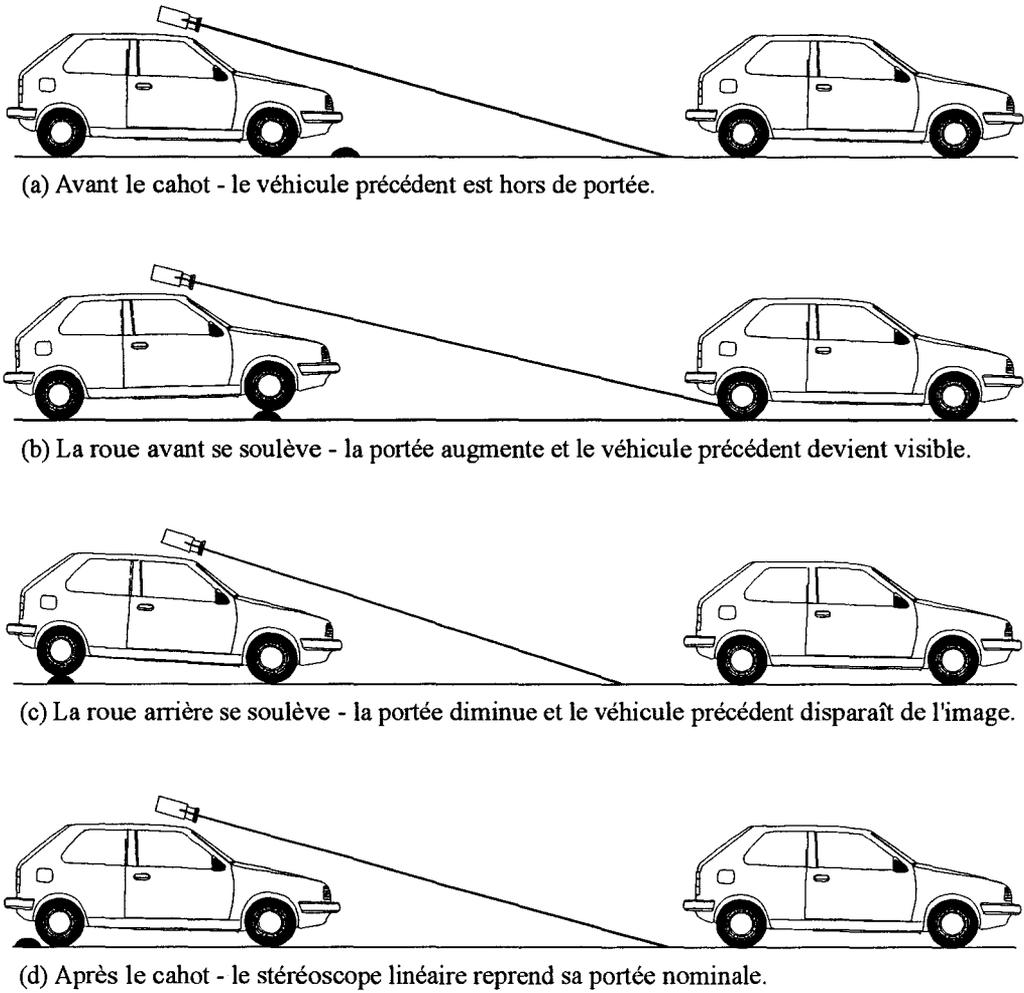


Figure 5.16: Explication de l'aspect de la séquence de la Figure 5.15.

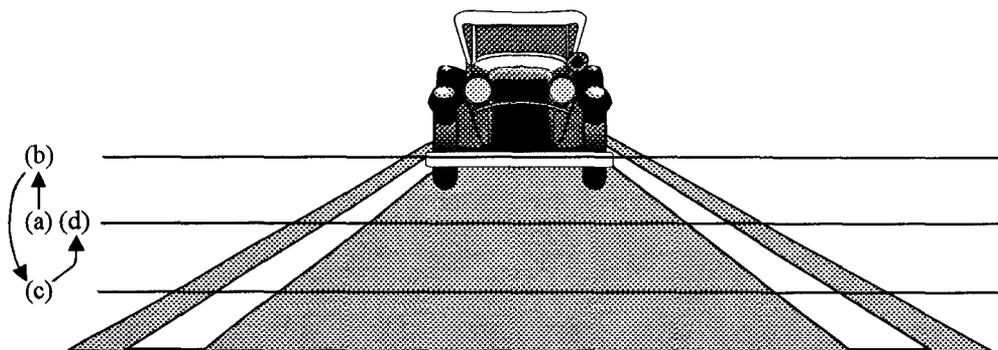


Figure 5.17: La scène de la Figure 5.16, vue dans l'axe du stéréoscope linéaire. Le mouvement du plan de visée au cours du temps est matérialisé par les flèches.

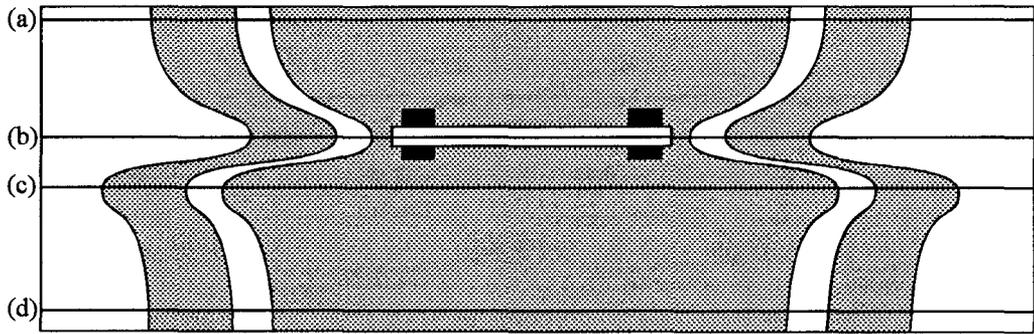


Figure 5.18: Image de la scène de la Figure 5.16, prise par l'une des caméras du stéréoscope linéaire.

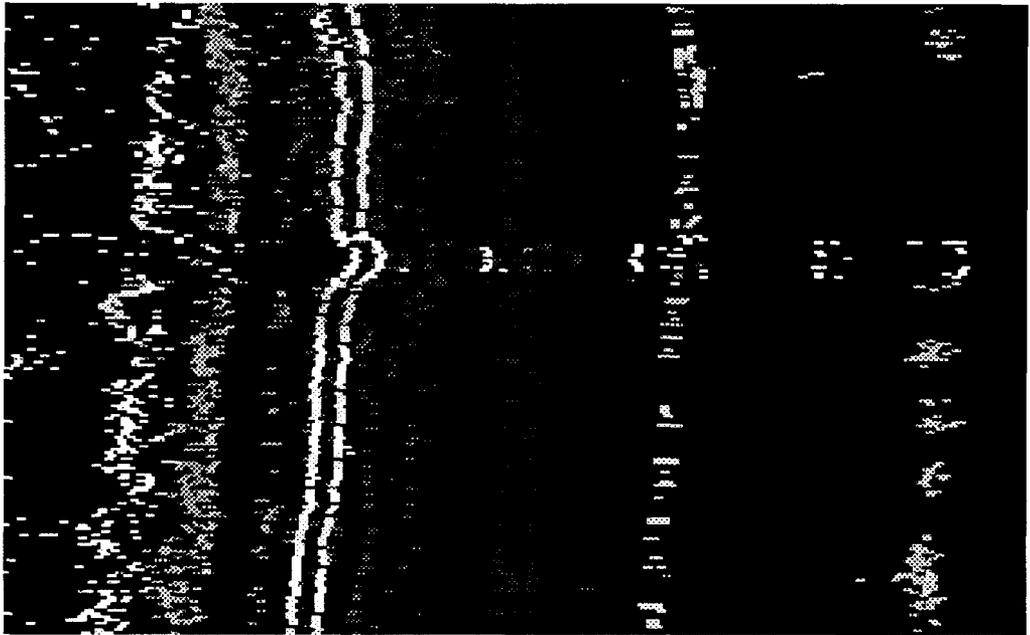


Figure 5.19: Scène reconstituée à partir de la séquence de la Figure 5.15.

La scène reconstituée (cf. Figure 5.19) montre clairement le terre-plein central, la ligne blanche continue qui le délimite, celle qui marque la bande d'arrêt d'urgence — du moins lorsqu'elle est dans le champ des deux caméras — et la ligne blanche discontinue qui sépare les deux voies de l'autoroute. On retrouve ici le bon comportement déjà constaté dans la séquence précédente. Le léger déport latéral trouvé lors du cahot (les bandes blanches se décalent vers la droite) est confirmé par l'examen de la bande vidéo.

La voiture située sur la voie de gauche est bien détectée et localisée, de même que celle située sur la voie de droite, qui mord légèrement sur la bande d'arrêt d'urgence.

Cette deuxième voiture n'est pas correctement détectée lors de sa réapparition (en bas de l'image). Ceci s'explique si l'on compare les deux images prises par le stéréoscope linéaire à ce moment (cf. Figure 5.15): les vibrations des objectifs au moment du cahot ont entraîné un décalage suffisant pour que les images gauche et droite de la voiture soient différentes et ne puissent pas être appariées. On voit ici l'importance de la coplanéité des champs de visée des deux caméras qui composent le stéréoscope linéaire (cf. Chapitres 2 et 3).

Concernant l'origine des vibrations responsables du décalibrage des caméras, nous avons constaté que le modèle d'objectif retenu à l'époque de ces prises de vue était pourvu d'une bague de mise au point extrêmement souple, ce qui est obtenu par un jeu important de la rampe hélicoïdale de mise au point. Cette rampe, solidaire du bloc optique, vibrait sous l'effet des trépidations de la voiture. Nous avons résolu le problème par l'utilisation d'un autre modèle d'objectif, choisi pour la raideur de sa bague mise au point, et par le blocage de cette dernière, obtenu en l'entourant d'une quantité généreuse de toile adhésive appelée *gaffer*, disponible chez les fournisseurs de matériel photographique et cinéma professionnel<sup>1</sup>. Ces deux mesures se sont montrées très efficaces sur la nouvelle version du stéréoscope linéaire, décrit dans les Chapitres précédents.

---

<sup>1</sup> Pour ce type d'usage, nous déconseillons fortement les rubans de toile adhésive vendus par les magasins de bricolage, dont la rigidité — et, partant, l'efficacité — est bien moins bonne, et dont l'adhésif mal équilibré peut arracher le revêtement caoutchouté de l'objectif quand on retire le ruban.

### 5.4.6. Approche et dépassement d'un véhicule

Nous roulons plus vite que la R4 qui nous précède. Les plans de visée croisent l'arrière de la R4 à une hauteur croissante à mesure que la distance décroît. Ce qui explique l'inversion de l'image de la R4, dont on commence par voir les roues pour remonter jusqu'au toit. Noter également son aspect évasé, dû à la distance décroissante.

Devant la R4 se trouve un autre véhicule, en train de se rabattre. Comme on l'a vu au § 5.4.5, la brièveté de l'apparition de ce véhicule est due au mouvement de tangage de notre véhicule.

Par ailleurs, nous sommes nous-mêmes dépassés par une voiture qui apparaît à la fin de la séquence, quoiqu'elle n'entre pas entièrement dans le secteur de stéréovision (l'image droite ne la montre qu'en partie).

On remarquera le mouvement de tangage de notre véhicule, visible au début de la séquence comme une variation de la largeur de l'image de la chaussée.

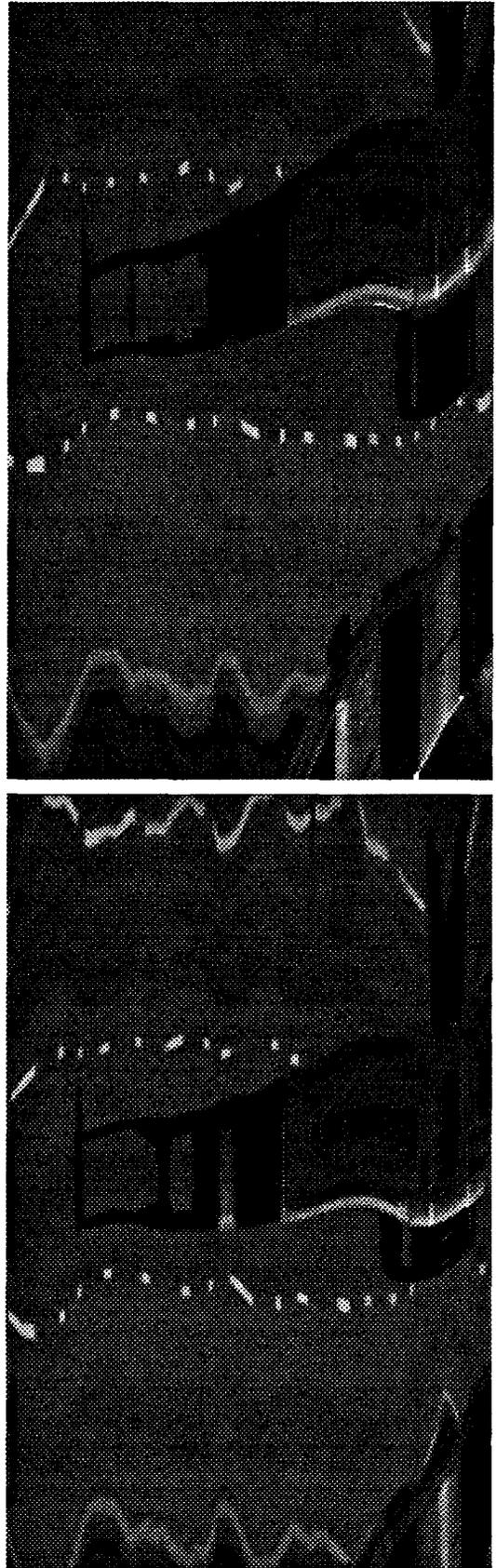


Figure 5.20: Nous nous approchons de la R4 qui nous précède.

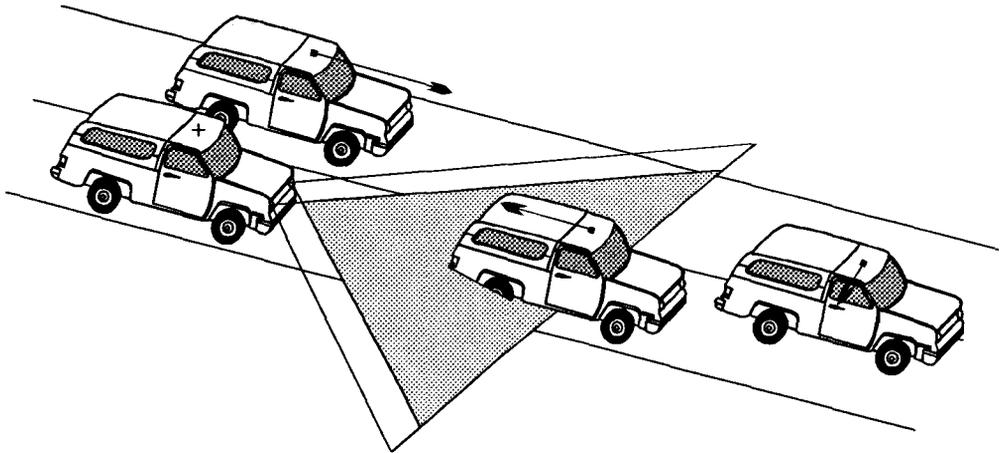


Figure 5.21: Les déplacements des différents véhicules au cours de la séquence. Les flèches indiquent le mouvement relatif à notre véhicule (marqué d'une croix).

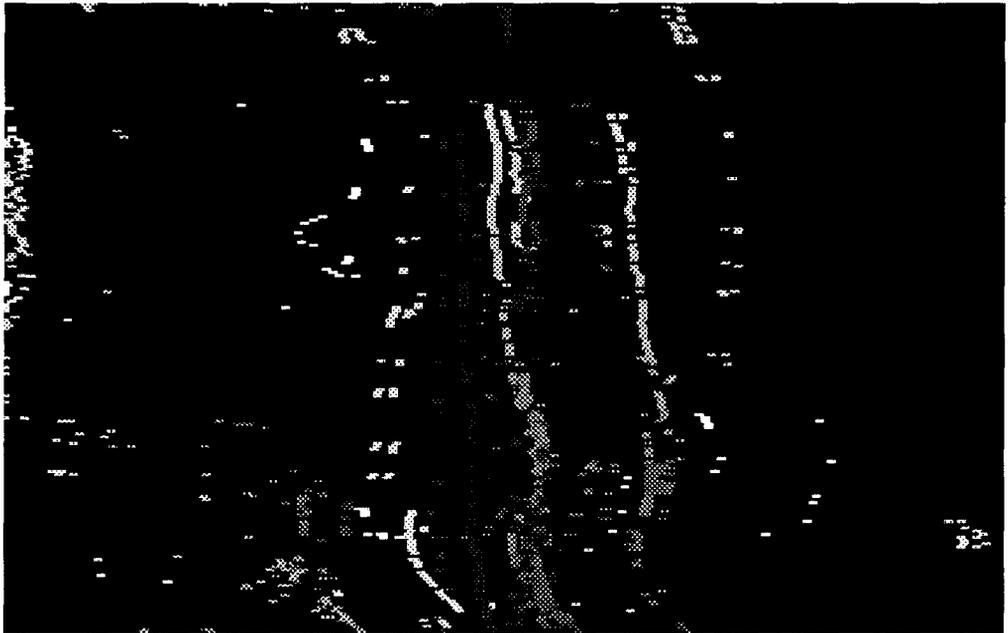


Figure 5.22: Scène reconstituée par les traitements.

Sur la scène reconstituée (cf. Figure 5.22), on reconnaît bien la forme de la R4. On notera que, contrairement à ce que l'on constatait sur les images du stéréoscope, sa largeur apparente est ici constante dans le temps. Ceci est dû à la définition de l'axe horizontal, qui représente la largeur réelle des objets, et non pas celle de leur projection sur les capteurs. La teinte des contours de la R4, de plus en plus foncée, traduit son rapprochement au cours du temps.

On voit également, en pointillé, la ligne blanche qui sépare les deux voies de l'autoroute. On notera son ondulation, qui rappelle que le repère de travail est lié à notre véhicule, et non à l'axe de la chaussée: notre propre mouvement de lacet par rapport à cette dernière est à l'origine de cette ondulation. Le terre-plein central est également visible, à l'extrémité gauche de l'image.

Le véhicule qui se rabat devant la R4 est clairement visible en bas de l'image, sous la forme d'une ligne claire continue correspondant à l'extrémité gauche du véhicule. Son extrémité droite n'est pas visible en raison d'une occlusion par la R4 (elle est masquée à la caméra droite). L'extrémité de sa longue ombre est visible en bas, à l'extrême droite de l'image.

La voiture qui nous double est visible, en bas et à gauche de l'image, comme une ligne oblique (très sombre en raison de sa proximité). Son faible contraste par rapport à la chaussée n'entraîne que très peu de détections manquées au cours du temps, ce qui indique une bonne robustesse de la détection de contours aussi bien que de la mise en correspondance. Les quelques erreurs de mise en correspondance paraissent dues à l'apparition momentanée, sur l'image droite, de la ligne blanche qui délimite le terre-plein central: l'algorithme semble confondre cette ligne avec le bord droit de la voiture.

### 5.4.7. Dépassement par un véhicule

Cette séquence peut être considérée comme l'inverse de la situation précédente: cette fois c'est nous qui sommes dépassés par une voiture.

L'action se décompose en trois phases:

- Le dépassement est en cours. La voiture qui nous double est à notre gauche ou un peu en arrière: elle n'est donc pas encore visible pour le stéréoscope linéaire, qui vise vers l'avant.
- La voiture termine son dépassement. Elle entre alors progressivement dans le secteur de stéréovision. On peut d'abord localiser son extrémité droite, ce qui suffit à déduire son mouvement relatif. Ensuite apparaît son extrémité gauche. La voiture est alors entièrement visible, et sa largeur peut être calculée.
- Le dépassement est terminé. La voiture est entièrement visible pour les deux caméras, et son éloignement est rendu visible par la diminution de la largeur de son image et par celle de la disparité. Quand la voiture atteint une distance supérieure à la portée du stéréoscope, elle disparaît de l'image.

Un véhicule que nous suivons apparaît brièvement sur la voie de droite. Ici encore, cette brève apparition d'un véhicule situé à la limite de portée du stéréoscope linéaire est due au mouvement de tangage de notre voiture.

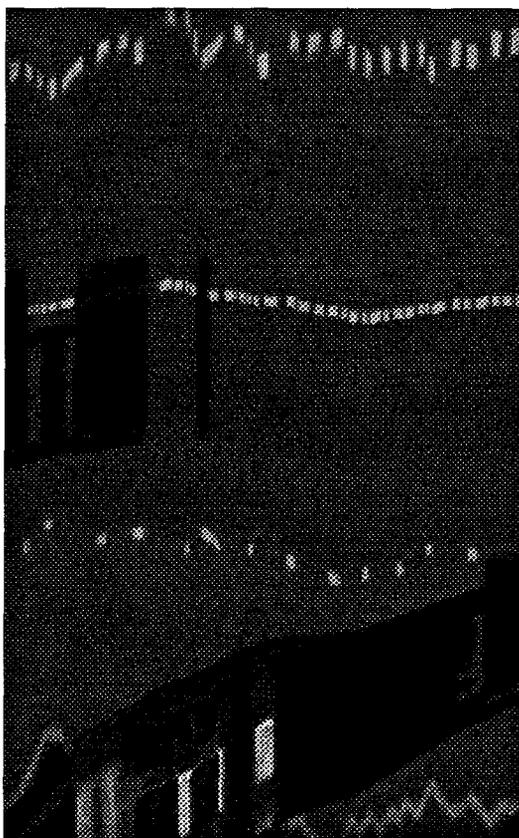
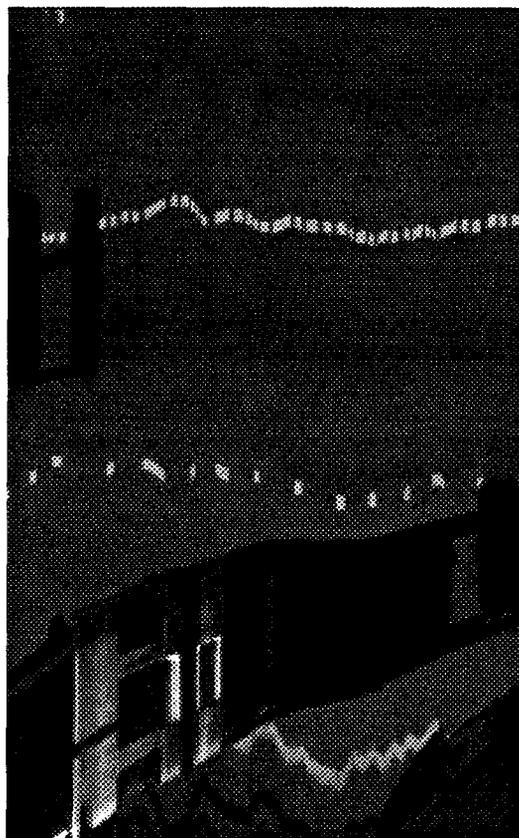


Figure 5.23: La voiture de gauche est en train de nous dépasser.

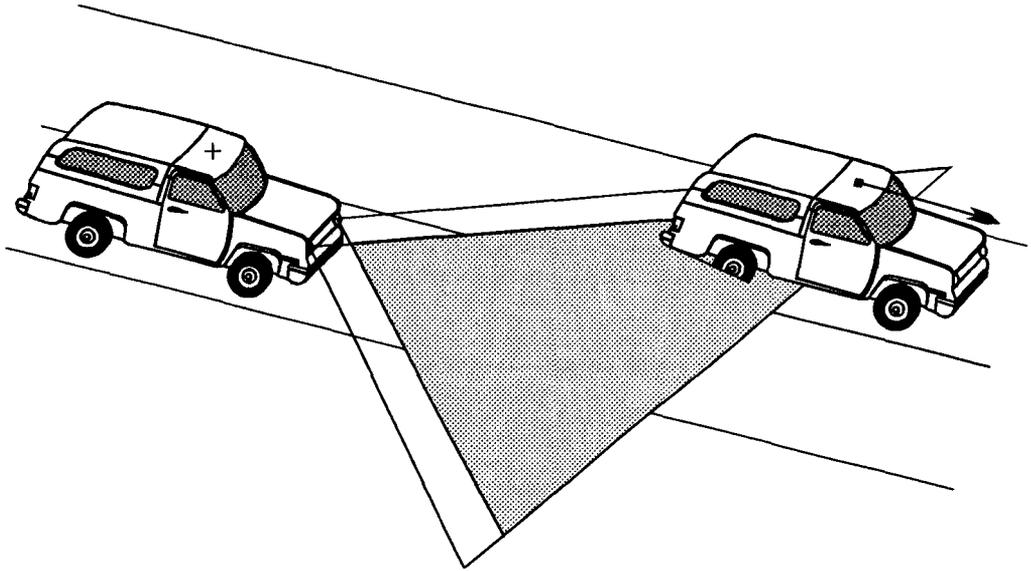


Figure 5.24: Les déplacements des différents véhicules au cours de la séquence.



Figure 5.25: Scène reconstituée par les traitements.

Dans la scène reconstituée (cf. Figure 5.25), on voit clairement la bande blanche continue qui sépare la voie droite de la chaussée d'une voie de dégagement, et la bande discontinue qui sépare les voies gauche et droite.

La voiture qui nous double est d'abord visible par son bord droit, qui est dans le secteur de stéréovision (c'est à dire présent sur les images gauche et droite) dès le début de la séquence. Ce bord est très bien détecté et localisé, de même que la plaque d'immatriculation qui apparaît brièvement à l'image. Le bord gauche de la voiture n'apparaît que plus tard dans le secteur de stéréovision: son apparition donne lieu à quelques fausses détections — probablement par suite d'une confusion avec la bande continue qui délimite le terre-plein central — après quoi il est correctement détecté et localisé jusqu'à la fin de la séquence.

Le début de la séquence montre le véhicule qui nous précède sur la voie de droite. On voit aussi, sur la gauche, une fausse détection, sorte de pic pointé vers la gauche qui semble causé par une confusion entre la ligne blanche qui délimite le centre de la chaussée et le bord droit de la voiture qui nous dépasse.

Enfin, on note un fort décalibrage vertical, causé par la vibration des objectifs — visible sur la séquence originale, notamment sur le véhicule qui nous précède — mais qui ne semble pas avoir de conséquences notables sur la détection ou sur la localisation. Ce point est particulièrement intéressant d'un point de vue pratique: dans des cas courants tels que celui-ci, même un stéréoscope légèrement déréglé peut fonctionner correctement.

### 5.4.8. Croisement d'un vélomoteur sur une route de campagne

Nous roulons sur une route de campagne entourée d'herbes hautes. Noter la forme hyperbolique vers le bord du cadre, caractéristique de l'image d'un objet vertical en rapprochement [Colle 90].

Un vélomoteur nous croise sur la voie de gauche, à une vitesse relative assez élevée (estimée à environ 120 km/h) puisque les deux véhicules roulent en sens inverse.

L'image du vélomoteur est visible comme une grande "corne" dont une extrémité est sur la route, tandis que l'autre extrémité se perd dans les hautes herbes. Cette image est intéressante pour tester la robustesse des traitements, car son exploitation impose de distinguer un objet devant un fond complexe, fortement texturé, qui génère un nombre important de primitives, et de le mettre en correspondance correctement dans ces conditions difficiles.

On notera par ailleurs un poteau, faisant partie d'une clôture, visible comme un trait blanc dans le coin inférieur gauche de l'image.

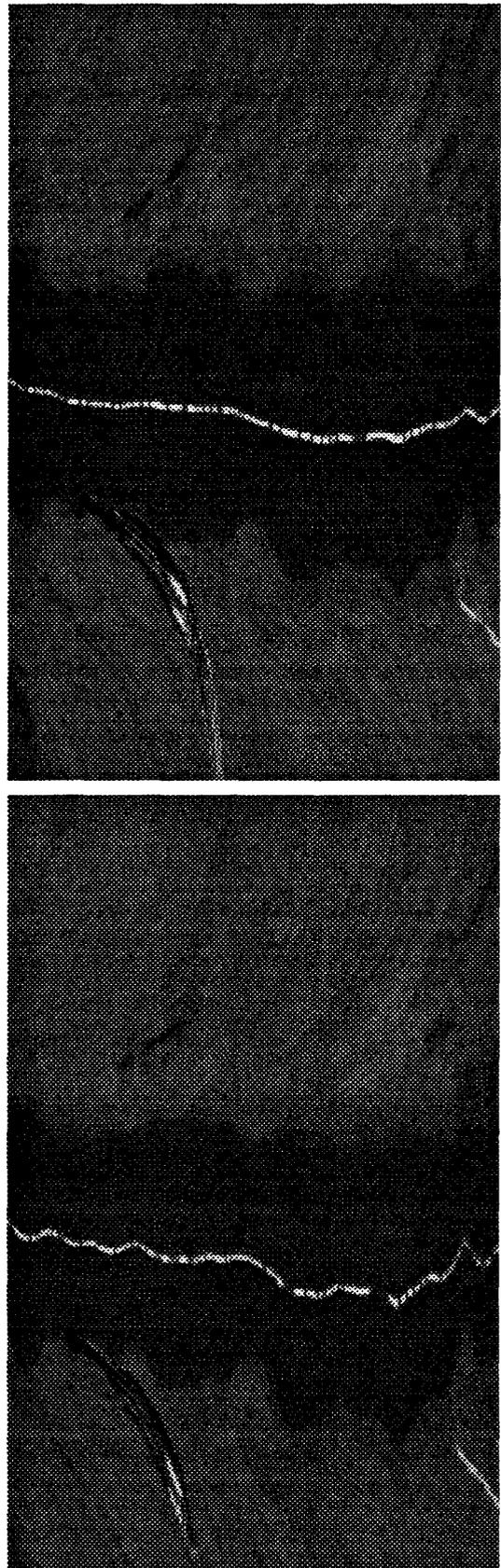


Figure 5.26: Croisement d'un vélomoteur - séquence originale.

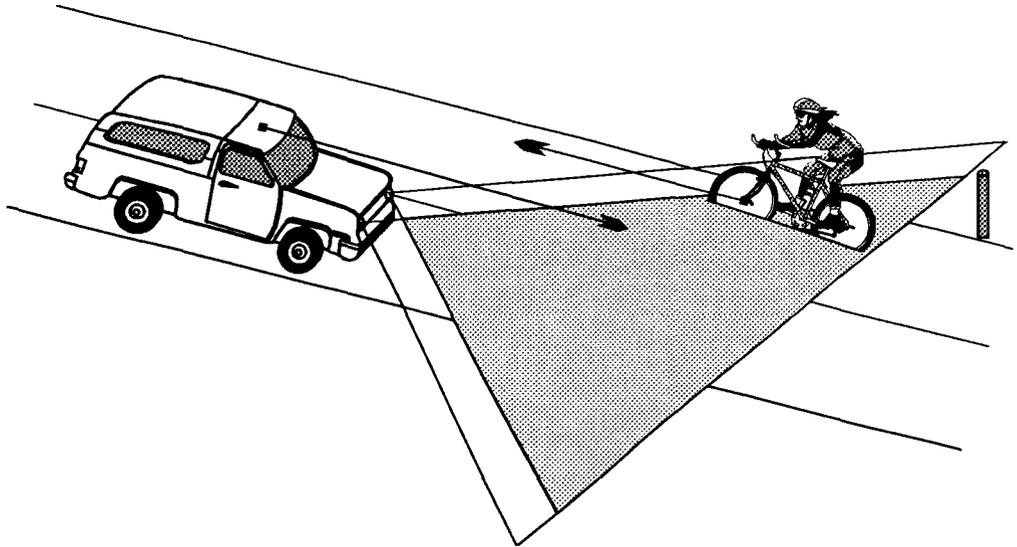


Figure 5.27: Ce qui se passe dans la séquence de la Figure 5.26. Les mouvements sont représentés par rapport au sol.

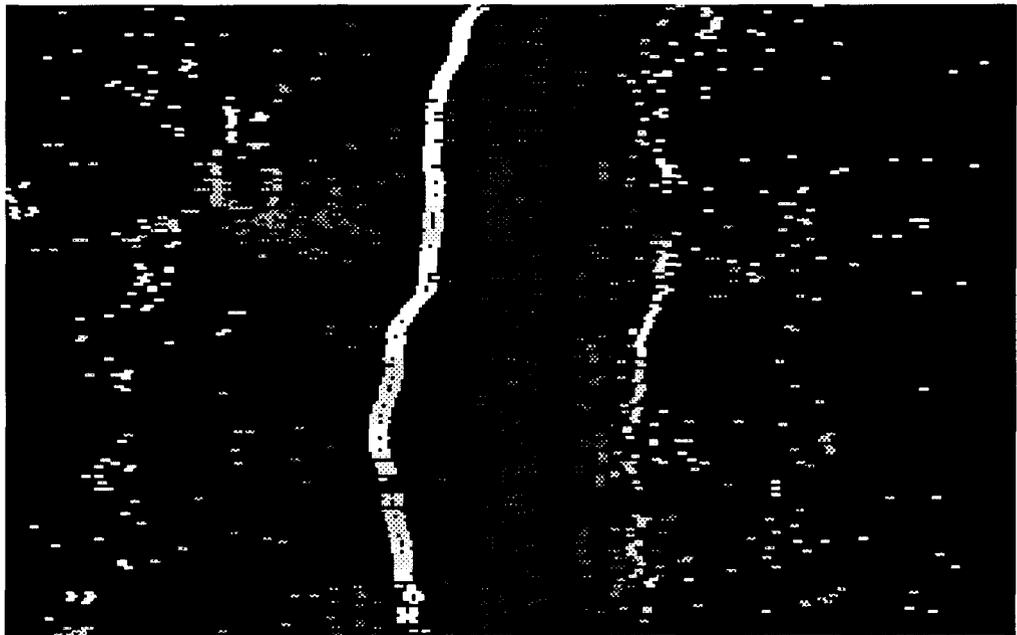


Figure 5.28: Reconstitution par les traitements de la scène de la Figure 5.26.

La bande blanche centrale, fortement contrastée et nettement marquée sur le fond uniforme de la chaussée, est très bien détectée et mise en correspondance (cf. Figure 5.28).

Le bord droit de la route, qui se distingue moins bien en raison d'un faible contraste avec le bas-côté de couleur vert sombre, est également détecté, bien que plus difficilement. Curieusement, le bord gauche de la chaussée n'est que rarement détecté par la procédure de détection de contours: il semble que l'on atteigne la limite de contraste en dessous de laquelle la détection ne se fait plus.

Le vélomoteur est visible en haut à gauche de l'image. Il est correctement localisé et détecté sur le fond uniforme de la chaussée, puis sur celui, beaucoup plus texturé, du bas-côté. Sa sortie du secteur de stéréovision, lorsqu'il n'est plus vu que par la caméra gauche, entraîne des appariements erronés avec les hautes herbes. Le poteau, visible en bas à gauche, est également bien détecté et mis en correspondance, mais avec une lacune à la fin de la séquence.

Les hautes herbes ne donnent pas lieu à beaucoup de faux appariements, comme on pouvait le craindre. En réalité, elles n'engendrent que peu de primitives, leur contraste étant relativement faible. Cette séquence ne constitue donc pas un test aussi sévère que nous l'espérions pour la robustesse des algorithmes.

---

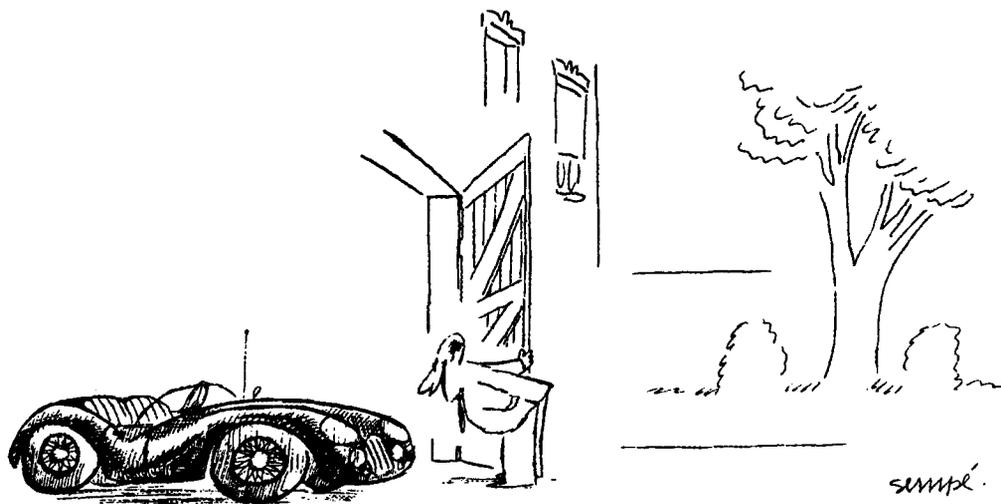
## 5.5. Conclusion

Les exemples présentés dans ce Chapitre permettent de tirer quelques conclusions quant aux qualités et aux lacunes de la version actuelle de nos algorithmes, et aussi de notre dispositif de prise de vue.

Tout d'abord, la détection et la localisation des obstacles sont bonnes même avec un certain décalibrage du stéréoscope linéaire. Ce n'est que dans des cas extrêmes (cf. § 5.4.5) que des problèmes peuvent apparaître. Cette grande robustesse permettra d'obtenir de bons résultats sans devoir recalibrer souvent le stéréoscope linéaire, ce qui est essentiel pour l'application envisagée.

La procédure d'extraction des points de contour décrite au Chapitre 4 est assez robuste, et donne de bons résultats dans la quasi-totalité des séquences enregistrées lors des essais - des séquences non représentés dans ce Chapitre, réalisés avec des expositions volontairement différentes pour les deux caméras, ont été correctement mises en correspondance. Cependant, dans quelques cas extrêmes, comportant des fonds fortement texturés (notamment dans la séquence du § 5.4.8), le choix du seuil de détection n'est pas indifférent. Ce point a été revu par Jean-Christophe Burie dans le but d'augmenter encore la robustesse du dispositif de détection d'obstacles.

La procédure de mise en correspondance des points de contours mis en évidence dans les images, bien qu'encore incomplète, donne également satisfaction. En particulier, nous n'avons pas trouvé d'appariements manqués du fait de cette procédure, ce qui indique une bonne pertinence du critère de ressemblance retenu. Les appariements multiples, qui apparaissent fréquemment dans les scènes reconstituées, seront résolus par la procédure d'analyse de la cohérence spatio-temporelle de la scène 3D, développée par Jean-Christophe Burie.



*«Bonne nuit petit monstre!»*

## CONCLUSION

Ayant constaté, par l'analyse des enquêtes REAGIR sur les causes d'accidents de la route, l'importance des facteurs liés à l'inattention du conducteur, nous avons envisagé, comme moyen de prévention de ces accidents, l'utilisation de dispositifs de détection d'obstacles destinés à émettre une alarme pour attirer l'attention du conducteur. Un passage en revue des méthodes de détection d'obstacles existantes nous a amené à conclure à l'inadaptation d'un grand nombre d'entre elles pour cet usage, essentiellement en raison d'un temps de réponse excessif et, pour beaucoup, d'une fiabilité insuffisante.

Nous avons donc choisi d'utiliser une approche basée sur la stéréovision, qui permet d'évaluer la distance des objets détectés. Cette caractéristique permet d'augmenter la robustesse du système de détection, en évitant les fausses alarmes dues notamment aux ombres et à la signalisation horizontale. La stéréovision étant très coûteuse en temps de calcul, nous avons développé un capteur stéréoscopique utilisant des caméras linéaires au lieu des caméras matricielles généralement adoptées pour ce type d'application. Ce choix permet de diminuer très sensiblement la quantité de données à traiter. De plus, la géométrie de ce capteur contribue à diminuer la complexité des algorithmes: l'objectif des traitements en temps réel peut ainsi être atteint avec une architecture informatique relativement simple et peu coûteuse, construite autour d'un ordinateur personnel de type PC-AT.

Nous avons conçu et réalisé ce capteur, que nous avons baptisé *stéréoscope linéaire*. Nous avons également conçu une méthode de calibrage spécifique qui permet de lui donner les caractéristiques géométriques choisies à l'avance en fonction de l'application envisagée et, surtout, d'assurer la fusion des plans de visée des deux caméras linéaires qui le composent. Cette condition, nécessaire pour travailler sur des scènes quelconques, n'avait jamais été exploitée jusqu'à présent, car elle est difficile à réaliser en pratique avec les méthodes de calibrage préexistantes.

Nous avons développé des méthodes spécifiques de traitement des images acquises par le stéréoscope linéaire, permettant la détection d'obstacles sur route dans les cas les plus courants.

Enfin, nous avons développé les méthodes d'enregistrement des images stéréoscopiques linéaires issues de notre capteur, et nous les avons appliquées lors des essais sur route du dispositif de détection d'obstacles basé sur le stéréoscope linéaire. Ces essais ont été présentés dans le Chapitre 5 de ce rapport, et leurs résultats sont jugés satisfaisants, sauf dans de rares cas très difficiles.

Le système a été intégré au véhicule expérimental Prolab 2 et présenté aux pouvoirs publics et à la presse à l'occasion des journées BMM'94, du 17 au 21 octobre 1994. Il a donné entière satisfaction durant plusieurs jours d'essais réalisés sur le circuit d'essais routiers de Mortefontaine (Oise).

Le stéréoscope linéaire n'est évidemment pas limité à la seule détection d'obstacles sur route. Tout d'abord, il peut également être utilisé à la détection d'obstacle devant des véhicules de transports guidés, tels que les trains ou les métros. Il peut également être utilisé en robotique mobile, pour le guidage et/ou la détection d'obstacles. Les travaux déjà publiés sur ce sujet (notamment [Inigo 85], [Inigo 87] et [Colle 90]) pourraient trouver des évolutions intéressantes en tirant parti du capteur stéréoscopique linéaire que nous avons développé, plus

précis que ceux utilisés à l'origine grâce notamment à un meilleur calibrage. Ceci lui permet d'élargir sa plage d'utilisation à toutes sortes de scènes naturelles, alors que [Colle 90], par exemple, se cantonnait à des scènes ne contenant que des arêtes verticales.

Les caractéristiques de rapidité, de précision et de souplesse qu'offrent les caméras linéaires quant aux paramètres de prise de vue (cadence et temps d'exposition notamment) permettent également d'envisager de nombreuses autres applications dans des domaines variés:

- La métrologie utilise déjà les caméras linéaires comme capteurs sans contact de haute précision. Le stéréoscope linéaire fournit en plus la localisation 2D des objets dans le plan de visée. La méthode de calibrage et d'étalonnage mise en oeuvre dans le présent travail n'est pas suffisamment précise pour ce type d'usage, mais on trouvera dans l'Annexe B une liste d'améliorations possibles dont la mise en oeuvre permettrait d'atteindre la précision maximale, limitée seulement par la résolution des capteurs.
- L'analyse des mouvements rapides utilise généralement des caméras vidéo à cadence élevées, très coûteuses et qui posent des problèmes d'acquisition, la capacité des mémoires des cartes d'acquisition étant limitée, et très rapidement saturée par les séquences d'images. Les seuls moyens utilisables en pratique pour l'acquisition à haute cadence d'images vidéo sont la réduction de la résolution — qui permet d'acquérir "à la volée" des séquences plus longues, mais diminue la précision spatiale — et l'utilisation d'enregistreurs spéciaux, sur bande ou cassette analogique, qui sont extrêmement coûteux.

Les caméras linéaires sont peu coûteuses (leur prix actuel équivaut à celui d'une bonne caméra de surveillance), et permettent des cadences de prise de vue très élevées, sans sacrifier la résolution, sans saturer la mémoire vive disponible. Ainsi, certaines applications industrielles fonctionnent couramment à 1000 images par seconde avec des caméras linéaires de 2048 pixels. L'utilisation du stéréoscope linéaire, avec le dispositif d'enregistrement proposé au Chapitre 5, permet d'enregistrer, avec une très grande précision, des mouvements 2D très rapides dans le plan de visée.

Les problèmes qui se posent encore concernent surtout les méthodes de traitement d'image destinés à la détection d'obstacle. Les algorithmes présentés ici n'utilisent que l'information contenue dans un seul couple d'images acquises simultanément par les deux caméras du stéréoscope. Or, leur contenu est souvent insuffisant pour permettre une mise en correspondance sans ambiguïté: même un opérateur humain, observant sur l'écran de contrôle les images acquises, est incapable de dire ce qu'il a devant les yeux (cf. Figure 1).



Figure 1: Image linéaire unique, étirée verticalement pour améliorer sa lisibilité. Il est difficile d'y reconnaître un piéton marchant entre deux lignes blanches tracées sur une chaussée .

Ce n'est qu'en observant ce qui se passe au cours du temps, lorsque les objets observés et/ou les caméras se déplacent dans la scène, que l'on peut distinguer

avec un minimum d'erreur la présence de différents objets solides, et aboutir à une mise en correspondance dépourvue d'ambiguïté (cf. Figure 2).

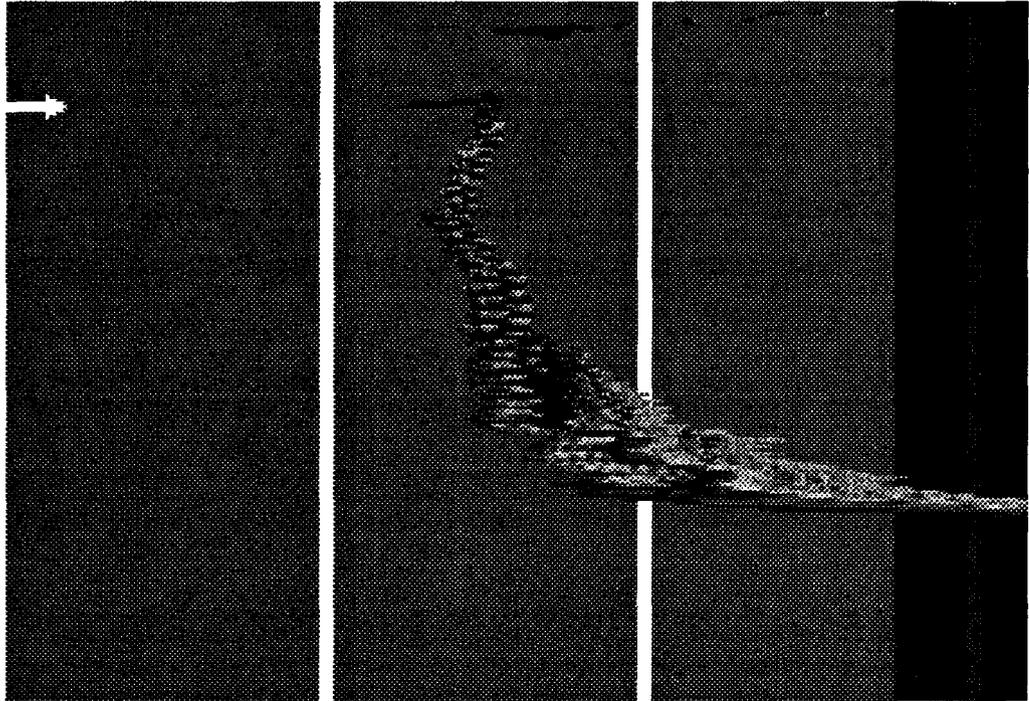


Figure 2: Séquence d'images linéaires dont est extraite la ligne de la Figure 1 (la flèche indique la position de cette dernière). L'information maintenant disponible permet de distinguer clairement le piéton du fond, et même d'estimer son mouvement (vitesse et direction).

Le développement de procédures de suivi temporel, adaptées à la stéréovision linéaire, apparaît donc souhaitable pour améliorer la robustesses des détections obtenues. Ce suivi temporel permettrait également de déterminer les vitesses de déplacement des objets relativement au stéréoscope. Notons que différentes approches dans ce domaine existent déjà en vision matricielle.

Cet aspect est actuellement étudié dans le cadre du programme Prometheus par Jean-Christophe Burie, chercheur du Centre d'Automatique de Lille, qui soutiendra prochainement une thèse faisant suite à celle-ci.

(A suivre)

## **BIBLIOGRAPHIE**

[Adams 88] J.-G. Adams:

***Risk homeostasis and the purpose of safety regulation.***

Ergonomics, vol. 31, n° 4, pp. 407-428, 1988.

[Adiv 85] G. Adiv:

***Determining three-dimensional motion and structure from optical flow generated by several moving objects.***

IEEE trans. Pattern Analysis and Machine Intelligence, vol. PAMI-7, n° 4, pp. 384-401, juillet 1985.

[Alizon 90] J. Alizon, J. Gallice, L. Trassoudaine & S Treuillet:

***Multi-sensory data fusion for obstacle detection and tracking on motorway.***

Proc. Pro-Art Workshop on Vision, Sophia Antipolis, pp. 179-187, 19-20 avril 1990.

[Aubert 89] D. Aubert:

***Mise en correspondance d'indices images en résolutions multiples.***

Thèse de doctorat, Institut National Polytechnique de Grenoble, janvier 1989.

[Auto-Moto 84]

***Sondage IFRES, Auto-Moto sur 27 panneaux usuels.***

in Auto-Moto, n° 24, pp. 8 et suivantes, février 1984.

[Ayache 89] N. Ayache:

***Vision stéréoscopique et perception multisensorielle.***

Ed. InterEditions, Paris, 1989.

[Ballard 82] D.H. Ballard & C.M. Brown:

***Computer vision.***

Ed. Prentice Hall, Englewood Cliffs (New Jersey), 1982.

[Barnard 80] S.T. Barnard & W.B. Thompson:

***Disparity analysis of images.***

IEEE trans. Pattern Analysis and Machine Intelligence, vol. PAMI-2, n° 4, pp. 333-340, juillet 1980.

[Beauviala 78]

***Entretien avec Jean-Pierre Beauviala.***

Cahiers du Cinéma n° 285, pp. 8-15, février 1978; n° 286, pp. 5-15, mars 1978; n° 287, pp. 5-17, avril 1978; n° 288, pp. 16-21, mai 1978.

[Beauviala 81]

***Entretien avec Jean-Pierre Beauviala.***

Cahiers du cinéma n° 325, juin 1981.

[Bhanu 84] B. Bhanu & O.D. Faugeras:

***Shape matching of two-dimensional objects,***

IEEE trans. Pattern Analysis and Machine Intelligence, vol. PAMI-6, n° 2, pp. 137-156, mars 1984.

[Boursier 91] D. Boursier & P.-A. Taufour:

***La technologie multimédia.***

Ed. Hermès, Paris, 1993.

- [Briot 79] M. Briot, J.-C. Talou & G. Bauzil:  
***Le système de perception du robot mobile autonome Hilare.***  
Proc. 2ème congrès AFCET/INRIA de Reconnaissance des Formes et Intelligence Artificielle, Toulouse, 12-14 septembre 1979.
- [Broughton 91] J. Broughton, H. Colk, B. Dryselius, H. Fontaine, G. Gastaldi, C. Hydén, J.-H. Klöckner, L. Neumann, R. Risser, I. Velding:  
***The Pro-Gen traffic safety checklist.***  
Rapport PROMETHEUS, Stuttgart, 1991.
- [Buil 89] C. Buil:  
***Astronomie CCD - Construction et utilisation des caméras CCD en astronomie amateur.***  
Ed. Société d'Astronomie Populaire, Toulouse, 1989.
- [Carlsson 90] S. Carlsson & J.-O. Eklundh:  
***Object detection using model based prediction and motion parallax.***  
Proc. Pro-Art Workshop on Vision, Sophia-Antipolis, pp.139-145, 19-20 avril 1990.
- [Chang 93] Y.-L. Chang, X. Lebègue & J.K. Aggarwal:  
***Calibrating a mobile camera's parameters.***  
Pattern Recognition, vol. 26, n° 1, pp. 75-88, 1993.
- [Chasseur 92] Chasseur d'Images:  
***Dossier-Tests objectifs Nikon.***  
Ed. Jibena, Paris, 1992.
- [Chich 87] Y. Chich:  
***La gestion de la limitation de vitesse: problèmes et perspectives.***  
Recherche Transports Sécurité, n° 15, pp. 5-12, septembre 1987.
- [Colle 90] O. Colle:  
***Vision stéréoscopique à l'aide de deux caméras linéaires - Application à la robotique mobile.***  
Thèse de doctorat, Institut National des Sciences Appliquées de Lyon, novembre 1990.
- [Couzy 81] A. Couzy:  
***La télédétection.***  
Coll. Que sais-je? Presses Universitaires de France, Paris, 1981.
- [Crowley 86] J.-L. Crowley & J. Coutaz:  
***Navigation et modélisation pour un robot mobile.***  
Technique et Science Informatique, vol. 5, éd. AFCET Gauthier-Villars, Paris, 1986.
- [Deloof 90] P. Deloof, N. Haese & S. Baranowski:  
***Recherche bibliographique sur les radars de prévention des collisions.***  
Recherche Transports Sécurité, n° 26, pp. 15-22, juin 1990.
- [Deriche 87] R. Deriche:  
***Optimal edge detection using recursive filtering.***  
Proc. 1st Int. Conf. On Computer Vision, Londres, 8-12 juin 1987.

- [Dobias 87] G. Dobias:  
***La voiture intelligente, un plus pour la sécurité routière?***  
Recherche Transports Sécurité, n° 15, pp. 13-18, septembre 1987.
- [Documentation 89]  
***Livre blanc de la sécurité routière.***  
La Documentation Française, Paris, 1989.
- [Dreschler 82] L. Dreschler & H.-H. Nagel:  
***Volumetric model and 3D trajectory of a moving car derived from monocular TV frame sequences of a street scene.***  
Computer Graphics and Image Processing, vol. 20, pp. 199-228, 1982.
- [du Buf 90] J.M.H. du Buf, M. Kardan & M. Spann:  
***Texture feature performance for image segmentation.***  
Pattern Recognition, vol. 23, n° 3/4, pp. 291-309, 1990.
- [Duvieubourg 93] L. Duvieubourg, T. Carpentier & J.-G. Postaire:  
***Monitoring changes in length with a linear CCD camera in mechanics of deformable solids.***  
Proc. 5th Int. Congress on Condition Monitoring and Diagnostic Engineering Management, Bristol, Juillet 1993.
- [Dyck 89] R.H. Dyck, Y.S. Abedini, J.S. Kim & K.K. Shah:  
***Time-Delay and Integration (TDI) image sensor (128x1024 elements) with small, high performance pixels.***  
in 1989 Fairchild Weston CCD sensors, systems and developmental technology databook, pp. 251-255, 1989.
- [Enkelmann 90] W. Enkelmann:  
***Obstacle detection by evaluation of optical flow fields.***  
Proc. Pro-Art Workshop on Vision, Sophia Antipolis, pp.146-155, 19-20 avril 1990.
- [Faig 75] W. Faig:  
***Calibration of close-range photogrammetry systems: mathematical formulation.***  
Photogrammetric Engineering and Remote Sensing, vol. 41, n°12, pp. 1479-1486, 1975.
- [Ferrandez 86] F. Ferrandez, D. Fleury & G. Malaterre:  
***L'étude détaillée d'accidents (EDA) — Une nouvelle orientation de la recherche en sécurité routière.***  
Recherche Transports Sécurité, n° 9-10, pp. 17-20, juin 1993.
- [Fremy 93] D. & M. Fremy:  
***Quid 1994.***  
Ed. Robert Laffont, Paris, 1993.
- [Gaillard 89] P. Gaillard & S. Roques:  
***Images de synthèse en relief.***  
Micro Systèmes, n°101, pp. 149-154, octobre 1989.
- [Gallagher 81] N. C. Gallagher & G. L. Wise:  
***A theoretical analysis of the properties of median filters.***  
IEEE trans. Acoustics, Speech, Signal proc., vol. ASSP-29, n° 6, pp. 1136-1141, juin 1981.

- [Gentile 94] F. Gentile:  
***La sécurité routière.***  
Coll. Que sais-je? Presses Universitaires de France, Paris, 1994.
- [Got 89] C. Got:  
***L'alcool et la route.***  
Haut Comité d'Etude et d'Information sur l'Alcoolisme, Paris, 1989.
- [Graefe 90] V. Graefe, U. Solder & U. Regensburger:  
***Recognition of objects on highways.***  
Proc. Pro-Art Workshop on Vision, Sophia Antipolis, pp. 133-138, 19-20 avril 1990.
- [Hecht 75] E. Hecht:  
***Theory and problems of optics.***  
Schaum's outline series, éd. McGraw-Hill, New York, 1975.
- [Holzbaur 93] H. Holzbaur & J. Hurd:  
***Lab report: 32 high-speed hard drives.***  
*in* Byte, vol. 18, n° 10, pp. 176-193, octobre 1993.
- [Horaud 88] R. Horaud & T. Skordas:  
***Structural matching for stereo vision.***  
Proc. 9th int. Conf. on Pattern Recognition, Rome, 14-17 novembre 1988.
- [Horaud 89] R. Horaud & T. Skordas:  
***Stereo matching through feature grouping and maximal cliques.***  
IEEE trans Pattern Analysis and Machine Intelligence, vol. PAMI-11, n° 11, pp. 1168-1180, novembre 1989.
- [Horaud 93-1] R. Horaud, R. Mohr & B. Lorecki:  
***On single-scanline camera calibration.***  
IEEE Trans. Robotics and Automation, vol. 9, n° 1, pp. 71-75, 1993.
- [Horaud 93-2] R. Horaud & O. Monga:  
***Vision par ordinateur - outils fondamentaux.***  
Ed. Hermès, Paris, 1993.
- [Inigo 85] R.M.Inigo, T. Tkacik & E.S. McVey:  
***The application of linear image arrays to mobile robot guidance and navigation.***  
Proc. 3rd Int. Conf. on Automated Guided Vehicle Systems, Stockholm, pp. 157-168, octobre 1985.
- [Inigo 87] R.M.Inigo & T. Tkacik:  
***Mobile robot operation in real-time with linear image array based vision.***  
Proc. IEEE int. Symp. Intelligent Control. Philadelphie, pp. 228-233, janvier 1987.
- [Ivanov 93] O.P. Ivanov, A.V. Chesnokov, V.E. Stepanov, A.N. Sudarkin, A.S. Tyurin & A.V. Volkovich:  
***Gamma-vision camera real-time system for creation images of gamma-radioactive objects.***  
Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, vol. 3, pp. 1-6, Le Touquet, 1993.

- [Jerian 90] C. Jerian & R. Jain:  
*Polynomial methods for structure from motion.*  
IEEE trans. Pattern Analysis and Machine Intelligence, vol. 12, n° 12, pp. 1150-1166, décembre 1990.
- [Keast 93] C. L. Keast & C. G. Sodini:  
*A CCD/CMOS-based imager with integrated focal plane signal processing.*  
IEEE Journal of Solid-state circuits, vol. 28, n°4, pp. 431-437, 1993.
- [Khoumi 86] A. Khoumi & P. Migaud:  
*Amélioration des capacités comportementales d'Hilare, pilotage et contrôle d'exécution de mouvements.*  
Rapport LAAS n° 86002, Université de Toulouse, mars 1986.
- [Kite 90] D.H. Kite & M. Magee:  
*Determining the 3D position and orientation of a robot camera using 2D monocular vision.*  
Pattern Recognition, vol. 23, n° 8, pp. 819-831, 1990.
- [Lee 90] H.-J. Lee & W.-L. Lei:  
*Region matching and depth finding for 3D objects in stereo aerial photographs.*  
Pattern Recognition, vol. 23, n° 1/2, pp. 81-94, 1994.
- [Lee 91] H. Lee, M. A. Weissman & J. P. Powell:  
*Measuring highway inventory features using stereoscopic imaging system.*  
Proc. 2nd int. conf. Applications of Advanced Technologies in Transportation Engineering, Minneapolis, pp. 448-452, août 1991.
- [Lenz 88] R.K. Lenz & R.Y. Tsai:  
*Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology.*  
IEEE Trans. Pattern Analysis and Machine Intelligence, vol.10, n° 5, pp.713-720, 1988.
- [Lumia 83] R. Lumia, R.M. Haralick, O. Zuniga, L. Shapiro, T.-C. Pong & F.-P. Wang:  
*Texture analysis of aerial photographs.*  
Pattern Recognition, vol. 16, n° 1, pp. 39-46, 1983.
- [Malaterre 86] G. Malaterre & J.-F. Peytavin:  
*L'analyse de la tâche appliquée à l'étude détaillée d'accident.*  
Recherche Transports Sécurité, n° 9-10, pp. 21-26, juin 1986.
- [Malaterre 92] G. Malaterre & H. Fontaine:  
*Les aides à la conduite: quels enjeux pour la sécurité?*  
Recherche Transports Sécurité, n° 35, pp. 43-54, septembre 1992.
- [Martins 81] H.A. Martins, J.R. Birk & R.B. Kelley:  
*Camera models based on data from two calibration planes.*  
Computer Graphics and Image Processing, vol. 17, pp. 173-180, 1981.
- [Mautner 94] A. Mautner & R. Lange:  
*Trafic monstre.*  
Documentaire, vidéo, réal. A. Mautner, Prod. Sender Freies Berlin & Arte, 1ère diff. Arte, 18 septembre 1994.

- [Medioni 85] G. Medioni & R. Nevatia:  
***Segment-based stereo matching.***  
Computer Vision, Graphics and Image Processing, vol. 31, pp. 2-18, 1985.
- [Meygret 90] A. Meygret & M. Thonnat:  
***Objects detection in road scenes using stereo data.***  
Proc. Pro-Art Workshop on vision, Sophia Antipolis, pp. 119-130, 19-20 avril 1990.
- [Mitchell 77] O.R. Mitchell & S.G. Carlton:  
***Image segmentation using a local extrema texture measure.***  
Proc. IEEE Comp. Soc. Conf. on Pattern Recognition and Image Processing, p. 386. *Aussi dans* Pattern Recognition, vol. 10, pp. 205-210, 1978.
- [Moore 89] M. Moore,  
***Roger & Me.***  
Documentaire, film 35 mm, réal. M. Moore, prod. Dog Eat Dog Films,  
© Warner Bros, 1989.
- [Murphy 87] T. Murphy & D. Debs:  
***Anti-blooming and integration control.***  
*in* 1987 Fairchild Weston CCD databook, pp. 232-234, 1987.
- [Nomura 92] Y. Nomura, M. Sagara, H. Naruse & A. Ide:  
***Simple calibration algorithm for high-distorsion-lens camera.***  
IEEE trans. Pattern Analysis and Machine Intelligence, vol. 14, n° 11, pp. 1095-1099, 1992.
- [Ohta 85] Y. Ohta & T. Kanade:  
***Stereo by intra- and inter-scanline search using dynamic programming.***  
IEEE trans. Pattern Analysis and Machine Intelligence, vol. PAMI-7, n° 2, pp. 139-154, mars 1985.
- [Okamoto 84] A. Okamoto:  
***The model construction problem using the collinearity condition.***  
Photogrammetric Engineering and Remote Sensing, vol. 50, n°6, pp. 705-711, 1984.
- [Okutomi 93] M. Okutomi & T. Kanade:  
***A multiple-baseline stereo.***  
IEEE trans. Pattern Analysis and Machine Intelligence, vol. 15, n°4, pp. 353-363, 1993.
- [Parey 87] C. Parey & A. Lauer:  
***La voiture intelligente.***  
Supplément au n° 190 de La Recherche, pp. 46-55, juillet-août 1987.
- [Petica 93] S. Petica:  
***Risque d'usage du radiotéléphone en voiture et ses effets sur la sécurité routière.***  
Recherche Transports Sécurité, n° 37, pp. 45-56, mars 1993
- [Philip 93] P. Philip - cité dans:  
***La moitié des automobilistes manquent de sommeil avant de prendre la route.***  
La Voix du Nord, p. 24, 21 octobre 1994.

- [Plastock 87] R.A. Plastock & G. Kalley:  
*Infographie - Cours et problèmes.*  
Ed. McGraw-Hill, New York, Série Schaum, 1987.
- [Pollard 85] S.B. Pollard:  
*Identifying correspondences in binocular stereo.*  
PhD Thesis, Université de Sheffield, U. K., novembre 1985.
- [Ramet 86] M. Ramet & G. Vallet:  
*L'analyse des accidents dans la perspective de la sécurité secondaire.*  
Recherche Transports Sécurité, n° 9-10, pp. 43-46, juin 1986.
- [Robe 91] C. Robe:  
*Segmentation d'images par analyse de connexité.*  
Thèse de doctorat, Université de Lille 1, octobre 1991.
- [Rostow 60] W. W. Rostow:  
*Les étapes de la croissance économique.*  
Ed. Cambridge University Press, 1960 pour la version originale. Ed. du Seuil, Paris, 1963 pour la version française.
- [Rumsey 91] F. Rumsey & T. McCormick:  
*Son et enregistrement - théorie et pratique.*  
Ed. Eyrolles, Paris, 1994.
- [Saguez 94] G. Saguez,  
*Trafic monstre.*  
Soirée thématique, réal. G. Saguez, prod. La Sept/Arte & Sender Freis Berlin, 1ère diff. Arte, 18 septembre 1994.
- [Sauvy 84] J. Sauvy:  
*L'industrie automobile.*  
Coll. Que sais-je? Presses Universitaires de France, Paris, 1984.
- [Skordas 88] T. Skordas:  
*Mise en correspondance et reconstruction stéréo utilisant une description structurelle des images.*  
Thèse de doctorat, Institut National Polytechnique de Grenoble, octobre 1988.
- [Sobel 74] I. Sobel:  
*On calibrating computer controlled cameras for perceiving 3D scenes.*  
Artificial Intelligence, vol. 5, pp. 185-198, 1974.
- [Tan 93] C. L. Tan & S.K.K. Loh:  
*Efficient edge detection using hierarchical structures.*  
Pattern Recognition, vol. 26, n° 1, pp. 127-135, 1993.
- [Texereau 52] J. Texereau:  
*Construction du télescope d'amateur,*  
Ed. Société Astronomique de France, Paris, 1° édition 1952, 2° édition 1961.
- [Tsai 85] R.Y. Tsai:  
*A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses.*  
IBM Research report RC51342, 8 mai 1985. Aussi dans 1986 IEEE Int. Conf. on Computer Vision and Pattern Recognition. Aussi dans IEEE Journal of Robotics and Automation, vol. RA-3, n° 4, pp. 323-344, 1987.

- [Tukey 74] J. W. Tukey:  
***Nonlinear (nonsuperposable) methods for smoothing data.***  
Conf. Rec. 1974, EASCON, p. 673, 1974.
- [UER 82]: Union Européenne de Radiodiffusion:  
***Codes temporels de commande de l'UER pour les magnétoscopes (systèmes de télévision à 625 lignes).***  
Norme technique 3097-F, 3<sup>e</sup> édition, éd. Union Européenne de Radiodiffusion, 1982.
- [Van Eslande 92] P. Van Eslande & C. Nachtergaele:  
***Fonctionnement humain et accidents de la route: une contribution à l'étude des limitations potzntielles à la prise en compte des aides informatives.***  
Recherche Transports Sécurité, n° 35, pp. 55-64, septembre 1992.
- [Villain 91] D. Villain:  
***Le montage au cinéma.***  
Ed. Cahiers du cinéma, Paris, 1991.
- [Wane 93] S. Wane:  
***Contribution au pilotage des robots mobiles.***  
Thèse de doctorat, Université de Technologie de Compiègne, 17 mars 1993.
- [Wang 91] L.-L. Wang & W.-H. Tsai:  
***Camera calibration by vanishing lines for 3-D computer vision.***  
IEEE trans. Pattern Analysis and Machine Intelligence, vol. 13, n° 4, pp. 370-376, 1991.
- [Welch 84] T.A. Welch:  
***A technique for high-performance data compression.***  
Computers, vol. 17, n° 6, pp. 8-19, june 1984.
- [Wilde 90] G. Wilde:  
***Véhicule informatisé et sécurité routière.***  
Recherche Transports Sécurité, n° 26, pp. 29-36, juin 1990.
- [Yakimovsky 78] Y. Yakimovsky & R. Cunningham:  
***A system for extracting three-dimensional measurements from a stereo pair of TV cameras.***  
Computer Graphics and Image Processing, vol. 7, n° 2, pp. 195-210, 1978.
- [Ziv 77] J. Ziv & A. Lempel:  
***A universal algorithm for sequential data compression.***  
IEEE trans. Information Theory, vol. IT-23, n° 3, pp. 337-342, 1977.
- [Zhou 92] J.Y. Zhou:  
***Contribution aux méthodes d'étalonnage des capteurs d'images.***  
Thèse de doctorat, Ecole Centrale de Paris, octobre 1992.

## **Annexe A**

# **HYPERFOCALE ET PROFONDEUR DE CHAMP**

Le faisceau lumineux émergent d'un objectif photographique *réel* (et donc, non assimilé à un sténopé) est conique, et converge (dans le cas idéal) en un point unique appelé *point focal* (cf. Figure A.1).

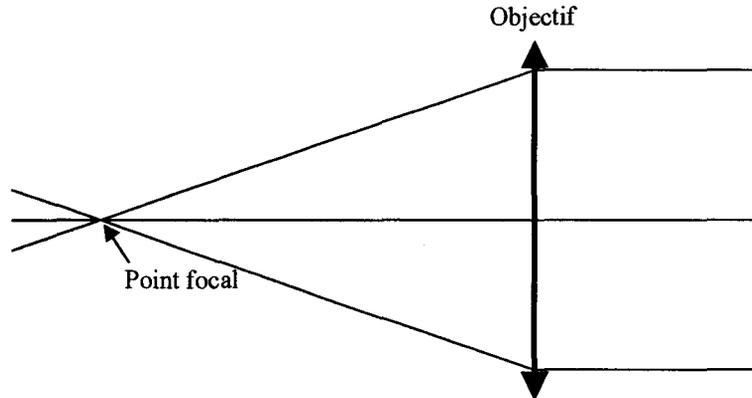


Figure A.1: Faisceau émergent d'un objectif (l'objet est ici supposé à l'infini).

Le capteur — supposé plan et normal à l'axe optique — doit se trouver au niveau du point focal pour donner une image ponctuelle, et donc parfaitement nette. Dans le cas contraire, (cf. Figure A.2), le faisceau émergent forme sur le capteur une image non ponctuelle — une *tache image* — qui se présente comme un disque dont le diamètre  $E$  dépend de la distance  $\delta$  du capteur au point focal, de la distance focale  $f$  et du diamètre utile  $D$  de l'objectif, selon la formule:

$$E = \frac{D \cdot \delta}{f} = \frac{\delta}{n} \quad \text{Equation 1}$$

où  $n = \frac{f}{D}$  est l'ouverture utile de l'objectif, c'est à dire le chiffre indiqué sur la bague de diaphragme (par exemple,  $n$  égale 8 si on travaille avec un objectif réglé sur l'ouverture 8).

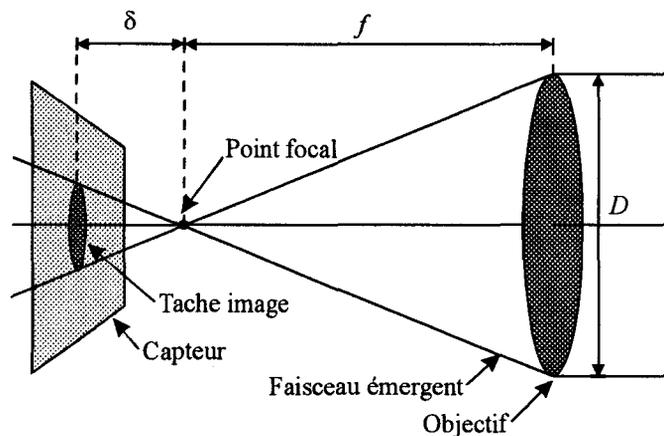


Figure A.2: Diamètre de la tache image sur le capteur.

Dans la pratique, il est impossible d'amener le capteur sur le point focal (on dit *mettre au point*, ou *faire le point*) pour l'ensemble des points de la scène si ceux-ci se trouvent à des profondeurs différentes. En effet, la position  $d'$  du point focal est alors liée à la distance  $d$  du point objet par la relation suivante, valable si l'on assimile l'objectif à une lentille mince [Hecht 75]:

$$\frac{1}{d} + \frac{1}{d'} = \frac{1}{f} \quad \text{Equation 2}$$

L'objectif doit alors être focalisé (c'est à dire déplacé le long de l'axe optique en direction de la scène) pour ramener le point focal sur le capteur (cf. Figure A.3).

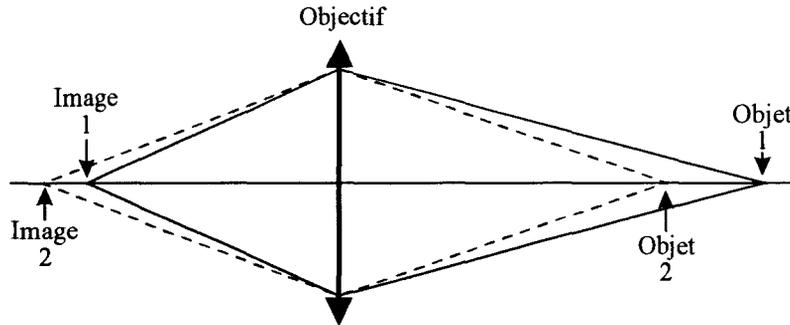


Figure A.3: Influence de la distance  $d$  du point-objet sur la position  $d'$  de son image.

L'Equation 1 devient alors:

$$E = \frac{\delta}{n'} \quad \text{Equation 3}$$

où  $n' = \frac{d'}{D}$  est l'ouverture utile pour la distance de mise au point considérée.  $d'$  est appelée *tirage optique*.

Dans la pratique, on peut remarquer qu'un capteur parfait, de résolution infinie, n'existe pas – les capteurs de nos caméras ont des pixels de  $11\mu\text{m}$  de côté – et qu'il est donc inutile d'obtenir une tache-image parfaitement ponctuelle, celle-ci ne se distinguant pas en pratique d'une tache-image défocalisée dont le diamètre est inférieure aux dimensions du photosite (cf. Figure A.4).

Par conséquent, l'image reste nette tant que le capteur est situé entre les distances  $d'_2$  et  $d'_1$  du centre optique de l'objectif (cf. Figure A.5).  $d'_1$  et  $d'_2$  correspondent aux tirages optiques pour des objets situés respectivement aux distances  $d_1$  et  $d_2$  ( $d_1 < d_2$ ,  $d'_1 > d'_2$ ).

Un cas particulièrement intéressant est celui où  $d_2$  est infinie. On obtient alors une image nette (au sens donné plus haut) pour toutes les distances supérieures à  $d_1$ . La distance  $H$  à laquelle l'objectif doit être focalisé pour obtenir ce résultat est appelée *distance hyperfocale*, ou plus souvent *hyperfocale*. En combinant l'Equation 2 et l'Equation 3, et en nommant  $e$  le diamètre maximal admissible de la tache-image (souvent appelé *cercle de confusion*), on obtient:

$$H = \frac{f^2}{n \cdot e} \quad \text{Equation 4}$$



Lorsque la mise au point est faite sur l'hyperfocale  $H$ , on obtient une image nette de tout objet situé à une distance  $d > \frac{H}{2}$ . Pour une distance de mise au point  $d_0$  quelconque,  $d_1$  et  $d_2$  sont données par les équations suivantes:

$$d_1 = \frac{H \cdot d_0}{H + d_0} \qquad d_2 = \frac{H \cdot d_0}{H - d_0}$$

L'intérêt pratique de l'hyperfocale est évident: connaissant la netteté minimale que l'on désire obtenir, l'ouverture et la distance focale de l'objectif, il suffit de faire le point sur la distance hyperfocale, et on est assuré d'une image nette entre la distance  $d_1$  et l'infini, auquel cas la profondeur de champ  $P$  est évidemment maximale. Pour un dispositif tel que le stéréoscope linéaire, qui doit fonctionner de manière autonome et qui ne peut s'adapter à un système de mise au point automatique (toute action sur la bague de mise au point compromettrait la coplanéité des champs de visée), une telle solution est la seule possible.

Pour l'anecdote, signalons qu'un très grand nombre d'appareils photographiques de faible prix (notamment les jetables) ont un objectif à mise au point fixe réglée sur l'hyperfocale, ainsi que beaucoup d'objectifs "fish-eye" dont l'énorme profondeur de champ permet même des agrandissements en grand format dans ces conditions.

---

## A.1. Choix pratique du cercle de confusion

Le diamètre  $e$  du cercle de confusion est un paramètre très important lorsque l'on cherche à déterminer la profondeur de champ et/ou l'hyperfocale dans des conditions de prise de vue données, et sa détermination est rarement simple. Dans le cas d'un capteur CCD, dont la résolution est parfaitement connue, le bon sens amène à choisir  $e$  égal aux dimensions du pixel. Ce n'est cependant pas toujours une très bonne idée, car une exigence aussi élevée entraîne une profondeur de champ très faible, et une hyperfocale trop lointaine pour être exploitable. Par exemple, avec nos caméras (pixels de  $11\mu\text{m}$ ) et nos objectifs de 35mm ouverts à  $n=5,6$ , l'hyperfocale serait à 28m et la plage de netteté s'étendrait de 14m à l'infini, ce qui est un peu juste pour des obstacles proches. L'ajustement de  $e$  permet de choisir une plage de netteté plus réaliste. Dans l'exemple précédent, on aura une plage de netteté s'étendant de 6,50m à l'infini en se contentant de  $e=22\mu\text{m}$ . La connaissance de  $e$  permettra, au besoin, d'ajuster les paramètres des traitements, par exemple pour la détection de contours qui est très sensible à la netteté de l'image.

## **Annexe B**

# **AMELIORATIONS POSSIBLES DE LA METHODE D'ETALONNAGE DU STEREOSCOPE LINEAIRE**

La précision obtenue par le stéréoscope linéaire calibré par la méthode décrite au Chapitre 3 est suffisante pour les besoins du programme Prometheus, nous n'avons donc pas voulu la rendre plus précise au prix d'une complexité accrue. Cependant, il serait possible de l'améliorer de plusieurs manières pour augmenter la précision du stéréoscope. Voici quelques pistes que nous n'avons pas suivies, mais qui pourrait être utiles pour des applications plus exigeantes:

## B.1. Mieux garantir la coplanéité des champs de visée

Le dispositif de calibrage présenté au Chapitre 3 est, à l'évidence, minimaliste. La précision que l'on peut espérer sur la coplanéité des capteurs n'est que de l'ordre du millimètre, en raison des mouvements verticaux que l'opérateur fait effectuer aux trois vis poussant de chaque caméra lorsqu'il recherche le meilleur calibrage.

Il est possible d'espérer mieux si l'on complique la mire de calibrage en lui adjoignant un deuxième plan plus proche, évidé en son centre, qui porterait deux repères de calibrage vertical et deux jeux de lignes-guides supplémentaires, *parfaitement* alignés sur ceux de la première mire (cf. Figure B.1).

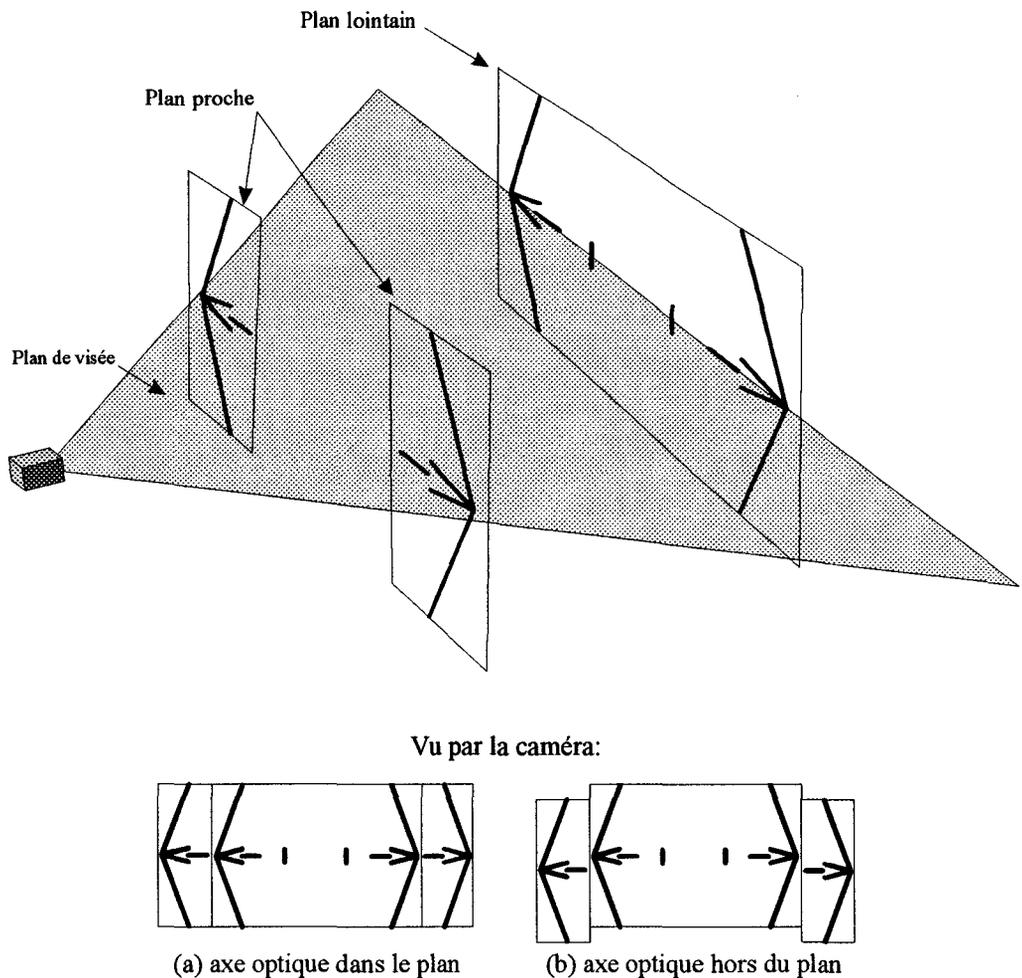


Figure B.1 : Mire à 2 plans pour mieux assurer la coplanéité des plans de visée

A l'aide de cette deuxième mire, l'opérateur doit, une fois la coplanéité réalisée sur le plan lointain, effectuer une translation verticale du plan de visée sans modifier son orientation — en tournant les trois vis poussant A, B et C *de la même hauteur exactement* ou, plus simplement, en plaçant les caméras sur un support de hauteur variable disponible chez les fournisseurs de matériel optique ou mécanique de précision — afin de retrouver les repères de calibrage vertical du plan proche (cf. Figure B.2).

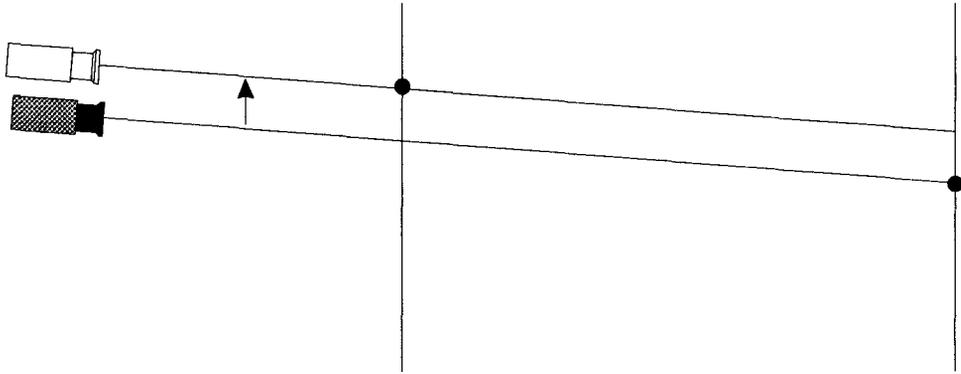


Figure B.2 : Translation verticale pour placer la mire proche dans le plan de visée

Bien entendu, les repères de calibrage vertical du plan lointain sont alors sortis du plan de visée du stéréoscope puisque ce dernier a bougé (cf. Figure B.3). En jouant sur la vis poussant C pour modifier l'angle de site  $\theta$ , il faut alors les retrouver. Les repères de calibrage proches sont alors eux-mêmes sortis du plan de visée, et une nouvelle translation en hauteur est nécessaire pour les retrouver.

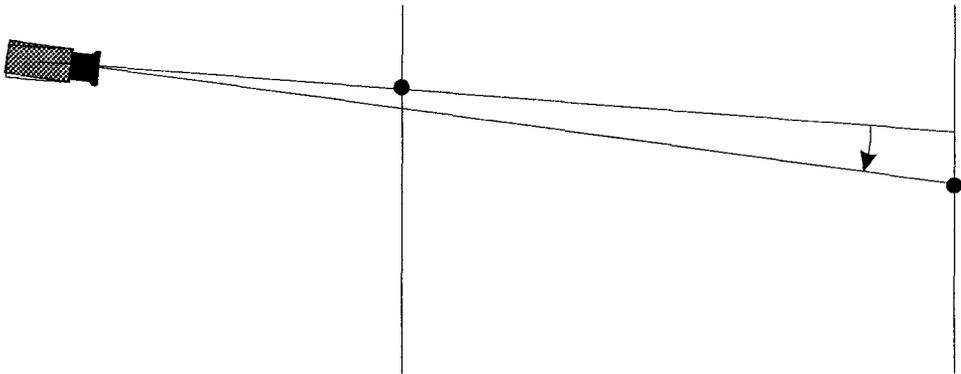


Figure B.3 : Changement d'inclinaison pour placer la mire lointaine dans le plan de visée

On continue ainsi à retrouver sur l'image les repères de calibrage vertical proches par translation, puis les repères de calibrage vertical lointains par rotation de la vis C, jusqu'à ce que les deux plans soit ajustés. On aura reconnu, à quelques détails près, la méthode itérative d'ajustage décrite plus haut (cf. § 3.5.2).

Nous n'avons pas testé cette procédure qui, outre qu'elle exige beaucoup de temps — la translation en hauteur semble très difficile à réaliser sans perturber les angles  $\psi$ ,  $\theta$  et  $\rho$ , même avec des doigts de fée, et un nouveau réglage selon la procédure du § 3.5.1 est nécessaire après chaque translation — n'a pas paru nécessaire: les essais d'étalonnage (cf. § 3.6.2) et de prise de vue (cf. Chapitre 4) n'ont jamais fait apparaître un décalage des plans de visée suffisamment important pour nécessiter un recalibrage avec une méthode plus précise que celle du § 3.5.1.

---

## **B.2. Déterminer la distance focale des objectifs**

Nous avons vu plus haut (cf. Chapitre 2) que les normes en vigueur permettent au fabricants d'objectifs de spécifier une focale avec  $\pm 5\%$  d'erreur. Cette tolérance recouvre deux notions bien différentes:

- Une erreur systématique due au calcul de l'objectif, la distance focale étant arrondie pour des raisons de simplicité (on n'indique pas les décimales) ou pour des raisons commerciales. C'est ainsi, par exemple, qu'un appareil photographique assez connu a été vendu par une marque avec un objectif donné pour 40 mm, et par une autre avec le même objectif donné pour 38 mm, le supplément de champ — imaginaire — ainsi procuré étant supposé attirer la clientèle.
- Une erreur aléatoire due aux dispersions de caractéristiques des verres, aux tolérances sur les rayons de courbures des lentilles ou sur leur montage, etc. Cette erreur est généralement faible, bien inférieure à la tolérance de 5%. Il suffit généralement de disposer de deux objectifs ayant des numéros de série proches pour éliminer quasiment ce défaut, voire le rendre négligeable si les exigences de précision ne sont pas trop élevées.

Cette incertitude sur la distance focale provoque, à elle seule, une incertitude de  $\pm 5\%$  sur la position des objets. Il est donc possible d'améliorer sensiblement la précision de mesure en mesurant la distance focale des objectifs.

Cette mesure est facile à effectuer avec notre installation de calibrage: il suffit, une fois le terminée la procédure d'alignement des plans de visée (cf. § 3.5.1 ou § 1), de mesurer la distance  $X$  entre les deux points les plus écartés de la mire, qui sont les croisements des lignes-guides de gauche et de droite (cf. Figure B.4) puis, à l'aide d'un logiciel de gestion adapté, le nombre de pixels  $n$  qui sépare leurs images sur le capteur de l'une des caméras du stéréoscope.

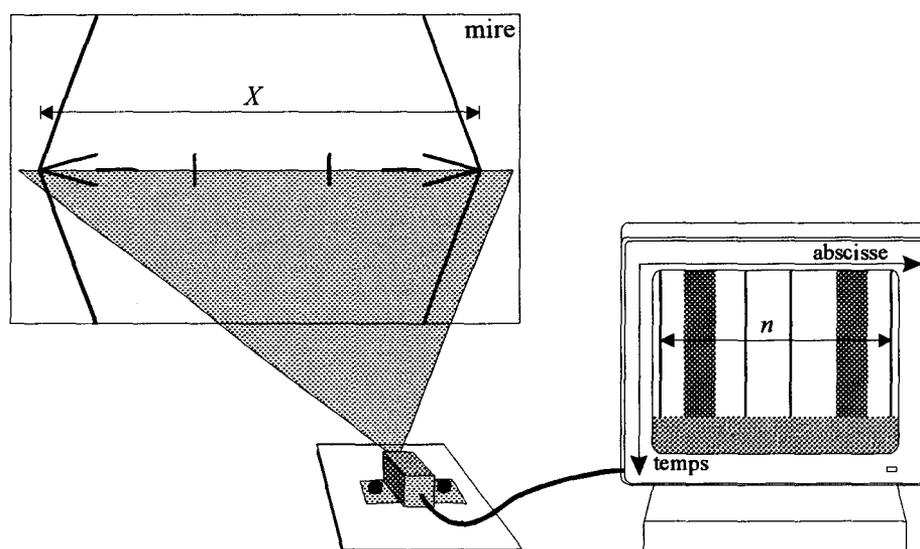


Figure B.4 : Détermination de la distance focale d'un objectif

La distance focale de l'objectif est alors obtenue par une simple règle de trois:

$$f = \frac{n \cdot p}{X \cdot D} \quad (\text{équation 1})$$

où

- $p$  est la largeur d'un photosite (donnée par le fabricant du capteur)
- $D$  est la distance caméra-cible, mesurée le long de l'axe optique.

La prise en compte de la distance focale doit être faite séparément pour chaque objectif, ce qui impose une mesure de distance focale pour chaque caméra, lors du calibrage, et une correction des images avant les calculs de reconstruction géométrique de la scène (cf. Chapitre 1), d'où une légère perte de temps.

Notons cependant que, si les deux objectifs ont des distances focales très proches (cas fréquent), une correction des images n'est pas nécessaire: il suffira de remplacer, dans les équations, la distance focale "officielle" par la moyenne des distances focales mesurées.

### B.3. Déterminer la distorsion des objectifs

Nous n'avons pas voulu prendre en compte la distorsion des objectifs, et ceci pour deux raisons:

- La prise en compte des distorsions dans les calculs prend un temps non négligeable, ce qui compromet l'objectif de l'exploitation en temps réel.
- Nous utilisons des objectifs photographiques de très bonne qualité dont la distorsion, de l'ordre de 0,4% [Chasseur 92], est jugée négligeable dans notre application dont les exigences de précision restent assez modestes.

Cependant, il serait possible d'ajouter sur la mire de calibrage existante des lignes verticales régulièrement espacées qui permettraient de déterminer la distorsion des objectifs par une méthode dérivée de celles qui existent en stéréovision matricielle (voir par exemple [Tsai 85], [Lenz 87], [Nomura 92]).

## **Annexe C**

# **MESURE DES HAUTEURS A L'AIDE DU STEREOSCOPE LINEAIRE**

Les équations du Chapitre 2 permettent, en fonction de la position ( $X, Z$ ) des objets présents dans le champ, de déterminer leur hauteur  $Y$ . Cependant, la hauteur ainsi calculée n'est pas la hauteur réelle des objets, mais seulement celle de leur intersection avec le plan de visée du stéréoscope, qui est toujours inférieure (cf. Figure C.1).

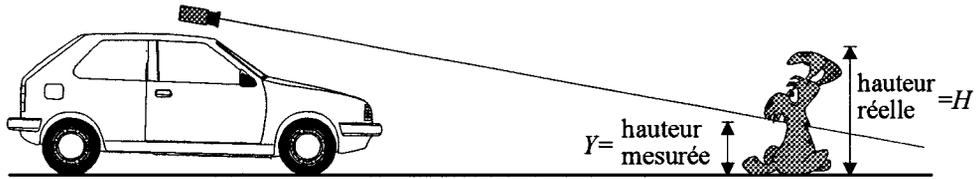


Figure C.1 : La hauteur mesurée par le stéréoscope n'est pas la hauteur de l'objet.

Il existe cependant plusieurs moyens de déterminer la hauteur d'un objet (au sens optique du terme, qui s'applique tout aussi bien à un véhicule, à un animal ou à un piéton).

## C.1. Mesure directe par le stéréoscope

La méthode la plus simple consiste à suivre ce dernier dans le temps, et de mémoriser la valeur maximale de sa hauteur mesurée  $Y_{max}$ . Il est cependant impossible de savoir si  $Y_{max}$  correspond réellement à la hauteur de l'objet: elle ne lui est qu'inférieure ou égale, par définition. L'intérêt pratique de cette méthode est donc faible.

Notons de plus que la hauteur mesurée  $Y$  dépend de plusieurs paramètres difficiles à mesurer:

- Les mouvements de tangage du véhicule entraînent une variation permanente de l'angle de site  $\theta$ , qui intervient dans les calculs de  $Y$ . La prise en compte de  $\theta$  nécessiterait une mesure de l'assiette du véhicule, ce qui ne semble pas possible actuellement: bien que des capteurs existent déjà pour les besoins des suspensions actives, leurs performances (en précision notamment) sont encore insuffisantes.
- L'hypothèse de la chaussée plane n'est souvent correcte qu'en première approximation, notamment dans les côtes (cf. Figure C.2). Or, il est très difficile de déterminer le profil vertical de la chaussée devant un véhicule, surtout sur les quelques dizaines de mètres de portée du stéréoscope, et avec la précision requise pour que la hauteur  $Y$  mesurée soit exploitable.

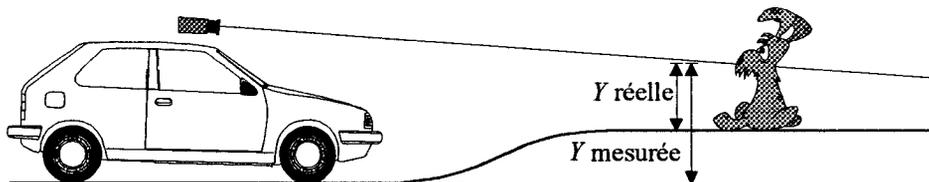


Figure C.2 : Erreur sur les mesures de hauteur sur une chaussée non plane.

## C.2. Mesure à l'aide d'une caméra matricielle calibrée

Une solution plus réaliste serait d'utiliser une caméra matricielle calibrée par rapport au stéréoscope linéaire. En supposant l'objet vertical, sa hauteur sur la chaussée  $H$  serait facile à obtenir à partir de sa profondeur  $D$ , de la hauteur  $h$  de son image sur le capteur matriciel, et de la distance focale  $f$  de l'objectif (cf. Figure C.3).

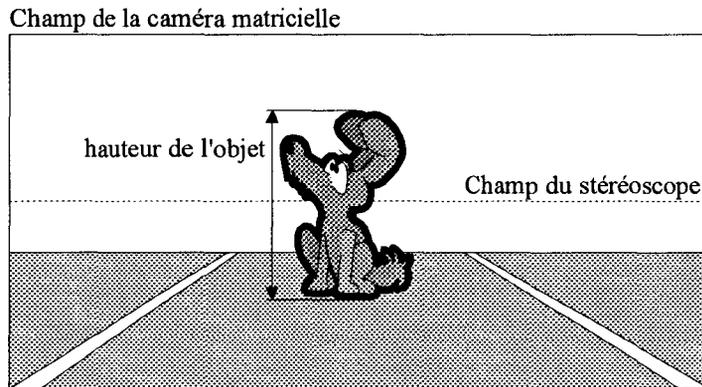


Figure C.3 : Mesure de hauteur avec le stéréoscope linéaire et une caméra matricielle calibrée.

A cet usage, la caméra matricielle devrait être située dans le plan de visée du stéréoscope, de préférence au centre du repère  $OXYZ$  pour éviter les calculs liés à un changement de référentiel. Les contours de l'objet seraient fournis par une segmentation de l'image autour de la "tranche" vue par le stéréoscope, sa hauteur serait donnée par l'équation de projection inverse classique:

$$H = h \frac{D}{f}$$