

THESE

présentée à

**L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE
FLANDRES ARTOIS**

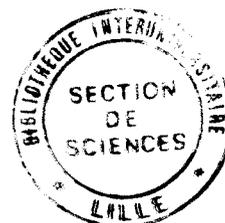
pour l'obtention du titre de

DOCTEUR

En Productique : Automatique et Informatique Industrielle

par

Luc DUVIEUBOURG



**ANALYSE DE SEQUENCES D'IMAGES LINEAIRES
STEREOSCOPIQUES. APPLICATION A LA
REALISATION D'UN SYSTEME DE DETECTION
D'INTRUSIONS INTELLIGENT POUR LES
TRANSPORTS GUIDES**

Soutenue publiquement le 23 Janvier 1991 devant la commission d'examen:



MM.	P. VIDAL	Président
	Y. DAVID	Rapporteur
	B. DOLPHIN	Rapporteur
	J.-G. POSTAIRE	Directeur de Recherche
	J.-P. PERRIN	Examiteur
	J.-P. DEPARIS	Examinateur

Ce travail a été effectué au sein du Centre d'Automatique de Lille (CAL), en collaboration avec, l'Institut National de REcherche sur les Transports et leur Sécurité (INRETS), le Centre de Recherche et d'Evaluation des Systèmes de Transports Automatisés (CRESTA) et le Groupement Régional Nord-Pas de Calais pour la Recherche dans les Transports (GRRT). L'auteur remercie ces trois derniers organismes pour leur soutien matériel et la confiance qu'ils lui ont accordée tout au long de cette étude.

REMERCIEMENTS

Je tiens tout d'abord à remercier Monsieur le Professeur **P. VIDAL**, directeur du Centre d'Automatique de Lille qui m'a permis d'effectuer mes recherches au sein de son laboratoire.

Je remercie sincèrement Monsieur le Professeur **J.-G. POSTAIRE**, pour l'intérêt qu'il a porté à mes travaux, pour sa disponibilité ainsi que pour ses encouragements pendant les moments difficiles.

J'exprime ici ma gratitude à Monsieur **Y. DAVID**, de l'institut National de REcherche sur les Transports et leur Sécurité, pour avoir soutenu cette étude et d'avoir accepté de juger mon travail.

Je suis très honoré de la présence dans ce jury de Monsieur le Professeur **B. DOLPHIN**, de l'université d'Amiens, et je tiens à lui présenter mes remerciements.

Je remercie également Monsieur **J.-P. PERRIN**, du service Recherche et Développement de la RATP, pour sa participation au jury.

Je suis très reconnaissant à Monsieur **J.-P. DEPARIS**, d'avoir bien voulu siéger à cette commission d'examen.

Mes remerciements vont ensuite à l'ensemble des chercheurs, techniciens et personnels du Centre d'Automatique de Lille.

Enfin une mention particulière est attribuée aux chercheurs C. Vieren, F. Cabestaing et L. Macaire pour leur amitié.

Table des Matières

CHAP I	- LES SYSTEMES DE SECURITE DANS L'AUTOMATISATION DES TRANSPORTS GUIDES.	p. 7
I) INTRODUCTION		p. 8
II) CAHIER DES CHARGES		p. 10
II.1) Configuration des stations.		p. 10
II.2) Types d'intrusions à détecter.		p. 13
II.3) Minimisation des fausses alarmes.		p. 13
II.4) Conclusion.		p. 14
III) SOLUTIONS EXISTANTES		p. 15
III.1) Détection infrarouge.		p. 15
III.1.1) Nappe infrarouge passif.		
III.1.2) Rideau infrarouge actif.		
III.2) Vision artificielle.		p. 18
III.3) Autres solutions.		p. 20
III.3.1) Tapis sensible.		
III.3.2) Câble rayonnant.		
III.4) Conclusion.		p. 23
IV) ORGANISATION DE LA THESE .		p. 23
CHAP II	- LA PERCEPTION VISUELLE D'UN ENVIRONNEMENT 3D	p. 25
I) INTRODUCTION.		p. 26
I.1) Les méthodes monoculaires.		p. 27
a) Scène statique		
b) Scène dynamique		

c) <i>Caméra mobile</i>	
I.2) Les méthodes binoculaires.	p. 28
II) LA VISION BINOCULAIRE OU STEREOVISION.	p. 28
II.1) Description du principe	p. 28
II.2) Reconstitution des informations 3D.	p. 29
II.2.1) Détection des indices et appariement.	
a) <i>Appariement de points caractéristiques.</i>	
b) <i>Appariement de contours.</i>	
c) <i>Appariement de régions.</i>	
II.2.2) Calcul des distances.	
II.3) Conclusion.	p. 38
III) VISION MONOCULAIRE.	p. 39
III.1) Reconstitution de forme à partir d'une image	p. 39
a) <i>Par analyse de la réflectance.</i>	
b) <i>Par analyse de la texture.</i>	
III.2) Mesure de profondeur par focalisation automatique.	p. 43
III.3) Vision active avec lumière cohérente.	p. 44
III.4) Analyse d'une séquence d'images.	p. 46
III.4.1) Méthode des primitives caractéristiques.	
a) <i>Extraction et mise en correspondance de primitives caractéristiques.</i>	
b) <i>Inférence 3D par estimation du mouvement.</i>	
c) <i>Conclusion.</i>	
III.4.2) Méthode du flot optique.	
a) <i>Calcul du flot optique.</i>	
b) <i>Reconstruction des informations 3D.</i>	
c) <i>Conclusion.</i>	

c) *Conclusion.*

IV) CONCLUSION. p. 55

CHAP III - BARRIERE OPTIQUE DE SURVEILLANCE p. 57

I) INTRODUCTION p. 58

II) DEFINITION D'UN PLAN DE SURVEILLANCE p. 59

II.1) Principe de détection p. 59

II.2) Focalisation. p. 61

II.3) Segmentation des images lignes. p. 68

III) DOUBLE PLAN DE SURVEILLANCE p. 71

III.1) Introduction p. 71

III.2) Caméras visant la même rampe lumineuse. p. 72

III.2.1) Principe.

III.2.2) Détermination de la position des objets.

a) *Calibration mécanique.*

b) *Calibration optique.*

III.2.3) Exemple de configuration en station.

III.3) Caméras montées tête-bêche. p. 78

III.3.1) Principe.

III.3.2) Détermination des attributs.

III.3.3) Exemple de configuration en station.

IV) CONCLUSION p. 81

CHAP IV - DECLENCHEMENT D'ALARME PAR ANALYSE DES SEQUENCES DE COUPLES D'IMAGES LIGNE p. 83

I) INTRODUCTION p. 84

II) ANALYSE SPATIO-TEMPORELLE. p. 87

III) STRATEGIE. p. 90

III.1) Représentation de l'information. p. 90

III.2) Modélisation par un automate. p. 92

III.2.1) Définition générale d'un automate.

III.2.2) Reconnaissance de séquences de transitions par automate.

a) *Définitions.*

b) *Reconnaissance d'une intrusion.*

III.2.3) Automate à états fini avec sortie.

a) *Définition générale d'une Machine de Moore.*

b) *Sorties de l'automate de reconnaissance d'intrusion.*

III.3) Analyse de différents types de séquences. p. 99

III.4) Analyse des limites de la procédure de détection. p. 100

III.4.1) Absence de phase d'entrée ou de sortie.

III.4.2) Absence de phase de présence.

III.5) Automate final. p. 104

CHAP V - IMPLANTATION TEMPS REEL ET RESULTATS EXPERIMENTAUX p. 107

I) INTRODUCTION. p. 108

II) IMPLANTATION. p. 108

II.1) Aspect matériel.	p. 108
II.1.1) Machine de vision linéaire	
a) <i>Une unité centrale: CPU.</i>	
b) <i>Une unité caméra ou interface intelligente caméra: carte IICAM.</i>	
c) <i>une unité de visualisation.</i>	
II.1.2) Configuration de l'application.	
II.1.3) Unité d'acquisition simultanée d'images ligne: carte SIAM.	
a) <i>Mode transparent.</i>	
b) <i>Mode acquisitions simultanées.</i>	
c) <i>Mode restitution.</i>	
d) <i>Mode lecture par le bus VME.</i>	
II.2) Aspects logiciels.	p. 114
II.2.1) Contrainte temps réel.	
II.2.2) Traitement multiprocesseur.	
III) RESULTATS EXPERIMENTAUX.	p. 117
a) <i>Effet de la vitesse de passage.</i>	
b) <i>Effet de la position longitudinale.</i>	
- CONCLUSION GENERALE	p. 131
- BIBLIOGRAPHIE	p. 136
- ANNEXES	p. 147

CHAPITRE I

**▣ LES SYSTEMES DE SECURITE DANS
L'AUTOMATISATION DES TRANSPORTS GUIDES.**

I) INTRODUCTION

Le travail présenté dans ce mémoire s'intègre dans un programme d'automatisation des moyens de transport guidé. L'automatisation d'un processus quelconque, consiste essentiellement à remplacer l'homme par un système susceptible d'agir à sa place. Dans le cas des moyens de transport guidé l'attention du conducteur est essentiellement absorbée par la fonction de surveillance des voies. L'automatisation de cette fonction fondamentale implique donc de concevoir et de réaliser des systèmes de surveillance automatique pour assurer que la voie est libre. Ce mémoire apporte une contribution à l'automatisation de cette fonction de surveillance des voies.

On constate, depuis le début de la décennie, que dans les nouveaux systèmes de transport guidé mis en chantier, des fonctions de plus en plus nombreuses sont progressivement prises en charges par des automatismes. Les premiers systèmes à automatisation intégrale ont été principalement installés sur des sites particuliers comme des aéroports, des parcs d'exposition. Ces systèmes automatiques sont maintenant étendus aux transports urbains. Plusieurs villes telles que Lille, Osaka, Vancouver et, plus récemment, Londres sont aujourd'hui équipées de nouveaux systèmes de transport à pilotage automatique.

Tous ces moyens de transport automatique circulent en sites protégés de telle sorte que les interactions avec le public sont très réduites. Le principal problème de ces systèmes à pilotage entièrement automatique se situe au niveau des stations où il s'agit d'assurer la sécurité des voyageurs circulant sur les quais et vérifier que la voie est libre.

Le métro de Lille est un exemple de transport guidé où il n'y a aucune possibilité d'interaction avec le public au niveau des stations. La protection des voies est réalisée par des portes palières, ce qui permet d'atteindre un niveau de sécurité bien supérieur à celui offert par les métros conventionnels. Dans ce cas, la nécessité d'un arrêt précis des véhicules devant les portes palières implique des équipements particuliers d'arrêt en station qui sont chargés de gérer la séquence d'arrêt et l'ouverture des portes.

Le coût important de réalisation des stations équipées de portes palières ainsi que celui des équipements électroniques nécessaires au pilotage précis des rames conduisent à une réduction du marché potentiel de ce système de protection. Pour le métro de Lille, l'équipement d'une station en portes palières coûte environ 400KF. Ce coût élevé s'explique par l'installation des portes palières elles-mêmes mais également par le fait que l'arrêt des rames exactement en face des portes du quai entraîne un certain nombre de contraintes.

Le système de conduite entièrement automatique pour le nouveau réseau de Lille a été conçu dès le départ, ce qui facilita son étude et sa réalisation. Par contre, pour les anciens réseaux, l'approche des problèmes de sécurité doit tenir compte de l'existence des infrastructures. Pour des raisons architecturales ou financières, il n'est pas toujours possible d'isoler en partie ou totalement le site de circulation des trains des espaces accessibles aux voyageurs. En effet, ce système, bien que très efficace, ne peut être installé sur les anciens métros où la longueur des stations entraînerait d'énormes coûts de construction.

Cependant, plusieurs métros entièrement automatiques en service ou en construction, comme ceux de Vancouver, Miami, Détroit et bientôt la ligne D du métro de Lyon ont été réalisés ou conçus sans portes palières. Avec la disparition d'une surveillance humaine de la conduite et en absence de portes palières, des dispositifs assurant que la voie est libre, de façon sûre et fiable, deviennent indispensables. Ces chantiers ont stimulé la recherche sur les systèmes de surveillance par barrières immatérielles, destinées à détecter les intrusions dans les zones protégées. Dans ce cas, la règle généralement admise est que l'on doit garantir aux passagers un niveau de sécurité au moins égal à celui qui est assuré par un conducteur.

La définition du cahier des charges de ces nouveaux systèmes de sécurité doit s'adapter à l'environnement particulier des stations de métro. En particulier l'importance des zones à surveiller et la diversité des accidents pouvant survenir sont des éléments primordiaux à prendre en compte lors de la conception de ces systèmes de surveillance.

II) CAHIER DES CHARGES

La conception de systèmes permettant d'atteindre un niveau de sécurité aussi élevé que possible sans utiliser de portes palières concerne essentiellement les fonctions suivantes:

- détection de chutes de voyageurs ou d'objets importants sur la voie en station,
- détection des entraînements de passagers au démarrage de la rame,
- détection d'intrusions de voyageurs sur les voies en dehors des stations et
- détection des intrusions dans les espaces entre le quai et la rame.

L'étude présentée dans ce mémoire concerne la première fonction. Il s'agit de mettre au point un système de détection d'intrusions sur les voies pour éventuellement arrêter les rames en cas d'obstacle sur leur passage. La détection de la présence d'un objet sur la voie d'un métro nécessite un recensement de toutes les situations possibles qui pourraient interférer avec le fonctionnement des rames.

Pour cela, le système de surveillance des voies devra s'adapter à la diversité géométriques des stations et réagir aux différentes intrusions de manière à minimiser le nombre de fausses alarmes tout en assurant une détection sans faille.

II.1) Configuration des stations.

La zone surveillée dans une station par un système de sécurité doit être au moins équivalente à celle que surveillerait un conducteur. Cette zone correspond à la voie et au bord du quai (Fig. 1-1). Elle varie avec la taille et la forme d'une station. On peut toutefois faire une distinction entre les stations des métros existantes et celles des métros conçus dès le départ pour un pilotage entièrement automatique. Pour cette dernière catégorie de métros, la standardisation des stations facilite le développement d'un système unique de détection de chutes sur les voies.

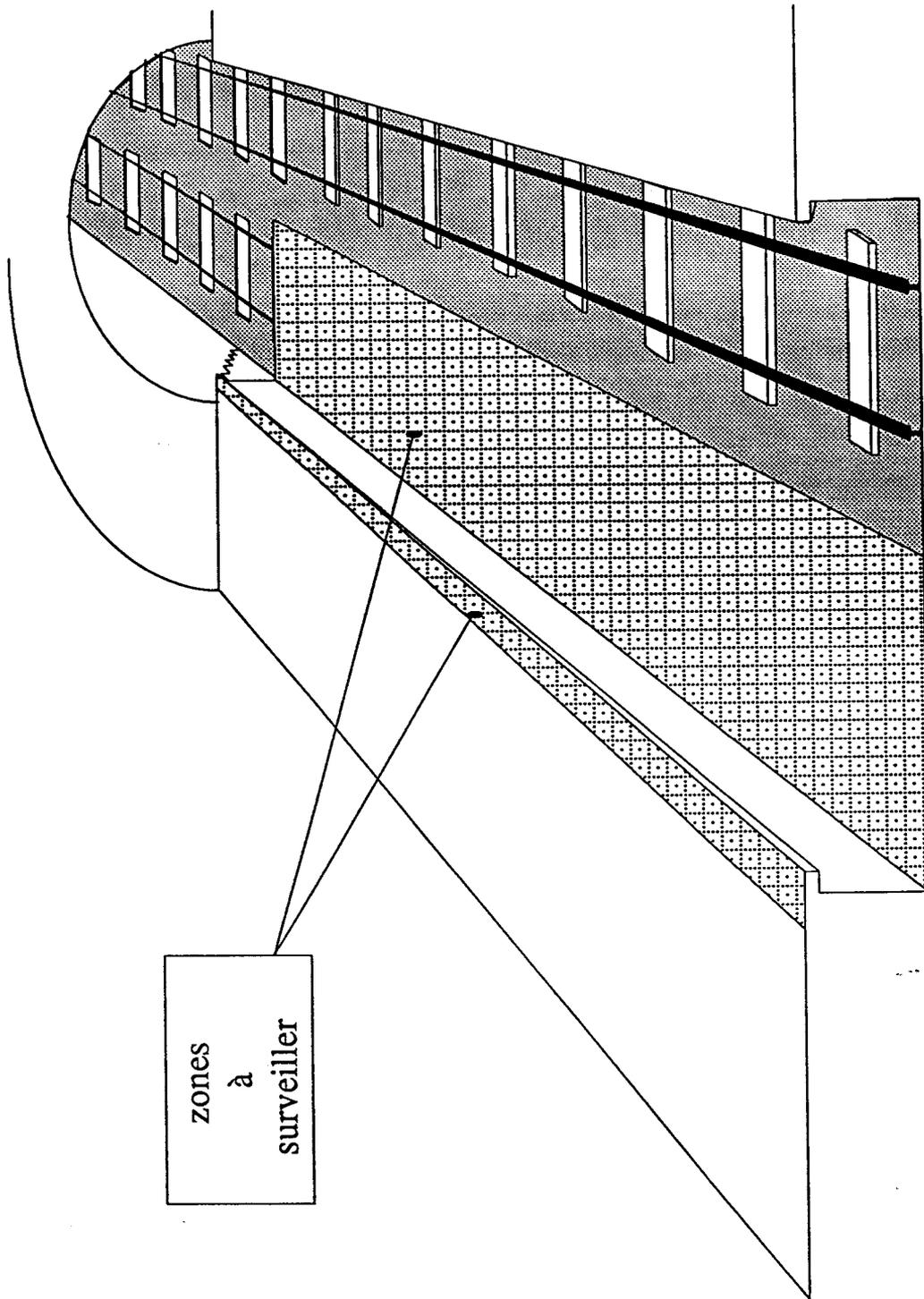


figure- 1-1 - Surveillance de la station

Par contre, pour les anciens réseaux, le système de sécurité devra répondre à certaines contraintes:

- Adaptation à la taille des stations.

En effet, les stations pouvant être de longueur variable, des aménagements éventuels du système devront être possibles sans remettre en question le principe de détection.

- Facilité d'intégration dans les infrastructures existantes.

Il n'est pas concevable de modifier l'architecture d'une station pour installer le système de sécurité. Celui-ci devra nécessiter le moins possible de travaux de génie civil.

- Rapidité de mise en place.

L'exploitation du réseau ne doit pas être perturbée pendant un intervalle de temps trop important. L'idéal serait de pouvoir réaliser les travaux pendant les interruptions normales du trafic.

D'autre part, quel que soit le type de réseau, le système de sécurité devra être insensible aux conditions d'éclairage des stations. Dans ce cas, on peut encore classer les stations en deux catégories:

- stations souterraines et

- stations à ciel ouvert.

Pour les stations souterraines, seule l'extinction d'une ou plusieurs sources de lumière artificielle provoque des variations de lumière. Par contre pour les stations extérieures, les passages de nuages provoquent des variations "rapides" de lumière et le cycle jour/nuit provoque une variation plus "lente" de luminosité. De plus, pour les stations extérieures les conditions atmosphériques (pluie, neige, vent ...) peuvent perturber le fonctionnement du système de sécurité. Celui-ci devra donc être insensible aux conditions atmosphériques pour ne pas générer de fausses alarmes intempestives.

II.2) Types d'intrusions à détecter.

La prise en compte de la sécurité est une question difficile à résoudre dans la mesure où il n'existe que peu de règles de portées générales, et où ce n'est bien souvent qu'*a posteriori*, en examinant les accidents, qu'on peut juger du caractère plus ou moins sûr d'un aménagement. Toutefois pour juger de la performance du système de protection on peut dresser une liste de cas types d'intrusions:

- intrusion partielle d'un usager (assis sur le bord du quai),
- intrusion d'un usager sur les voies,
- chute d'objets importants (valise, journal),
- chute de petits objets (ticket de métro, paquet de cigarettes),
- passage d'animaux (oiseau, rat, chien, chat ...).

Cette liste, non exhaustive, nous permet de constater que si toutes les intrusions sont détectées, la chute de petits objets ou le passage de petits animaux ne doit pas entraîner l'arrêt de la rame alors que celle d'un passager ou d'un objet important doit immédiatement déclencher une procédure d'alarme. Le système de surveillance devra donc faire la distinction entre ces différents cas. D'autre part, il est essentiel que toutes les intrusions sur les voies susceptibles de présenter un caractère dangereux pour la sécurité des usagers dans la station aussi bien que dans la rame soient détectées. Le système doit garantir un taux nul de détections manquées.

II.3) Minimisation des fausses alarmes.

Il est primordial de chercher à limiter la fréquence des perturbations sur le réseau de transport en procédant à une identification préalable des différents types d'intrusions. En effet, les petits objets ne doivent pas entraîner d'alarme, alors que les

gros objets doivent en général provoquer l'arrêt immédiat de la rame. Mais ce critère de taille reste très subjectif car un journal, bien qu'il puisse être considéré comme un gros objet, ne présente aucun danger pour la rame et ne doit pas déclencher d'alarme. Dans ce contexte la minimisation des fausses alarmes passe par une analyse plus poussée des différentes intrusions recensées au paragraphe précédent. La vitesse de pénétration d'un objet dans l'espace des voies et la localisation tridimensionnelle de cet objet peuvent être des critères supplémentaires qui permettent une meilleure discrimination entre les différents types d'intrusions.

II.4) Conclusion.

Les caractéristiques principales du système de sécurité sont la fiabilité de la détection, la discrimination entre les différents types d'intrusions, l'intégration aisée dans les infrastructures existantes pour les anciens réseaux, mais d'autres critères comme une maintenance facile et la durée de vie du dispositif pourront être pris en compte lors de l'évaluation des systèmes de sécurité.

L'intégration de toutes les caractéristiques détaillées ci-dessus dans un même système de protection est ambitieux. Chacune des solutions technologiques envisagées devra satisfaire autant que possible ce cahier des charges sans toutefois atteindre un coût trop important en investissement.

III) SOLUTIONS EXISTANTES

Plusieurs systèmes de détections ont été développés pour résoudre les problèmes de surveillance dans d'autres domaines que celui des transports. Des systèmes automatiques ont été installés pour la surveillance de locaux ou de zones à risques telles que les aires de travail des machines automatiques. Ces systèmes peuvent être adaptés à l'environnement des stations de métro et plusieurs d'entre eux ont été testés et évalués par différentes sociétés d'exploitation de réseaux de métros.

III.1) Détection infrarouge.

III.1.1) Nappe infrarouge passif.

Avec un dispositif utilisant une nappe infrarouge, la zone de détection est plane et de forme rectangulaire. Elle couvre la largeur d'une voie de métro, soit environ 4m, sur 6m, avec une épaisseur moyenne de 5cm (Fig. 1-2). Basé sur la détection de changements de température, le système réagit au passage d'un corps, mais aussi à l'arrivée d'un train en station. Ce système a été testé par la société SEMALY responsable de l'exploitation de la ligne D du métro de Lyon et fut rejeté car les objets de 20 cm à 30 cm sont aussi détectés lorsqu'ils passent près du capteur

III.1.2) Rideau infrarouge actif.

Ce système, élaboré par la société JAY ELECTRONIQUE, est dérivé du dispositif développé pour la remise en marche automatique des escaliers mécaniques. Il utilise la détection par coupure de faisceaux infrarouges modulés disposés perpendiculairement à l'axe des voies en dessous du niveau des quais. Les sens d'émission sont inversés entre deux faisceaux consécutifs, pour constituer deux ensembles, l'un de faisceaux "droits" et l'autre de faisceaux "gauches" (Fig.1-3). L'ensemble des cellules est raccordé à un microprocesseur qui les scrute périodiquement.

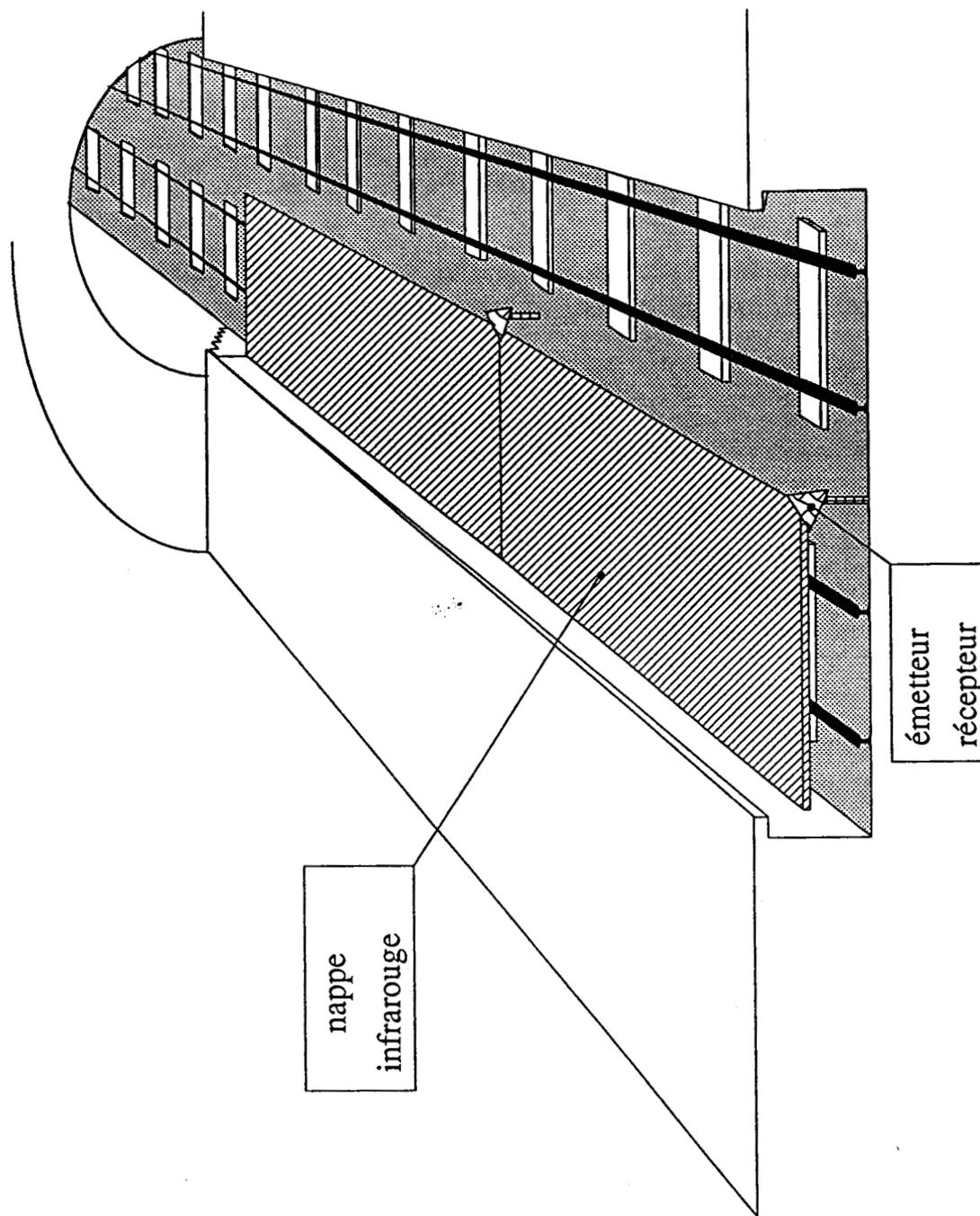


figure - 1 - 2 - Nappe infrarouge passif

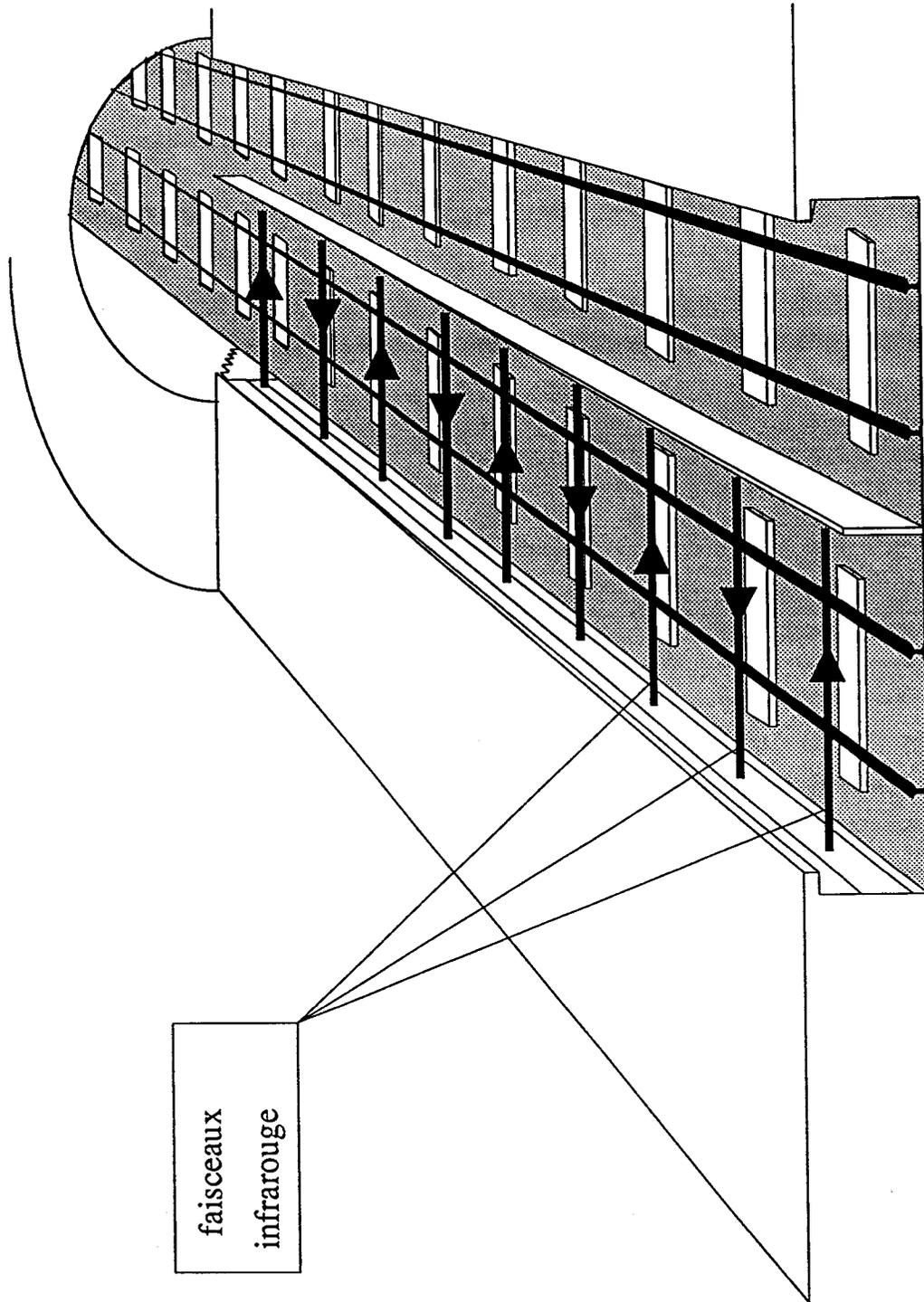


figure - 1-3 - Rideau infrarouge actif

On considère qu'il y a chute d'objet sur la voie lorsque deux faisceaux consécutifs, droit ou gauche, sont occultés simultanément, ce qui donne la résolution spatiale du système de détection. Cette procédure ne permet pas de différencier la chute d'un passager de celle d'un journal car le dispositif ne donne aucune information sur la vitesse de pénétration dans la zone surveillée. D'autre part, ces dispositifs nécessitent souvent des investissements importants compte tenu de la taille des stations de métro.

Toutefois, cette technique présente des possibilités très intéressantes. En effet, en prolongeant le dispositif dans le tunnel, on peut détecter l'arrivée de la rame et inhiber les faisceaux en avant de celle-ci au fur et à mesure de sa progression. Le système se réinitialise d'une manière identique après le passage de la rame. De plus l'occultation des faisceaux à chaque passage de rame permet un contrôle quasi permanent de leur état de fonctionnement.

III.2) Vision artificielle.

Une autre solution au problème de la détection des intrusions consiste à utiliser un système de vision artificielle. Le dispositif DAVID, conçu par la société Canadienne Computing Devices Systems, utilise cette technologie. Il a été développé pour la protection de sites stratégiques et a été adapté au problème de surveillance des interfaces quais-voies dans les stations de métros.

Il utilise des caméras matricielles de bonne définition, situées à la verticale du bord du quai, pour saisir l'image de l'interface quai-voie (Fig. 1-4). Une procédure d'installation est nécessaire pour s'adapter à la géométrie du site.

Pendant cette procédure l'image analysée est divisée en sous-images adjacentes. Les variations du niveau de gris moyen dans chacune des sous-images sont analysées, en fonction du temps, pour mettre en évidence les mouvements d'objets dans la scène observée. Une alarme est déclenchée dès qu'un nombre prédéterminé de sous-images changent consécutivement d'éclairement moyen.

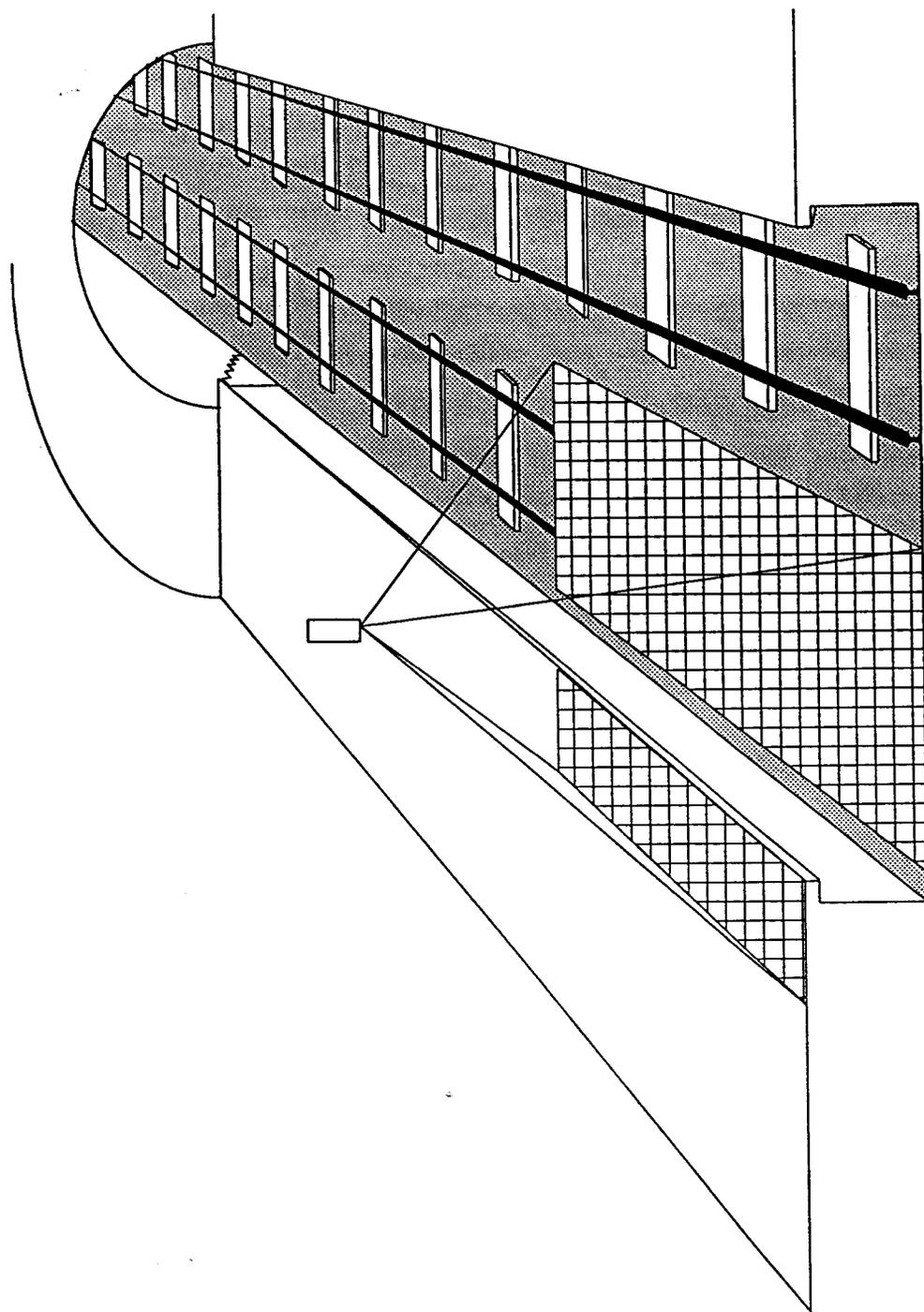


figure - 1-4 - Vision artificielle

Le système est adapté lors de son installation, à la configuration d'une station en sélectionnant, dans la scène observée, les sous-images actives qui correspondent aux zones à surveiller. Le nombre et la surface des sous-images qui doivent être affectées par des variations de niveau de gris pour déclencher une alarme permet d'ajuster la taille minimale des objets à détecter.

Après l'ajustement des paramètres internes, le système génère une alarme à chaque fois qu'une personne ou un objet de grande taille tombe sur la voie.

Mais le défaut majeur de ce système est de rester très sensible à toutes les variations de l'éclairage ambiant. En effet, le système DAVID est incapable de différencier une variation de l'intensité lumineuse due au mouvement d'un objet d'une variation due au reflet sur une vitre, à une ombre ou tout autre événement capable de modifier localement l'éclairage de la scène. Il en résulte un taux de fausses alarmes très important qui limite son intérêt pratique.

III.3) Autres solutions.

III.3.1) Tapis sensible.

Les tapis sensibles à la pression sont disposés respectivement sur les voies et en bordure de quai sur toute la longueur de la station (Fig 1.5). L'écrasement du tapis génère une alarme. Le principe de la détection repose soit sur la variation de la capacité électrique de tapis constitués de deux feuilles métalliques formant un condensateur, soit sur la variation de la pression de tapis constitués de capsules à air.

Ce dispositif est bien souvent inapproprié pour les stations de métro existantes car la configuration des voies et l'environnement sale et encombré des rails ne sont pas propices à l'installation de ce type de tapis.

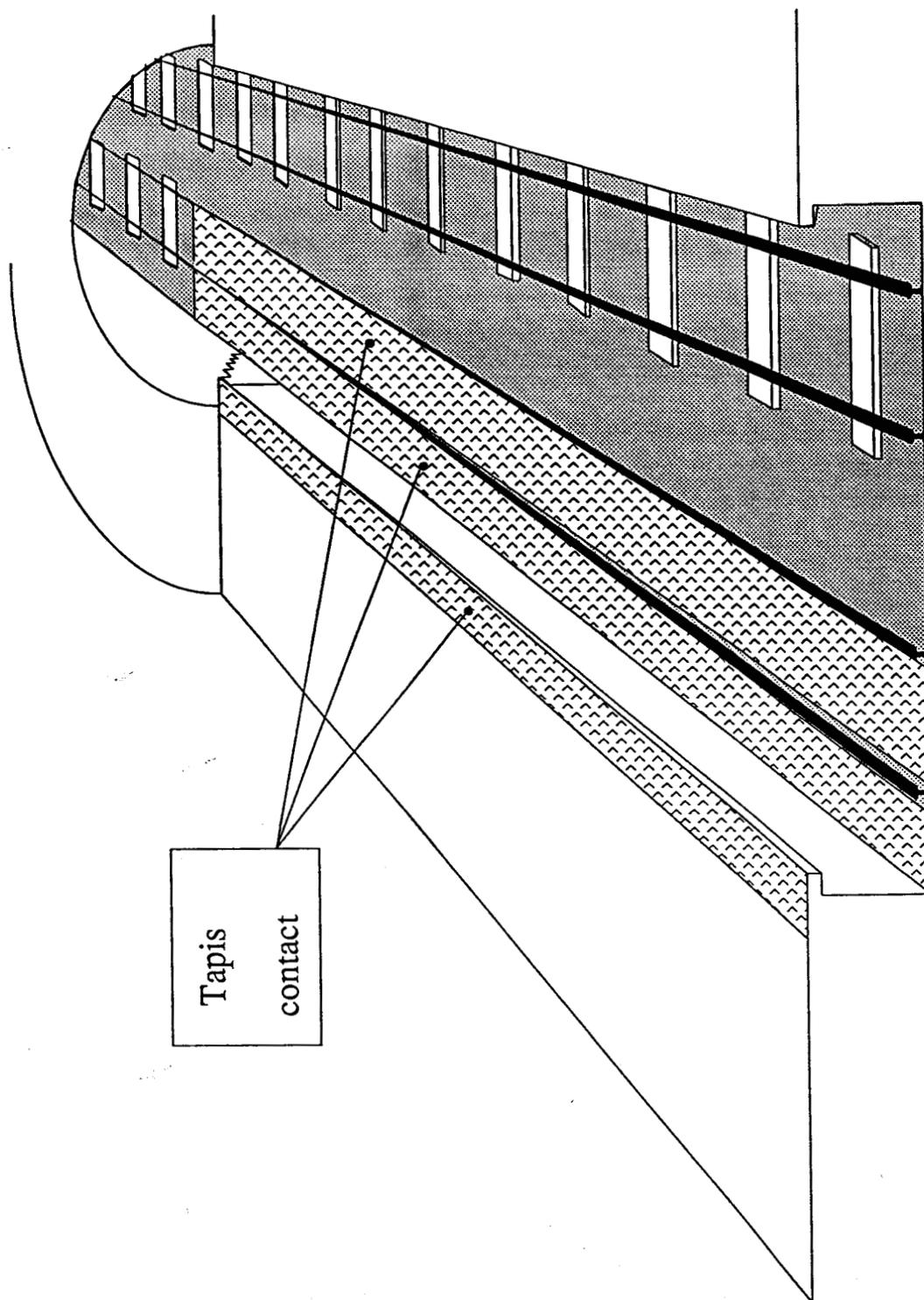


figure - 1-5 - Tapis sensible

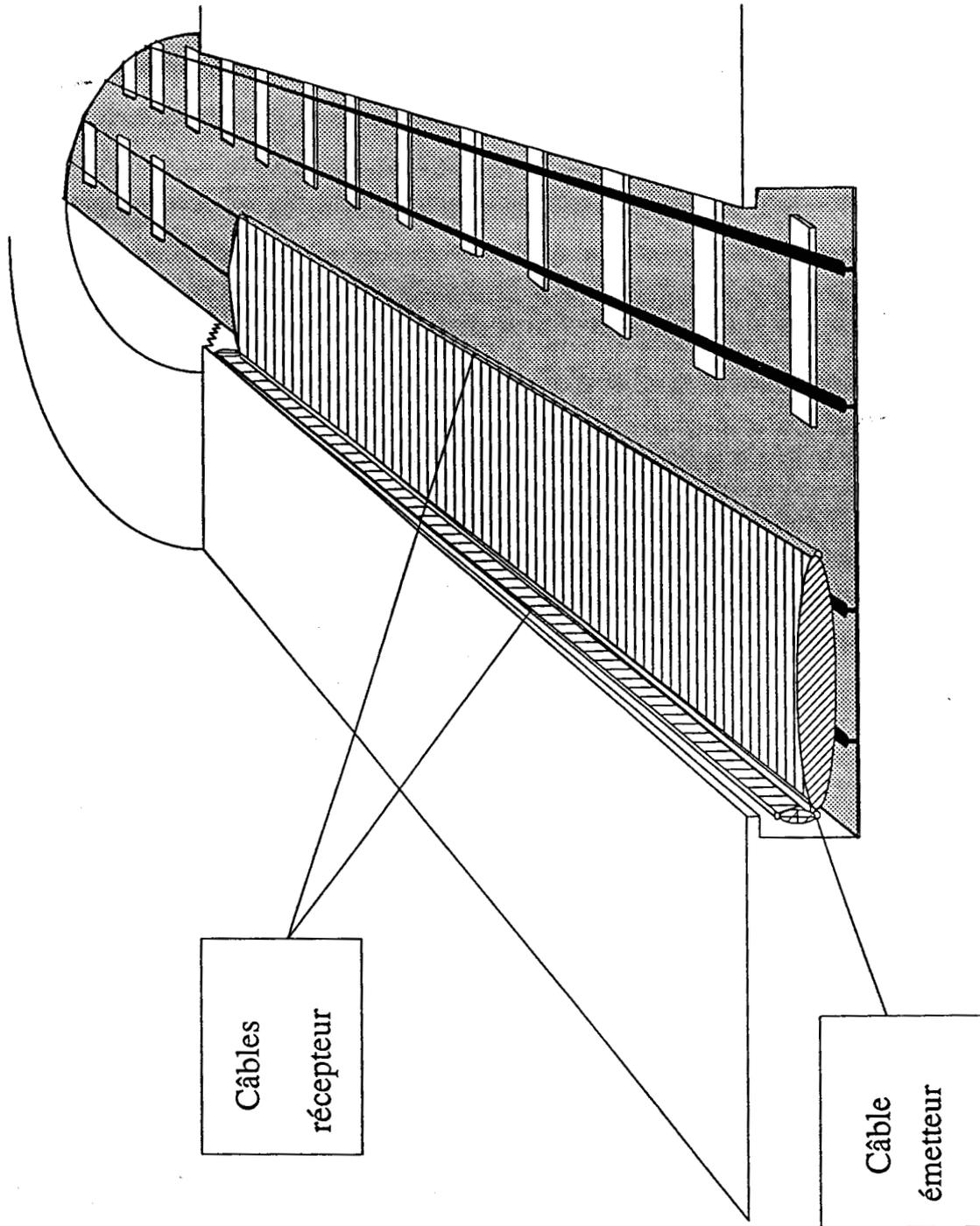


figure - 1 - 6 - Câble rayonnant

III.3.2) Câble rayonnant.

Un câble rayonnant émetteur est disposé sur le ballast le long du sous-quai. Deux câbles récepteurs sont installés respectivement le long de l'entre-voie et sous le nez de quai (Fig 1.6). Le système de protection utilise un principe de la détection électromagnétique. Les câbles émetteurs et récepteurs haute fréquence (40 Mhz) détectent la perturbation de la liaison haute fréquence entre les deux antennes formées par ces câbles.

Avec ce dispositif, la détection des intrusions théoriques risque d'être fortement perturbée par la présence d'une rame sur la voie opposée ainsi que par les câbles électriques installés sous le nez de quai.

III.4) Conclusion.

Tous ces systèmes ont été testés soit par la SEMALY, maître d'oeuvre du projet d'automatisation de la ligne D du métro de Lyon [BRA 89], soit par la R.A.T.P. (Régie Autonome des Transports Parisiens) [PIS 87]. Ils ont tous été rejetés à l'exception du rideau infrarouge actif qui a été retenu par la SEMALY pour le métro de Lyon, malgré ses limitations.

IV) ORGANISATION DE LA THESE .

Aucun des dispositifs de détection d'intrusion actuellement disponibles ne satisfait complètement le cahier des charges défini au paragraphe I -2 . Toutefois, à la suite de nombreux essais, la SEMALY a retenu le rideau infrarouge actif pour équiper la ligne D du métro de Lyon.

Cependant les responsables de la société d'exploitation du métro de Lyon ont conscience que ce dispositif ne permet pas de différencier les types très variés

d'intrusions que l'on peut observer dans les stations. En effet, il suffit d'occulter volontairement 2 faisceaux pour arrêter le métro. Rien n'empêche donc des passagers malveillants de bloquer le système.

Les systèmes de vision artificielle permettent, *a priori*, une surveillance plus fine de l'interface quai-voie. En effet, des caméras peuvent être utilisées pour obtenir des informations sur la taille des intrusions, sur leur position, sur leur sens de déplacement ainsi que sur leur vitesse.

C'est dans ce sens que nous présentons, au chapitre II de la thèse, une synthèse sur les techniques générales de perception visuelle tridimensionnelle en vision artificielle. Les diverses techniques existantes sont analysées en vue de les mettre en oeuvre pour l'application envisagée. Cette étude détaillée et critique fait ressortir le besoin de concevoir un autre système de détection pour répondre aux contraintes de la surveillance des stations.

Au chapitre III nous exposons un principe de détection visuelle basé sur l'utilisation d'une caméra linéaire. La combinaison de plusieurs caméras de ce type permet d'obtenir des informations de taille et de position sur les objets pénétrant dans le champ visuel des caméras. On débouche ainsi sur le concept de barrière optique.

Le chapitre IV est dédié à la réalisation matérielle de cette barrière. Dans un premier temps, nous présentons une stratégie de suivi des intrusions à partir de paramètres extraits des images recueillies. Cette procédure apporte l'information attendue sur le sens de déplacement et sur la vitesse des objets traversant la barrière. Cette stratégie s'articule autour d'un automate à états finis.

Le chapitre V est consacré à la description des solutions matérielles et logicielles mise en oeuvre pour répondre aux contraintes temps réel de la détection. Pour valider notre approche nous concluons ce chapitre en exposant des résultats, obtenus pour différents essais d'intrusions d'une personne sur les voies.

Dans la conclusion on expose les limites de la méthode et les améliorations envisagées pour y remédier. Des perspectives d'utilisation autres que celles de la détection de chute sur les voies d'un métro sont proposées.

CHAPITRE II

- ✧ LA PERCEPTION VISUELLE D'UN ENVIRONNEMENT 3D**

I) INTRODUCTION.

Le problème principal de la compréhension générale d'une image est de reconstituer l'information tridimensionnelle contenue dans la scène observée à partir d'une image numérique acquise par une caméra qui ne fournit qu'une information bidimensionnelle. La principale difficulté provient du fait que l'image numérique 2D ne contient pas les informations sur le relief nécessaires à la modélisation spatiale des objets et de leur environnement 3D.

Pourtant, la reconstruction tridimensionnelle est indispensable pour diverses applications. Par exemple, la manipulation de pièces pour l'assemblage automatique nécessite une description tridimensionnelle des objets. La perception de l'univers 3D est également nécessaire pour piloter des robots mobiles ou des véhicules guidés [NIT 88].

Il existe des capteurs à balayage qui, grâce à la mesure du temps que met un signal réfléchi pour revenir après son émission, autorisent la perception d'une scène 3D [JAR 83] [LEW 77]. Dans ce cas, les distances du capteur aux points de la scène sont immédiatement disponibles. Dans le même esprit, Boissier [BOI 85] effectue des mesures ponctuelles par télémétrie laser pour piloter le robot mobile HILARE. Toutes ces techniques évitent le contact direct entre le capteur et les objets.

Les méthodes utilisant ces capteurs, différents des caméras, ne seront pas détaillées dans ce chapitre consacré aux méthodes de perception du monde 3D à l'aide d'une ou plusieurs caméras.

Ces méthodes peuvent être regroupées schématiquement en deux catégories:

Les méthodes monoculaires qui utilisent une seule image ou une séquence d'images acquise par une caméra unique.

Les méthodes binoculaires qui utilisent deux images simultanées de la même scène saisie par deux caméras différentes.

I.1) Les méthodes monoculaires.

a) *Scène statique*

La perception de l'environnement 3D d'une scène statique par une caméra immobile fait appel à des méthodes indirectes. Seule une analyse poussée de l'image bidimensionnelle permet de reconstituer les informations tridimensionnelles.

Connaissant la position de la source lumineuse par rapport à la caméra, il est en effet possible, en s'appuyant sur les lois de réflectance [HOR 77], ou encore sur l'analyse des textures observées, de reconstruire l'information 3D [KAN 89].

On peut aussi obtenir la position des objets dans une scène réelle par mise au point de l'image à partir de l'analyse des variations du module du gradient de l'image en fonction de la focalisation [GRO 87].

b) *Scène dynamique*

On peut également accéder à la troisième dimension d'un objet au travers d'un plan lumineux, obtenu par utilisation d'une lumière cohérente, par le biais du déplacement soit de l'objet, soit du plan lumineux [SHI 71].

D'autres techniques permettent de reconstruire le modèle 3D d'un objet se déplaçant dans le champ de la caméra sous éclairage normal. On obtient, par analyse d'une séquence d'images, des mesures locales ou globales tridimensionnelles sur les parties mobiles de la scène: profondeur relative et orientation des surfaces (structure), mouvement dans l'espace des objets [NAG 86 a].

c) *Caméra mobile*

Une approche 3D est également possible en exploitant le mouvement de la caméra qui prend des vues d'un objet fixe sous différents angles. Cette méthode ressemble étroitement à la stéréoscopie qui fait appel à la vision binoculaire [BRI 83].

I.2) Les méthodes binoculaires.

Ces méthodes utilisent deux images d'une même scène, obtenues simultanément par deux caméras. Le principe de reconstruction de l'information tridimensionnelle consiste à comparer les deux images recueillies [YAK 78].

Nous allons maintenant détailler chacune des approches évoquées ci dessus, en commençant par la stéréovision.

II) LA VISION BINOCULAIRE OU STEREOVISION.

II.1) Description du principe

Dans le but d'obtenir des informations tridimensionnelles, cette approche consiste à mesurer la distance qui sépare les caméras de points particuliers extraits d'une scène. Possédant deux vues d'un même objet, prises de deux endroits différents, il est possible de mesurer la distance de la caméra à ces points par une simple triangulation. Cette technique est connue sous le terme de "vision stéréoscopique" ou encore "vision binoculaire". Elle constitue la principale source d'information que possède l'être humain pour percevoir son environnement tridimensionnel. La figure 2.1 représente une paire de caméras observant la même scène.

L'image d'un point-objet A de la scène dans chacune des vues est formée par l'impact sur le plan image d'un rayon de lumière ayant pour origine ce point-objet et passant par le centre O de l'objectif. Inversement, le centre de l'objectif et le point-image P sur le plan image déterminent la ligne droite (D) le long de laquelle se trouve le point-objet A, dans la scène considérée.

La question à laquelle il s'agit de répondre est de déterminer, dans l'espace 3D, où est situé le point-objet A le long de cette droite (D). Cette question peut facilement être résolue à l'aide de l'image prise à partir de la seconde caméra. Si nous pouvons

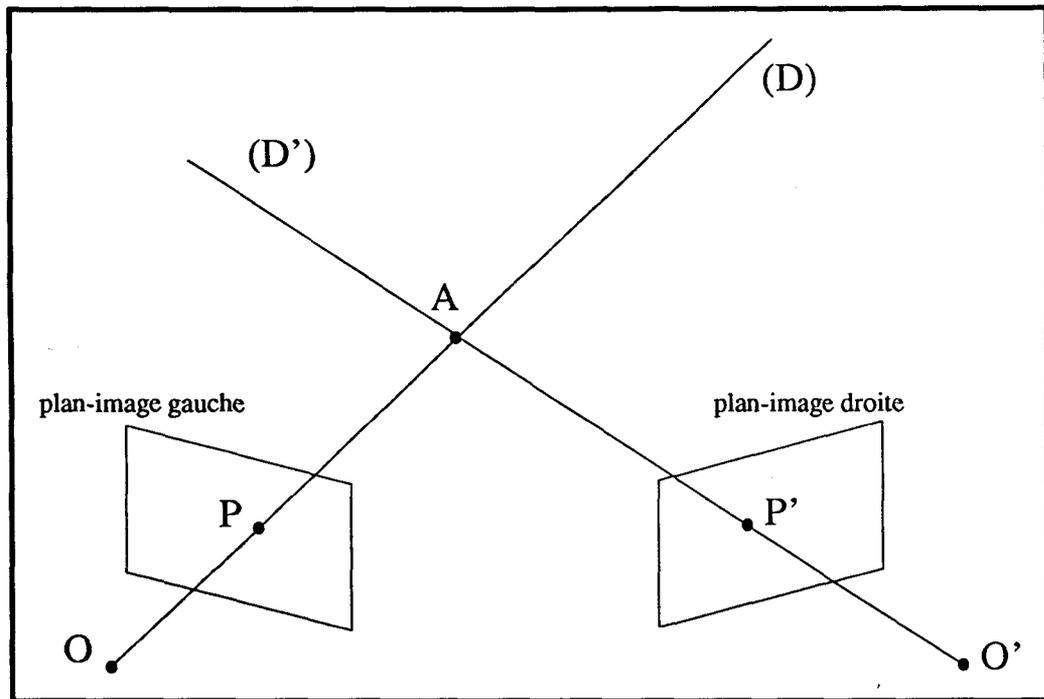


Fig. -2-1- Stéréovision.

découvrir dans la seconde image le point-image P' qui correspond au même point-objet A , alors ce point-image et le centre O' du second objectif déterminent une seconde droite (D') . La connaissance des positions relatives des deux caméras permet alors de localiser le point-objet A à l'intersection des deux droites (D) et (D') dans l'univers 3D.

II.2) Reconstitution des informations 3D.

Dans la pratique, on cherche à identifier dans les deux images, non pas de simples points dont la mise en correspondance risque d'être difficile, mais des "indices" plus représentatifs de la structure des objets observés. L'extraction de l'information de profondeur d'une paire d'images stéréoscopiques peut alors se décomposer en trois phases principales:

- la détection d'indices, facilement identifiables dans les deux images;
- la mise en correspondance de ces indices entre les deux images;

- le calcul des distances entre les caméras et les objets dont les indices caractéristiques ont été reconnus et appariés dans les deux images.

Dans le cas où l'on connaît la position relative des deux caméras, la recherche pour la mise en correspondance d'un indice donné d'une image est restreinte dans la seconde image à l'exploration d'une ligne, plutôt que de toute l'image. La figure 2.2 montre que l'objet responsable de l'indice I dans l'image gauche doit se situer le long d'une droite (D) partant du centre O de l'objectif de la caméra, et passant par l'indice I. L'image de l'objet dans l'image de droite se trouve donc sur la projection de la droite (D) sur le plan de l'image droite qui est le segment BC. Cette ligne est appelée "ligne épipolaire".

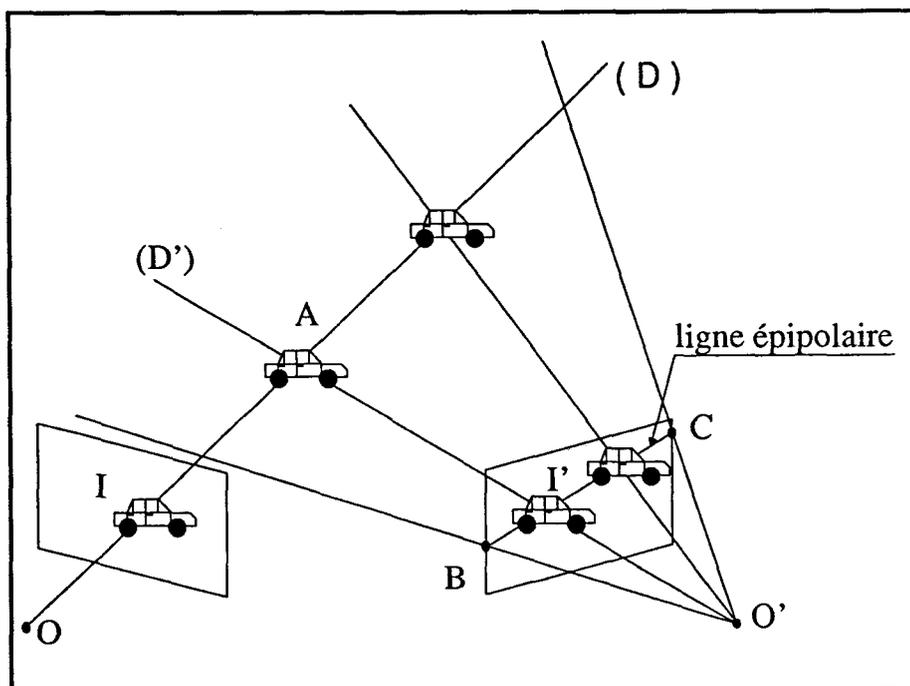


Fig. -2-2- Ambiguïté d'appariement.

Si la position des deux caméras n'est pas connue avec une grande précision, la recherche d'une mise en correspondance peut tout de même être effectuée en utilisant la contrainte épipolaire, non plus sur une ligne mais sur une bande de l'image contenant la ligne épipolaire.

D'autres contraintes peuvent être utilisées pour faciliter la mise en correspondance des indices.

La contrainte d'unicité implique qu'un indice dans la première image ne peut avoir qu'un seul correspondant dans la seconde image.

Une autre contrainte couramment utilisée est la continuité. Elle s'appuie sur l'hypothèse que les surfaces physiques sont localement continues et que la variation de leur profondeur par rapport aux caméras est généralement faible. Cette contrainte de continuité est liée à la notion de disparité. La disparité est définie comme l'écart entre les coordonnées du point de l'indice I dans le repère associé à l'image de gauche et les coordonnées de son correspondant I' dans le repère associé à l'image de droite (figure 2.2). Si A est un point de la scène 3D auquel correspondent les indices I et I' comme projections sur les images gauche et droite respectivement, avec une disparité d , alors la continuité entraîne qu'un indice voisin de I dans l'image de gauche doit trouver son correspondant dans l'image de droite avec une disparité proche de d . Lorsque les images des deux caméras sont dans le même plan (vision humaine) il existe une relation simple entre la distance et la disparité. Une disparité nulle implique que le point-objet se trouve à l'infini.

Nous présentons maintenant la détection des indices en liaison avec les techniques de mise en correspondance. En effet, les algorithmes de mise en correspondance en stéréovision diffèrent selon le type d'indice choisi et les contraintes utilisées.

II.2.1) Détection des indices et appariement.

Rappelons que l'objectif recherché est de trouver des indices d'une image qui peuvent être mis en correspondance dans l'autre image sans ambiguïté. Dans le cas idéal, la mise en correspondance doit s'établir entre tous les pixels des deux images, ce qui permet d'établir une carte de profondeur de même résolution que celle des images. Néanmoins, l'information sémantique contenue dans un pixel est trop pauvre pour résoudre le problème des appariements entre images. Par exemple, le fait que deux pixels de deux images possèdent les mêmes niveaux de gris ne présume en

rien du fait qu'ils puissent être les reflets du même point-objet de la scène. Une méthode consiste à trouver un appariement entre un ensemble de pixels voisins (ou fenêtre) de la première image et une fenêtre de la seconde image, de telle sorte que les deux fenêtres présentent le maximum de similitude. Une mesure de similitude typique est la somme des carrés des différences entre les niveaux de gris des deux fenêtres [BURT 82]. Dans certain cas, l'appariement des fenêtres est délicat, notamment lorsque la fenêtre se trouve dans une région homogène. D'autre part, lorsque les fenêtres sont sélectionnées sans tenir compte de leur contenu, la méthode risque d'être mise en défaut [CAS 84]. Il est préférable de considérer des indices qui véhiculent une information plus riche pour permettre des mises en correspondances plus sûres et plus fiables. On distingue alors deux classes d'approches, celle basée sur l'appariement de points caractéristiques et celle basée sur l'appariement de contours ou de régions extraits séparément des deux images.

a) Appariement de points caractéristiques.

Une technique pour trouver les points candidats pour la mise en correspondance consiste à utiliser l'opérateur de Moravec [MOR 77] [MOR 81]. Celui-ci est à la base d'une technique d'extraction des maximums de la variance locale des niveaux de gris dans l'image, bien adaptée au problème de la stéréovision. On calcule tout d'abord la somme des carrés des différences de niveau de gris de chaque pixel et de ses voisins dans chacune des directions horizontale, verticale et diagonales. La variance retenue pour chaque pixel est le minimum des quatre sommes ainsi obtenues. Les points remarquables sont ceux pour lesquels la variance atteint un maximum local. Ainsi les points intéressants possèdent une forte variance dans toutes les directions, ce qui favorise leur mise en correspondance avec les points caractéristiques extraits de l'autre image selon la même procédure. La mise en correspondance s'effectue sur des fenêtres centrées autour des points caractéristiques par une mesure de similitude de ces fenêtres.

Barnard et Thompson [BAR 80] utilisent une technique de relaxation pour trouver les appariements entre les voisinages des points caractéristiques. Ils utilisent l'opérateur de Moravec pour sélectionner ces points, indépendamment dans chaque image. Un ensemble d'appariements possibles est construit en associant chaque point candidat de l'image 1 avec tous les points candidats de l'image 2 situés à l'intérieur d'une zone de recherche définie par la position relative des caméras. Une

probabilité initiale de mise en correspondance correcte est donnée à chaque appariement en évaluant la ressemblance des voisinages des points candidats. La relaxation est alors utilisée pour augmenter ou diminuer les probabilités initiales pas à pas, en fonction de la cohérence des discontinuités calculées à partir des appariements d'un point de vue géométrique. Après chaque itération, les appariements possibles qui ont une probabilité d'exactitude inférieure à un certain seuil sont exclus de l'ensemble des appariements. Au terme de quelque itérations, les appariements ayant une forte probabilité d'exactitude sont retenus pour reconstruire la scène 3D. Cette technique présente parfois des difficultés d'évaluations des probabilités initiales et la convergence de la méthode n'est pas toujours garantie.

Les méthodes de mise en correspondance de points caractéristiques sont mises en défaut lorsque la scène présente des surfaces homogènes, sur lesquelles il est très difficile de mettre des points caractéristiques en évidence. De plus, la présence de structures répétitives ou d'occlusions risque de conduire à de mauvais appariements.

b) Appariement de contours.

Avec l'approche contour, on tente de rechercher des correspondances entre les contours détectés dans chacune des images stéréoscopiques. Comme un contour représente surtout des propriétés géométriques de la scène, cette approche contour est moins sensible aux caractéristiques photométriques des éléments mis en correspondance. De plus les appariements de contours sont en général plus précis et moins coûteux en temps de calcul que les appariements de points caractéristiques. En effet, les contours étant moins nombreux que les points caractéristiques dans une image, ils sont repérés plus précisément.

Les contours sont généralement localisés au niveau des maximums de l'amplitude du gradient de l'image de niveau de gris. Ces maximums du gradient correspondant aux passages par zéro des dérivées secondes, on peut également détecter les contours par les passages par zéro du Laplacien de l'image. Ces opérateurs différentiels étant très sensibles au bruit, un prétraitement tel qu'un lissage peut être utilisé pour supprimer le bruit dans l'image.

Une théorie quantitative de la vision stéréoscopique humaine fut formulée par Marr et Poggio [MAR 79]. A partir de cette théorie, Marr et Hildreth [MAR 80] proposent d'utiliser des distributions gaussiennes comme filtre de lissage ainsi que le Laplacien de la distribution gaussienne pour la détection des contours. Les deux images sont respectivement convoluées avec cet opérateur. Les points de contour sont localisés par les passages à zéro de la réponse dans les deux images. Toutefois, cet algorithme ne donne pas de bons résultats sur des scènes réelles et Grimson [GRI 81] propose une version améliorée. La modification, qui concerne surtout l'étape de mise en correspondance, conduit à de meilleurs résultats sur des scènes réelles. Cet opérateur laplacien correspond approximativement à une différence de deux gaussiennes (DOG : Difference Of Gaussian) d'écart type différent. Ces méthodes utilisent une stratégie de recherche "coarse to fine". Une recherche à basse résolution (coarse) est suivie de recherches plus fines en augmentant la résolution (fine). Ceci est réalisé en convoluant l'image avec des filtres de résolutions différentes. Les filtres de Marr et Hildreth nécessitent des noyaux de convolution larges 13x13, voir 15x15, ce qui les pénalise dans une application orientée temps réel. Crowley [CRO 84] utilise un opérateur (DOLP: Difference Of Low Pass) semblable à celui de Marr et Hildreth pour calculer le Laplacien à diverses résolutions. L'avantage majeur de son opérateur est le développement d'un moyen rapide de calcul de la transformée de DOLP. Cette approche échoue pour des scènes complexes parce que les contours détectés par les filtres de basses résolutions ne correspondent pas toujours à de vrais contours, ce qui entraîne l'introduction d'appariements incorrects. Les appariements corrects comme les appariements incorrects se propagent des bas niveaux jusqu'au niveau le plus haut [AUB 89].

Baker et Binford [BAK 81] proposent des méthodes similaires pour la mise en correspondance stéréo basée sur une technique de programmation dynamique. Ils tentent de trouver un appariement optimum pour chaque paire de lignes épipolaires dans l'image en supposant que l'ordre des contours mis en correspondance est préservé le long de ces lignes épipolaires. L'appariement optimum est obtenu en maximisant une mesure dérivée des propriétés locales des contours et en utilisant des critères d'orientation et d'intervalle entre les contours appariés. Ainsi, deux indices proches dans une image s'associent à deux indices proches dans l'autre image. Ohta et Kanade [OHT 85] utilisent la même hypothèse de contours ordonnés pour les appariements.

La plupart des mises en correspondance de caractéristiques en stéréovision appartiennent des indices de contours de bas niveau. Medioni et Nevatia [MED 85], comme Ayache et Favergon [AYA 85] proposent une technique pour apparier des indices de type contour de plus haut niveau. Les passages à zéro dans l'image convoluée avec des Laplaciens fournissent les points de contours. Les chaînes de contour sont alors approchées par des segments de droite. Chaque segment est représenté par un certain nombre d'attributs: des attributs géométriques tels que les coordonnées du milieu, la longueur, l'orientation et d'autres basés sur l'intensité, tels que le contraste le long du segment, l'intensité moyenne du gradient de l'image le long du segment. Les segments à apparier sont recherchés dans une fenêtre dont deux des cotés sont parallèles à l'orientation du segment et les deux autres horizontaux. En utilisant une telle fenêtre, on considère toutes les lignes épipolaires qui coupent le segment à apparier et une disparité maximum est autorisée pour chaque appariement. Grâce à une technique itérative, la qualité de chaque appariement possible est évaluée en fonction de la disparité. L'appariement retenu pour chaque segment est celui pour lequel il y a le plus de ressemblance avec les disparités du voisinage.

L'avantage de cette approche est qu'elle utilise des primitives géométriques très riches en information, donc faciles à mettre en correspondance. D'autre part la grande stabilité des segments d'une image à l'autre garantit de bons résultats.

Le principal inconvénient de cette approche est que la mise en correspondance de contours ou de segments engendre des cartes de profondeur définies uniquement sur quelques points de l'image. Cette perte de résolution au niveau de l'information 3D nécessite une reconstruction des surfaces par interpolation.

c) Appariement de régions.

Un dernier type d'indice qui peut être mis en correspondance est constitué par les régions, définies comme des zones connexes fermées, formées d'un ensemble de points homogènes [COQ 87]. La mise en correspondance de régions est très séduisante car les régions portent une information intrinsèque, comme le niveau de luminance ou la texture, plus riche que les contours. Le principal inconvénient de

cette méthode est justement d'extraire correctement ces régions des images [MON 87].

II.2.2) Calcul des distances.

Considérons l'illustration de la figure 2.3 décrivant une configuration géométrique typique en stéréovision. On associe à chaque caméra un repère local. Dans chaque repère, l'axe des z coïncide avec l'axe optique de la caméra associée et les deux plans images sont parallèles aux plans (Ox, Oy) et $(O'x', O'y')$. Ils coupent les axes Oz et $O'z'$ à une distance unité de O et O' . Ceci revient à prendre des distances focales unités: $f = f' = 1$.

Le repère associé à la caméra de droite est obtenu en effectuant une rotation d'un angle Φ autour de l'axe Oy et une translation de vecteur $\vec{OO'}$ de composante b_x suivant Ox et b_z suivant Oz sur le repère associé à la caméra de gauche. Cette transformation est représentée par la matrice A telle que:

$$A = \begin{pmatrix} \cos \Phi & 0 & \sin \Phi & b_x \\ 0 & 1 & 0 & 0 \\ -\sin \Phi & 0 & \cos \Phi & b_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

Soit P un point de la scène observée par les deux caméras et (X, Y, Z) ses coordonnées par rapport au repère associé à la caméra de gauche. Exprimons les coordonnées du point P dans le repère associé à la caméra de droite (X', Y', Z') :

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = A \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2)$$

L'image p du point réel P est la projection perspective de centre O de P sur le plan image gauche. L'image p' est la projection perspective de centre O' du point P sur le plan image droite. A partir des expressions données par l'équation (2) nous

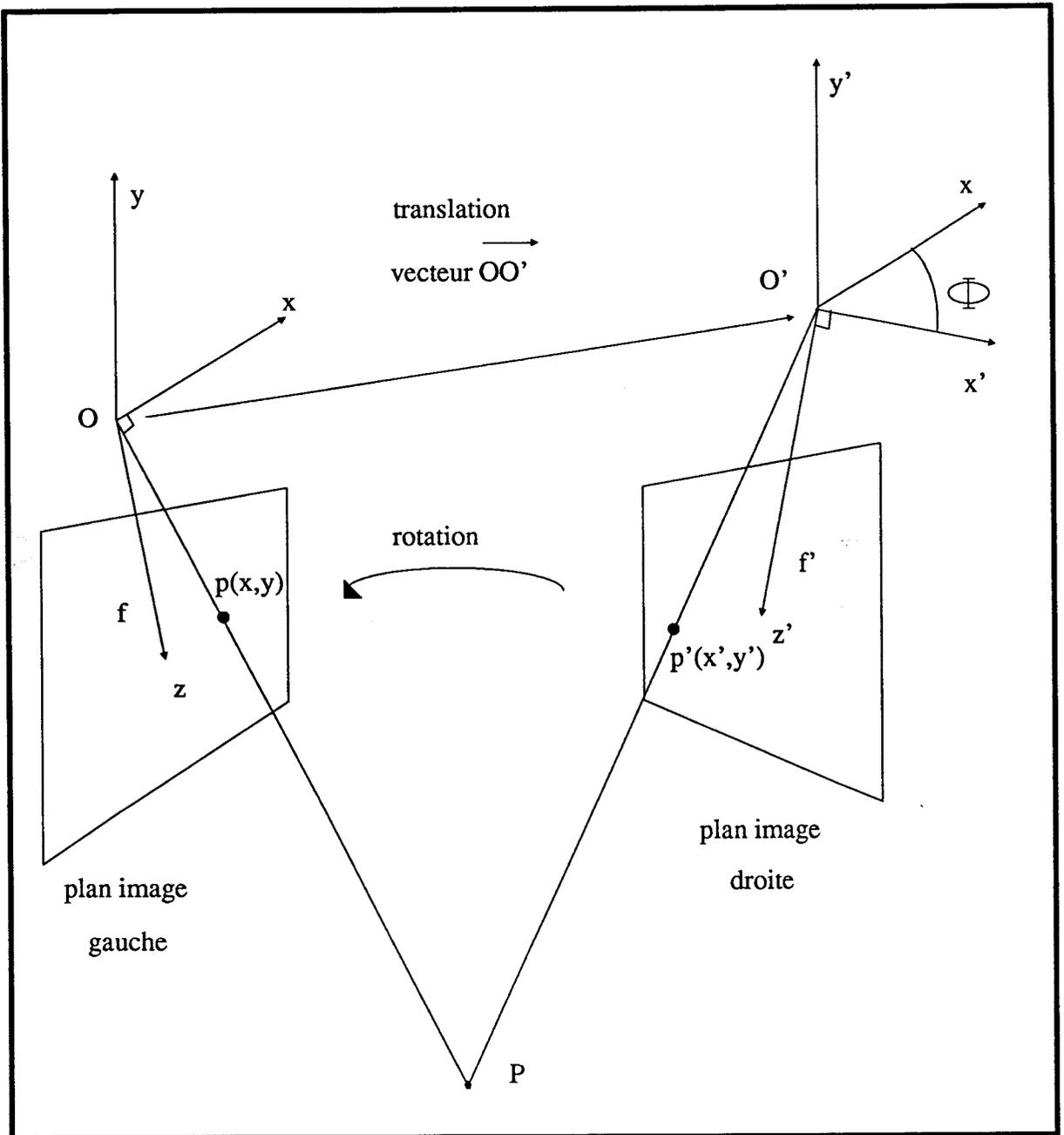


Fig. -2-3- Géométrie de stéréovision.

pouvons exprimer les coordonnées (X', Y', Z') du point P dans le repère de la caméra de droite par exemple, en fonction des coordonnées x' et y' du point p' , des

coordonnées x et y du point p , et des paramètres b_x , b_z et Φ définissant la géométrie du système:

$$X' = \frac{\cos\Phi (x' b_x - x x' b_z) - \sin\Phi (x' b_z + x x' b_x)}{\cos\Phi (x' - x) - \sin\Phi (1 + x x')}$$

$$Y' = y' \frac{\cos\Phi (b_x - x b_z) - \sin\Phi (b_z + x b_x)}{\cos\Phi (x' - x) - \sin\Phi (1 + x x')}$$

$$Z' = \frac{\cos\Phi (b_x - x b_z) - \sin\Phi (b_z + x b_x)}{\cos\Phi (x' - x) - \sin\Phi (1 + x x')}$$

Les calculs permettant d'obtenir ces résultats sont décrits en annexe (ANNEXE 1).

Les coordonnées (X', Y', Z') déterminent la position du point P dans l'espace tridimensionnel par rapport au repère de la caméra de droite.

II.3) Conclusion.

La stéréovision semble être une source d'informations très séduisante pour la perception de l'environnement. Il est admis que l'étape clef de la stéréovision est le processus de mise en correspondance des indices détectés dans les deux images. Pour résoudre ce problème, plusieurs méthodes ont été étudiées, telles que la méthode de programmation dynamique, la méthode de relaxation, et la méthode multi-résolution (coarse to fine). Cependant les algorithmes restent lents et complexes. Certains auteurs ont utilisé une séquence d'images binoculaires pour confirmer les appariements dans les paires d'images successives [NEV 76]. D'autres utilisent une troisième caméra pour résoudre le problème de mise en correspondance [GER 86],[GUR 86],[ITO 86],[OHT 86] et [YAC 86]. L'ajout d'une caméra permet de vérifier sur la troisième image si l'appariement effectué entre indices des deux premières images est correct.

Nous allons, dans le paragraphe suivant, détailler les systèmes de vision monoculaire capables de fournir une information tridimensionnelle.

III) VISION MONOCULAIRE.

III.1) Reconstitution de forme à partir d'une image

Une image bidimensionnelle ne fournit que des informations incomplètes sur la forme des objets constituant la scène observée. Pour reconstituer la forme 3D des objets observés, des hypothèses reliant les caractéristiques de l'image et celles de la scène sont nécessaires. Les méthodes discutées dans ce paragraphe sont souvent appelées en anglais " Shape from ".

a) Par analyse de la réflectance.

La détermination de l'orientation d'éléments de surface à partir des nuances de niveaux de gris (shape from shading) fut étudiée en premier par Horn [HOR 75]. La quantité de lumière réfléchiée par un élément de surface dépend de sa microstructure, de ses propriétés optiques mais aussi de la distribution et de la polarisation de la lumière incidente. Mais pour certaines surfaces, la fraction de la lumière incidente réfléchiée dans une direction particulière dépend seulement de l'orientation de la surface. Les caractéristiques de la réflectance de telles surfaces peuvent être représentées comme une fonction $f(i, g, e)$ des angles incident i , de phase g et émergent e , comme définis sur la figure 2-4.

La fonction de réflectance $f(i, g, e) = \rho \cos i$ traduit le modèle le plus simple de réflexion d'une surface, où ρ est une constante (albédo de surface) qui dépend de la nature de la surface. Cette fonction de réflexion correspond à un modèle phénoménologique d'une surface parfaitement diffuse (surface lambertienne), la lumière réfléchiée étant la même dans toutes les directions.

Sous projection orthogonale, la direction de vue et l'angle de phase sont constants pour tout élément de surface. Ainsi pour une source de lumière fixe et une caméra fixe, la fonction de réflectance dépend seulement du vecteur normal à la surface ($f(i, g, e) = R(p, q)$). L'orientation dans l'espace tridimensionnel du vecteur \vec{n} de composantes p et q , normal à la surface, peut être représentée à l'aide

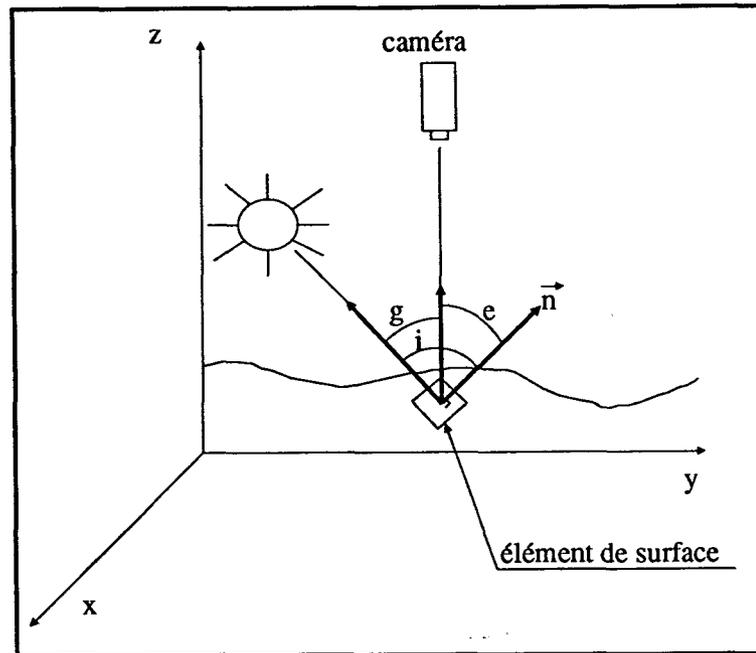


Fig. -2-4- Analyse de la réflectance.

des composantes du gradient de la fonction de réflectance. En supposant que la fonction $z = f(x, y)$ représente la surface, le gradient (p, q) est défini comme:

$$p = \frac{\delta f}{\delta x} \quad q = \frac{\delta f}{\delta y}$$

Les expressions pour $\cos i$, $\cos e$ et $\cos g$ se calculent facilement à partir du vecteur $(p, q, -1)$ normal à la surface, du vecteur $(p_s, q_s, -1)$ qui représente l'orientation de la source de lumière et du vecteur $(0, 0, -1)$ qui pointe dans la direction de l'observateur. L'intensité I au point (x, y) de l'image, uniquement fonction de la normale à la surface dans l'hypothèse de surface lambertienne, s'identifie à la fonction de réflectance :

$$I(x, y) = R(p, q) = \frac{\rho(1 + pp_s + qq_s)}{\sqrt{(1 + p^2 + q^2)}\sqrt{(1 + p_s^2 + q_s^2)}}$$

La fonction de réflectance $R(p, q)$ permet de déterminer une carte de réflectance [HOR 77] qui représente les courbes d'iso-intensité de l'image en fonction de p et q , pour des valeurs connues de p_s, q_s et de ρ . Cependant, les contraintes géométriques de l'image ainsi que les contraintes sur la réflectivité de la surface ne

sont pas suffisantes pour déterminer l'orientation de l'élément de surface au point considéré.

L'introduction de contraintes supplémentaires d'unicité, de telle sorte qu'à chaque point de l'image soit associé au plus une orientation de surface, et de contraintes de continuité sur une même surface, qui précisent que l'orientation varie peu sauf au voisinage des contours de l'objet, permet alors d'obtenir les paramètres p et q d'orientation des surfaces observées.

En éclairant une scène avec la même source de lumière, mais sous deux directions différentes, on obtient deux cartes de réflectance [IKE 81]. A partir de ces cartes on détermine les valeurs de p et q en recherchant les intersections des courbes d'iso-intensité.

b) Par analyse de la texture.

L'être humain peut aisément reconnaître la forme des objets grâce aux changements de texture dus à la distance et à l'orientation. Il est clair que le changement de texture fournit une importante source d'information sur l'orientation locale des surfaces visibles.

En général, les distorsions introduites par la projection sur le plan image peuvent être considérées comme provenant des effets suivants:

- l'effet de distance: les objets dans le champ de la caméra, apparaissent sur le plan image plus grand, s'ils sont près;

- l'effet de position: la distorsion d'un motif dépend aussi de l'angle entre l'axe optique et la ligne de vue qui relie le motif au centre optique de l'objectif;

- l'effet d'écrasement: la distorsion d'un motif dépend de l'angle entre la normale à la surface qui le supporte et la ligne de vue;

Pour reconstruire la forme de l'objet dans la scène, ces effets de distorsion peuvent être différenciés selon les propriétés géométriques de l'élément de texture, appelé textel, sur lequel agit la distorsion.

Par exemple, l'image d'un textel circulaire est en général une ellipse. L'orientation de son grand axe et le rapport du petit diamètre au grand diamètre peuvent être utilisés pour calculer respectivement la rotation et l'inclinaison de la surface par rapport à l'axe optique de la caméra [STE 79]. Le problème majeur de cette approche est que l'on doit d'abord reconnaître la forme du textel qui n'est pas toujours connue *a priori*.

Une nouvelle approche, qui ne requiert pas de connaissance *a priori* sur la texture, se base sur des hypothèses statistiques de la distribution des textels. Par exemple, si la vraie texture est distribuée de façon isotropique, la forme 3D peut être calculée à partir de l'observation de son image dans une direction donnée. Cette approche fut étudiée par Witkin [WIT 81], puis améliorée par Davis [DAV 83].

Une autre hypothèse utilisée concerne l'homogénéité de la texture. On suppose que la nature de la surface est homogène et que toute variation observée dans l'image est due au changement soit de forme, soit de point de vue. La texture apparaît plus lâche sur une surface éloignée de l'observateur et plus serrée près de celui-ci [BAJ 76]. Avec ces hypothèses, les propriétés des textures observées, comme la distorsion et le changement d'aspect, ne sont pas en nombre suffisant pour déterminer les orientations des surfaces observées de façon unique. Des hypothèses d'unicité et de continuité sont là encore nécessaires pour propager les contraintes et faciliter la recherche d'une solution globalement correcte.

Notons finalement que les méthodes "shape from" fournissent des informations sur la forme et l'orientation des objets observés mais ne donnent aucun renseignement sur la position tridimensionnelle des objets.

III.2) Mesure de profondeur par focalisation automatique.

Quand l'oeil accommode pour voir avec un maximum de netteté les objets situés à une distance déterminée, les objets plus proches ou plus éloignés ne forment plus une image nette sur la rétine. La technique de focalisation automatique s'inspire de cette constatation et remplace la notion subjective de "netteté" par une notion plus concrète pour le système de vision, basée sur la mesure de l'intensité du gradient: plus l'image d'un objet est nette, plus l'intensité du gradient de cette image est élevée. La première étape consiste donc en une segmentation de l'image original par l'emploi d'un opérateur gradient ou laplacien. Les contours sont ensuite analysés dans une deuxième étape pour déterminer la netteté de l'image des objets dans la scène [BER 86].

On considère la relation optique: $\frac{1}{f} = \frac{1}{r} + \frac{1}{s}$

où r et s sont les distances de l'objet et de son image respectivement au centre de la lentille et f la distance focale (Fig. 2.5).

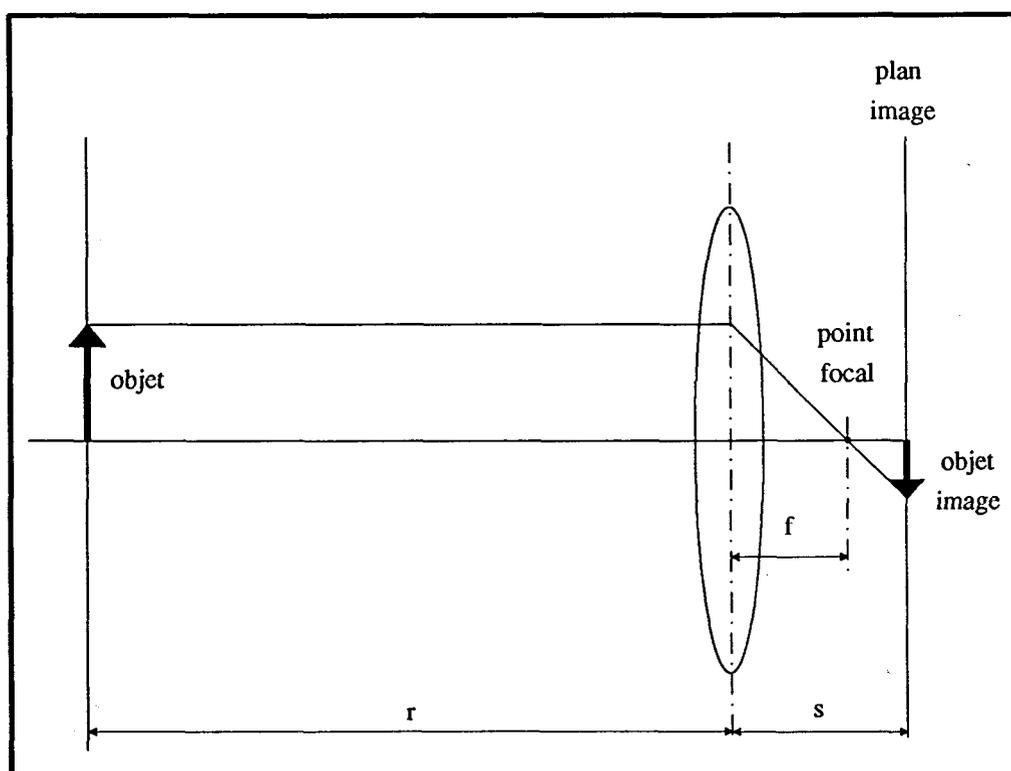


Fig.-2-5-Construction graphique de l'image d'un objet

En focalisation automatique, on recherche la valeur de s , en réglant la mise au point de l'objectif, qui amène un objet au maximum de netteté dans le plan image, c'est à dire la valeur qui maximise le module du gradient calculé sur toute l'image ou sur une fraction de celle-ci. Connaissant la distance focale f et la distance s , il est possible de calculer la distance r de l'objet considéré au système optique [JAR 76] [TUC 81].

Cette méthode est très intéressante car il n'y a pas de problème de mise en correspondance à résoudre, mais elle manque de précision pour les objets proches de la caméra et pour les objets se situant dans la zone de profondeur de champ du système caméra-objectif.

III.3) Vision active avec lumière cohérente.

Une approche directe pour trouver la profondeur est la méthode de vision active ou triangulation optique, développée par Shirai, qui utilise une source de lumière parfaitement connue [SHI 71]. La méthode de triangulation est la même qu'en stéréovision, mais une des caméras est remplacée par une source de lumière cohérente, afin d'éviter les problèmes de mise en correspondance.

Cette méthode consiste à projeter sur un objet dont on veut déterminer la forme 3D, un ou plusieurs plans lumineux obtenus à partir d'une source de lumière monochromatique constituée d'un laser à faible intensité. Ce dispositif d'éclairage présente l'intérêt d'émettre dans une direction privilégiée un faisceau de lumière visible ayant une divergence très faible. Pour obtenir un plan lumineux, il suffit de dévier le rayon lumineux au moyen de deux miroirs situés dans deux plans orthogonaux (Fig. 2-6).

Une caméra munie d'un filtre passe-bande, centré sur la fréquence de la lumière cohérente, permet d'obtenir une image très contrastée de l'intersection du plan avec l'objet.

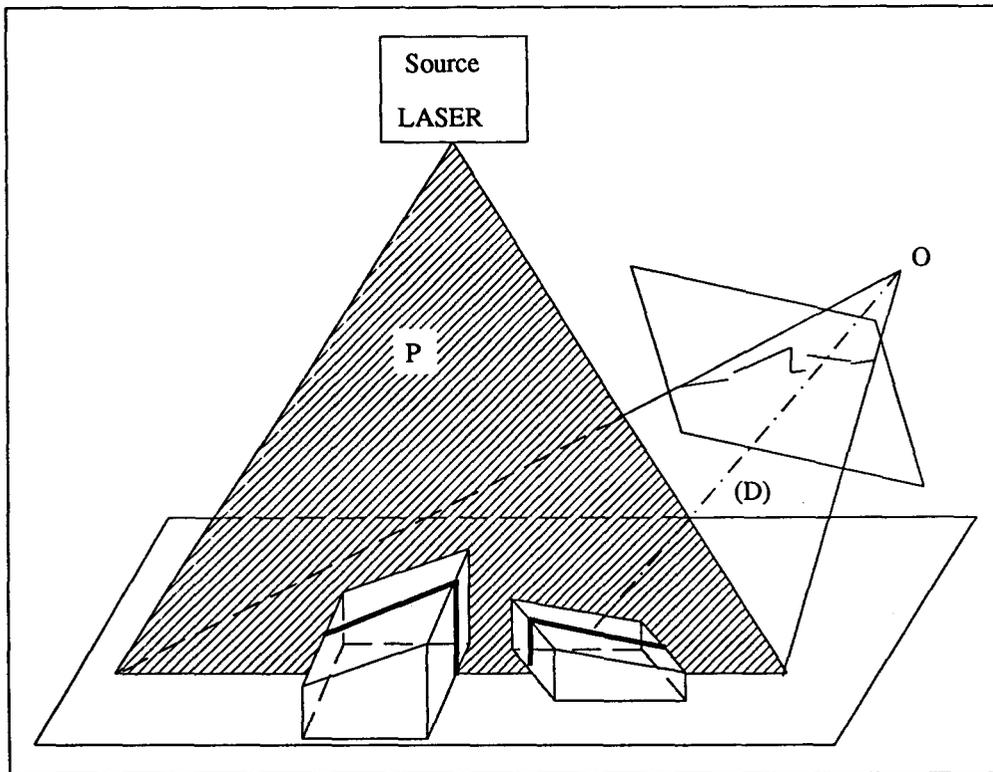


Fig. -2-6- Vision 3D avec lumière structurée.

Dans cette procédure, une caméra matricielle enregistre les positions des points éclairés par le plan lumineux (Fig. 2.6). La position de chaque point éclairé, qui apparaît en surbrillance dans l'image, est déterminée par l'intersection du plan lumineux (P), qui est connu avec précision, avec la droite (D) joignant ce point éclairé de la scène au point correspondant dans le plan image. Une fois l'identification du point dans le plan image accomplie, la distance du point correspondant de la scène à la caméra peut être calculée par simple triangulation. L'information 3D de la scène entière est acquise en déplaçant soit le plan lumineux, soit la caméra, et en enregistrant la position des points éclairés pour chaque configuration.

Cette technique donne de bons résultats dans des configurations bien contrôlées et à faible distance de la caméra. Elle ne peut être envisagée pour l'analyse de scène 3D réelles du monde extérieur.

III.4) Analyse d'une séquence d'images.

Une séquence d'images fournit une très grande quantité d'informations sur une scène dans laquelle se déplacent des objets. L'interprétation du mouvement de ces objets, déterminé à partir de l'analyse d'une séquence d'images monoculaires, donne des informations tridimensionnelles sur la scène. Traditionnellement, il existe deux approches principales pour l'analyse du mouvement.

La première est basée sur l'extraction d'un ensemble de caractéristiques bidimensionnelles dans l'image correspondant aux objets tridimensionnels de la scène. Elle consiste à mettre en correspondance ces caractéristiques entre les images successives pour ensuite donner une description 3D de la scène à partir de ces mises en correspondance. La tâche d'analyse peut être décomposée en trois étapes qui permettent de reconstituer la structure 3D des éléments en mouvement:

- extraction des primitives,
- mise en correspondance des primitives et
- inférence des informations 3D à partir du mouvement.

La deuxième consiste à calculer le flot optique, ou le champ des déplacements de l'image, et à déterminer le champ des déplacements tridimensionnels à partir de celui-ci, en supposant qu'il est la projection sur le plan image du champ des déplacements réels dans l'espace 3D.

III.4.1) Méthode des primitives caractéristiques.

a) Extraction et mise en correspondance de primitives caractéristiques.

L'extraction de primitives caractéristiques d'une image constitue une étape importante sur laquelle s'appuie la phase de mise en correspondance. La tâche pour établir et maintenir les correspondances entre primitives tout le long d'une séquence d'images n'est pas évidente. Le développement de techniques robustes pour résoudre le problème de mise en correspondance est un sujet de recherche commun

à d'autres domaines du traitement d'images, tel que la stéréoscopie que nous avons évoquée précédemment. En effet, quelques unes des techniques utilisées dans le domaine de la stéréovision peuvent être aussi bien appliquées à l'analyse d'une séquence d'images monoculaires. Nous présentons une brève description de quelques unes des approches classiques qui ont été plus spécifiquement développées pour l'analyse de séquence d'images.

Sethi et Jain [SET 87] décrivent une méthode pour mettre en correspondance des primitives entre deux images consécutives et les maintenir tant que celles-ci sont présentes dans l'image. Ils présentent des algorithmes basés sur la continuité du mouvement de ces primitives d'une image à la suivante. Ils obtiennent un ensemble optimum de trajectoires pour ces primitives en respectant des contraintes de direction et d'amplitude du mouvement. Une méthode est proposée en complément pour tenir compte d'éventuelles d'occlusions. Cette méthode formule des hypothèses d'occlusions si le nombre de primitives détectées dans une image est inférieur au nombre de primitives détectées dans une ou plusieurs images précédentes ou suivantes. La position de la primitive manquante est interpolée en utilisant les deux images précédentes; un test est effectué dans les deux images suivantes pour vérifier l'existence de l'occlusion. Des expérimentations illustrent, avec des primitives caractéristiques extraites manuellement, que la méthode est capable de donner des résultats satisfaisants quand le nombre d'occlusions reste limité.

Les méthodes de relaxation sont très utilisées pour éviter le problème des faux appariements. Ranade et Rosenfeld [RAN 80], utilisent des méthodes de relaxation pour chercher la correspondance entre des points caractéristiques sur deux images successives. La méthode de relaxation de Barnard et Thompson [BAR 80], détaillée dans le paragraphe consacré à la stéréovision, peut être appliquée à l'appariement de primitives caractéristiques de deux images monoculaires. La contrainte de disparité est ici basée sur la régularité du mouvement des points caractéristiques extraits de l'image, et choisis comme paramètres.

Pragger et Arbib [PRA 83] décrivent une technique similaire à celle de Barnard et Thompson, mais incluent une contrainte temporelle sur le déplacement des points caractéristiques. Beaucoup d'autres approches pour appairer les points caractéristiques entre les images d'une séquence peuvent être trouvées dans la littérature [JEN 86], [DRE 82].

Une autre approche consiste à limiter les mises en correspondance aux points ou régions affectés par le mouvement. La différence temporelle d'images fait apparaître des points de valeur non-nulle qui correspondent aux déplacements d'objets entre les images successives [JAI 81]. L'étape suivante consiste à agglomérer ces points et filtrer les détections parasites pour aboutir à une interprétation en termes de régions ou objets mobiles sur un fond fixe. La mise en correspondance de ces régions entre les images successives s'effectue par des techniques de corrélation. Stelmaszyk [STE 85] utilise les contours en mouvement à la place des régions mobiles comme indices de déplacements. Mais, lorsque le fond n'est pas uniforme, la méthode fait apparaître des contours sous la position de l'objet en mouvement dans les images précédentes. Vieren [VIE 88] lève cette difficulté en utilisant trois images successives de la scène à analyser. La mise en correspondance de ces contours en mouvement peut s'avérer plus facile que celle de points isolés si la détection des contours est efficace et que l'échantillonnage temporel est rapide.

Une fois la mise en correspondance des primitives réalisée, l'estimation du mouvement et de la structure 3D des objets dans la scène, peut être envisagée.

b) Inférence 3D par estimation du mouvement.

Dans ce paragraphe, nous discutons de l'approche structurelle, en anglais "feature based ", pour l'estimation de la structure 3D à partir d'une séquence d'images monoculaires acquise au moyen d'une seule caméra. On suppose que les primitives caractéristiques extraites d'images consécutives ont été correctement appariées. Les contraintes formulées sont basées sur le mouvement de corps rigides tels que la distance 3D entre deux primitives d'un tel corps reste constante durant le mouvement de l'objet. Ces hypothèses permettent d'aboutir à des systèmes d'équations non linéaires que l'on résout en utilisant les déplacements observés des primitives 2D dans la séquence d'images analysée.

Un modèle d'image, dit orthographique, est utilisé par Ullman [ULL 79] pour estimer la structure tridimensionnelle et le mouvement d'un objet rigide. La position et le mouvement de quatre points non-coplanaires dans l'espace sont retrouvés à partir de trois projections orthogonales distinctes de ces quatre points. Ullman aboutit à trois équations linéairement indépendantes qui possèdent deux solutions égales en

amplitude mais de signes opposés. Il choisit une des deux solutions pour déterminer les distances aux trois plans, et une configuration 3D unique peut alors être calculée.

Toutefois, le modèle de projection orthographique n'est pas bien adapté à certaines situations du monde réel qui demandent l'utilisation de la projection perspective. L'utilisation de la transformation perspective augmente substantiellement la complexité du problème. Roach et Aggarwal [ROA 79] [ROA 80] furent parmi les premiers à s'intéresser au mouvement et à la structure des objets d'une scène via la transformation perspective d'image. Ils montrent que l'on a besoin de cinq points en correspondance dans deux images pour retrouver les paramètres de translation et de rotation. Chaque point produisant deux équations par image, on obtient ainsi vingt équations non linéaires. Pour une simulation avec des images très peu bruitées, la méthode converge très lentement vers la bonne réponse. Si on introduit du bruit sur la position des points dans les plans images, la solution ne peut être trouvée que si l'on prend deux images avec 12 à 15 points chacune ou trois images avec sept à huit points. La résolution de ce système d'équations permet de reconstituer la surface 3D qui passe par les points considérés.

Tsai et Huang [TSA 81] proposent une méthode pour trouver le mouvement d'un élément de surface à partir de la projection 2D de la scène. Ils définissent un ensemble de huit paramètres et supposent un mouvement de corps rigide. Ces paramètres peuvent être déterminés de façon unique à partir de deux images successives en résolvant un polynôme du sixième degré.

Peu après, Tsai et Huang [TSA 84] étudièrent le problème des éléments de surface courbe en mouvement. Deux principaux résultats furent établis concernant l'existence et l'unicité des solutions. Une matrice E est spécifiée, telle que $E = T.R$, où T est une translation et R est une rotation. À partir de huit points mis en correspondance, de positions quelconques, la matrice E est déterminée de façon unique en résolvant huit équations linéaires. Cette approche est très sensible au bruit qui affecte la précision sur le calcul des coordonnées.

Les approches ci-dessus utilisent des points et leurs projections sur les plans image comme point de départ pour la formalisation du problème. Une autre approche consiste à utiliser des lignes 3D et leurs projections. Si on utilise des lignes ou des droites comme primitives caractéristiques, deux vues ne sont pas suffisantes pour

reconstruire la scène 3D et un minimum de trois vues est nécessaire. Ceci est dû au fait que les droites possèdent un degré de liberté supplémentaire par rapport aux points.

Yen et Huang [YEN 83] proposent une méthode itérative basée sur une projection sphérique et sur l'observation de sept lignes en correspondance dans trois vues pour le cas d'un mouvement quelconque.

Liu et Huang [LIU 86] ont utilisé les correspondances de droites pour formuler une méthode analogue à la précédente. Ils décomposent le mouvement de corps rigide en une rotation autour d'un axe passant par l'origine, suivie d'une translation. Dans le cas d'une rotation pure, deux droites en correspondance sur deux images sont suffisantes pour déterminer la matrice de rotation. Les équations non linéaires résultantes sont résolues de manière itérative. Dans le cas d'une translation pure, cinq droites en correspondance sur trois images produisent un système d'équations linéaires que l'on résout pour trouver la translation. Dans le cas général, Liu et Huang utilisent six droites en correspondance dans trois images pour déterminer la matrice de rotation puis la matrice de translation. Des simulations de l'algorithme sur des images synthétiques volontairement bruitées montrent que l'approche est très sensible au bruit. De plus, l'estimation du vecteur translation est très sensible aux erreurs d'estimation des paramètres de rotation.

Mitiche, Seida et Aggarwal [MIT 86] utilisent le principe de l'invariance angulaire entre deux droites d'un même corps rigide en mouvement. Après avoir déterminé l'orientation des droites, ils calculent les composantes de la rotation et finissent par établir les caractéristiques de la translation. L'observation de quatre droites dans trois vues suffit au calcul des paramètres de mouvement et de la structure 3D.

Les segments de droites sont des primitives qui possèdent l'avantage d'être plus aisées à extraire et à mettre en correspondance que les points caractéristiques. Il est possible de combiner les points et les segments de droites. Dans ce cas, quatre points et une droite sont suffisants pour calculer la structure 3D de la scène aussi bien que les déplacements entre ces images [AGG 87]. L'utilisation de points ou de droites, ou une combinaison de points et de droites, pour le calcul de la structure tridimensionnelle à partir du mouvement conduit souvent à des équations non

linéaires. D'autre part, le nombre minimum de points ou de droites à prendre en compte est très sensible au bruit présent dans les images de la séquence.

c) Conclusion.

Ces méthodes basées sur l'analyse des primitives caractéristiques exigent une mise en correspondance entre les primitives extraites d'une image avec celles extraites de l'image suivante de la séquence. Plusieurs méthodes ont été décrites pour extraire ces primitives et les mettre en correspondance. La tâche est difficile et seulement des solutions partielles, satisfaisantes pour des situations simples, ont été proposées. En général le processus se complique lorsqu'il y a occlusion des objets, ce qui entraîne la disparition temporaire de certaines primitives et la génération de faux appariements.

III.4.2) Méthode du flot optique.

Nous présentons les approches dans lesquelles les changements instantanés des valeurs de l'intensité de l'image sont analysés pour créer une carte de vitesse appelée flot image ou flot optique.

Le flot optique se calcule à partir des dérivées premières spatiales et temporelles de l'image. Il est possible, par l'addition de contraintes, d'estimer les paramètres de mouvement et la structure tridimensionnelle de la scène.

a) Calcul du flot optique.

On note $g(x, y, t)$ l'intensité de l'image au point (x, y) du plan image à l'instant t . On admet l'hypothèse que la brillance d'un point de l'espace 3D et, par là-même, la luminance de sa projection dans le plan image sont invariantes dans le temps, ou du moins entre deux instants t et $t + dt$ considérés. Si le point considéré se déplace de (dx, dy) de l'instant t à l'instant $t + dt$, alors nous pouvons écrire que:

$$g(x+dx, y+dy, t+dt) = g(x, y, t) \quad (1)$$

Si l'on suppose que la fonction intensité possède des propriétés de continuités et de dérivabilités jusqu'à l'ordre 2, on peut alors considérer le développement de Taylor de cette fonction intensité au point (x, y) à l'instant t .

$$g(x+dx, y+dy, t+dt) = g(x, y, t) + dx \frac{\delta g}{\delta x} + dy \frac{\delta g}{\delta y} + dt \frac{\delta g}{\delta t} + \text{termes d'ordre supérieur} \quad (2)$$

Si on néglige les termes d'ordre supérieur ou égal à 2 dans l'expression (2), on obtient, après quelques calculs simples [HOR 81]:

$$\frac{\delta g}{\delta x} \frac{dx}{dt} + \frac{\delta g}{\delta y} \frac{dy}{dt} + \frac{\delta g}{\delta t} = 0$$

ou, de manière plus concise:

$$g_x u + g_y v + g_t = 0 \quad (3)$$

Dans cette équation, les dérivées partielles g_x , g_y et g_t sont estimées à partir de la fonction intensité d'image en x , y et t respectivement, $u = \frac{dx}{dt}$ et $v = \frac{dy}{dt}$ sont les composantes du vecteur vitesse au point (x, y) . L'estimation des vitesses entre l'image $I(t)$ et l'image $I(t+dt)$ conduit à résoudre l'équation (3) pour chaque point (x, y) de l'image $I(t)$.

L'hypothèse initiale d'invariance de la luminance nous permet d'aboutir à la relation (3). Cette hypothèse implique que les variations temporelles des intensités dans l'image ne sont dues qu'au mouvement. Cela suppose que l'illumination de la scène reste invariante. Cependant, la relation (3) est insuffisante pour permettre d'estimer les deux composantes u et v du vecteur vitesse au point (x, y) .

L'introduction d'une contrainte de lissage a été originellement formalisée dans [HOR 81]. La scène observée étant principalement constituée d'objets avec des surfaces lisses, Horn suppose que la variation du champ des vitesses est faible d'un point à un autre. Cette variation se traduit par le carré de la norme du gradient du

vecteur vitesse, vecteur à quatre composantes $(\frac{\delta u}{\delta x}, \frac{\delta u}{\delta y}, \frac{\delta v}{\delta x}, \frac{\delta v}{\delta y})$ noté aussi (u_x, u_y, v_x, v_y) .

L'addition des contraintes d'invariance de luminance et de lissage mène à une fonction erreur totale, qu'il convient de minimiser pour déterminer les composantes u et v du vecteur vitesse. Horn [HOR 81] aboutit à un système d'équation aux dérivées partielles qu'il résout par une méthode itérative. Yachida [YAC 83] développe cette méthode itérative et ajoute une contrainte de lissage temporelle complémentaire. Horn montre que cette méthode itérative converge quand le flot optique est statique, c'est à dire lorsque le vecteur vitesse ne change pas avec le temps, comme c'est le cas lorsqu'une sphère tourne autour d'un axe fixe [HOR 81].

Quand cette condition n'est pas respectée, par exemple dans le cas d'une translation d'un objet fixe sur un fond stationnaire, il existe des frontières où la propriété de continuité est en défaut [ADI 85]. Nagel [NAG 86 b] intègre une matrice dans la fonction erreur à minimiser pour tenir compte de la discontinuité du champ des vitesses. Une autre solution consiste à détecter ces frontières et à ne calculer le flot optique que dans les régions homogènes, où le champ des vitesses est continu [SCH 86].

Une présentation plus complète de la plupart des méthodes avec une comparaison de leurs aspects communs et de leurs différences peut être trouvée dans [NAG 87].

b) Reconstruction des informations 3D.

Subbarao [SUB 87] et Waxman [WAX 84] donnent une explication détaillée sur la formulation des équations qui relient le flot optique à la structure 3D et aux paramètres de mouvement. Ils supposent la caméra en mouvement et des contraintes de continuité du flot optique pour l'estimation des informations tridimensionnelles et du mouvement. Ils aboutissent à un système d'équations non linéaires.

William [WIL 80] et Lawton [LAW 83] considèrent le calcul pour le cas où la caméra est uniquement animée d'un mouvement de translation.

Prazdny [PRA 81] décompose le mouvement de la caméra en une rotation et une translation. Il aboutit à une fonction erreur qu'il minimise pour estimer le mouvement.

Chou et Kanantani [CHO 87] proposent une technique itérative pour affiner l'estimation initiale du mouvement des objets.

c) Conclusion.

Dans cette section nous avons présenté l'utilisation du flot optique pour estimer des paramètres de mouvement à partir d'une séquence d'images monoculaires. Différentes techniques pour calculer le flot optique sont discutées ainsi que son exploitation pour reconstituer le mouvement 3D. La méthode du flot optique ne nécessite pas en général de mise en correspondance comme pour les méthodes des primitives caractéristiques. Mais la méthode du flot optique n'est valable que pour de faibles déplacements. Une façon de se placer au mieux dans ce contexte est de procéder à un échantillonnage temporel de la scène suffisamment fin. Le traitement du flot optique aussi bien que l'interprétation du mouvement et de la structure 3D demandent l'évaluation des dérivées de de l'image de niveaux de gris. Mais les images réelles sont en général bruitées et l'évaluation des dérivées est une opération très sensible au bruit. D'autre part, les discontinuités du flot optique dues aux occlusions demandent une bonne détection des régions homogènes sinon le non respect de l'hypothèse de continuité mettrait cette technique en défaut.

IV) CONCLUSION.

Nous avons analysé, dans ce chapitre, des méthodes classiques d'interprétation de scènes tridimensionnelles. L'étude critique des différentes méthodes qui ont pu être proposées pour répondre à l'application envisagée a porté sur les possibilités d'utilisation de chacune d'elles. Il ressort de cette étude qu'aucune de ces méthodes ne répond à l'ensemble des contraintes imposées par le cahier des charges. Pratiquement chaque méthode existante correspond à type d'application précis et il n'y a pas de méthode universelle. .

Notre objectif est d'obtenir la position dans l'espace et la taille des objets pouvant pénétrer sur les voies de métro. Les approches monoculaires du type "shape from" ne donnent qu'une information sur la forme et l'orientation des surfaces. Elles nécessitent une connaissance préalable de la nature des objets présents dans la scène ce qui les rend inadaptées à notre problème. Une méthode efficace pour obtenir l'information tridimensionnelle est la méthode de triangulation optique utilisant une source LASER. Cependant elle ne peut être utilisée que dans un environnement bien contrôlé, éliminant toute possibilité d'application sur des scènes réelles.

L'analyse d'une séquence d'images exploite le mouvement des corps rigides pour interpréter une scène tridimensionnelle. L'approche flot optique part de l'hypothèse d'invariance du niveau de gris d'un point dans le temps. Cette hypothèse peut être mise en défaut par les conditions régnant dans les stations de métro. La mise en correspondance des primitives caractéristiques extraites d'images successives doit être robuste pour limiter les faux appariements. Un échantillonnage temporel rapide permet en général de fiabiliser cette approche. Chaque nouvelle acquisition apporte un flot important de données, ce qui empêche un traitement temps réel.

La focalisation automatique est une des méthodes qui permettent de déterminer la distance d'un objet au système optique en utilisant des algorithmes relativement simples et efficaces. Elle souffre cependant d'une limitation de précision dans certaines zones du champ observé.

La stéréovision semble être la seule approche suffisamment précise et fiable, bien que la mise en correspondance des primitives reste un problème délicat à résoudre. Elle est la méthode la plus utilisée pour la détermination des distances objets-caméras dans le cadre de scènes réelles.

Combiner la stéréovision et le mouvement est un bon exemple d'une classe de problèmes dans lesquels plusieurs capteurs ou sources d'informations ont besoin d'être combinés. La théorie complète de l'intégration des capteurs reste pour l'essentiel à faire même si quelques travaux ont été récemment dédiés à ce problème [AYA 87]. Sur le problème spécifique de l'intégration de la stéréovision et du mouvement, très peu de choses ont été faites.

Toutes les méthodes décrites dans ce chapitre ne permettent pas de répondre aux exigences de l'application. La complexité d'une scène réelle prise avec une caméra matricielle ne facilite pas une interprétation tridimensionnelle rapide et efficace. L'analyse de notre problème présentée au chapitre I nous permet d'envisager l'étude d'autres systèmes de détection d'objet. En effet, le problème que nous souhaitons résoudre n'est pas forcément la présence d'un objet dans une zone dangereuse mais en fait d'analyser le passage d'un objet d'une zone dite autorisée à une zone interdite. Cette réflexion peut nous permettre de résoudre le problème non plus avec des caméras matricielles mais avec des caméras linéaires.

CHAPITRE III

✧ BARRIÈRE OPTIQUE DE SURVEILLANCE

I) INTRODUCTION

Dans le cahier des charges nous avons souligné l'intérêt de définir la position dans l'espace d'un objet détecté par le système de surveillance des voies de métros et des bords de quais. Le chapitre II fait une synthèse des méthodes de reconstruction tridimensionnelle d'une scène à partir d'une ou plusieurs caméras matricielles. Aucune de ces méthodes ne semble réellement adaptée à la fois à l'environnement d'une station de métro et à la contrainte temps réel. En effet, l'efficacité d'un système de surveillance dépend de son aptitude à répondre en temps réel aux événements extérieurs, tout en étant capable de différencier les différents types d'intrusions dans un environnement parfois difficilement contrôlable.

La stéréovision matricielle est la méthode qui donne les meilleurs résultats sur la position tridimensionnelle d'un objet dans une scène quelconque. Néanmoins le principal défaut de cette technique est la complexité des algorithmes efficaces de mise en correspondance qui interdit tout traitement temps réel. Dans cette optique temps réel, il apparaît donc souhaitable de minimiser le volume d'information à traiter tout en conservant le principe d'une localisation des objets en mouvement dans l'espace 3-D de la scène surveillée. Notre souci a été de concevoir une méthode qui tout en s'inspirant de la stéréovision classique, soit plus simple et plus rapide.

Nous proposons un nouveau système de détection mettant en oeuvre une machine de vision utilisant des caméras linéaires. Ce système est dédié au contrôle des intrusions dans des zones bien délimitées. L'approche proposée est basée sur la définition d'un plan de surveillance constitué d'une caméra linéaire visant une rampe de tubes fluorescents.

L'utilisation de caméras linéaires réduit considérablement la quantité d'information à traiter par rapport aux techniques mettant en oeuvre des caméras matricielles.

II) DEFINITION D'UN PLAN DE SURVEILLANCE

II.1) Principe de détection

Le plan défini par une caméra linéaire visant une rampe de tubes fluorescents constitue l'élément de base du système de détection d'intrusion proposé (Fig. 3- 1).

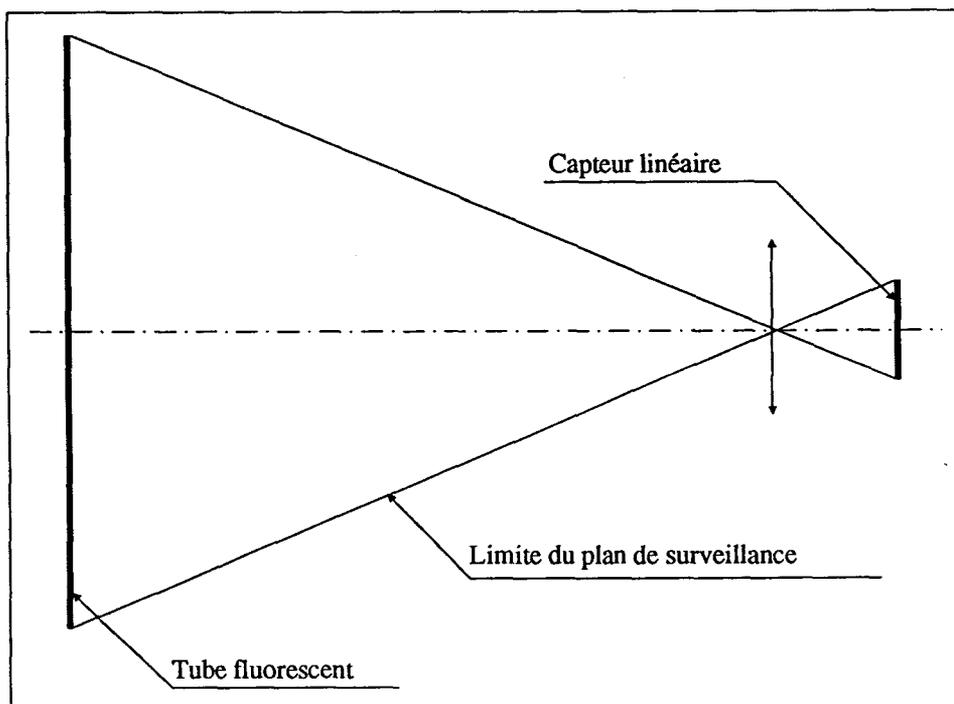


Fig. -3-1: Définition d'un plan de surveillance.

Si aucun objet ne coupe le plan, la rampe d'éclairage est entièrement vue par la caméra et fidèlement reproduite sur l'image ligne (Fig. 3-2). L'image ligne est constituée de 1728 valeurs correspondant au nombre d'éléments photosensibles (pixels) du capteur. Ces valeurs numériques ou niveau de gris sont codées sur les 6 bits de poids fort d'un octet. Ceci donne 64 niveaux de gris sur une dynamique de 256 niveaux possibles.

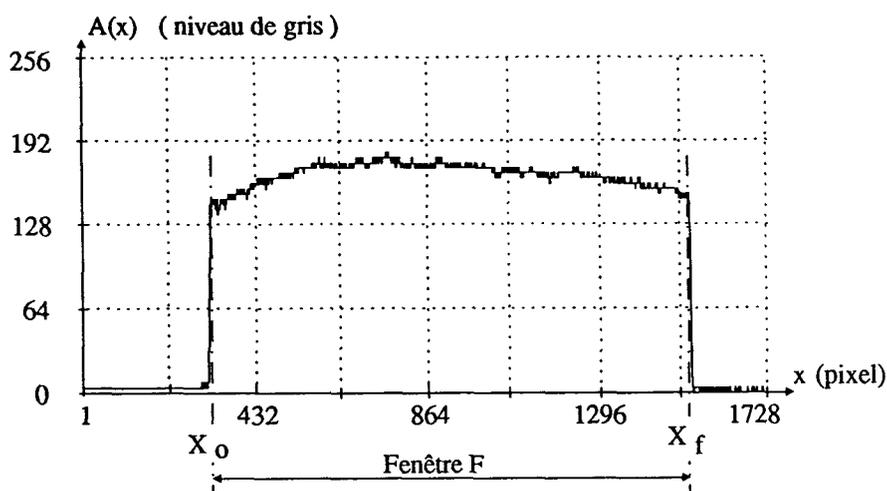


Fig. -3-2 : Image ligne sans objet dans le plan.

L'image ligne représentée par une fonction $A(x)$ traduit le niveau de gris du pixel d'abscisse x telle que:

$$0 \leq A(x) \leq 255 \text{ avec } 1 \leq x \leq 1728$$

D'autre part, on définit une fenêtre F sur l'image ligne correspondant à la partie du capteur sur laquelle est formée l'image du tube fluorescent (Fig. 3-2). Cette fenêtre est délimitée par une abscisse origine X_0 et une abscisse fin X_f telles que si le pixel x appartient à la fenêtre F alors, $X_0 \leq x \leq X_f$.

La détection ainsi que les calculs se feront dans cette fenêtre F .

Dés qu'un objet opaque traverse le plan de surveillance ainsi défini, il laisse une trace dans l'image ligne ce qui permet de détecter aisément sa présence (Fig. 3-3).

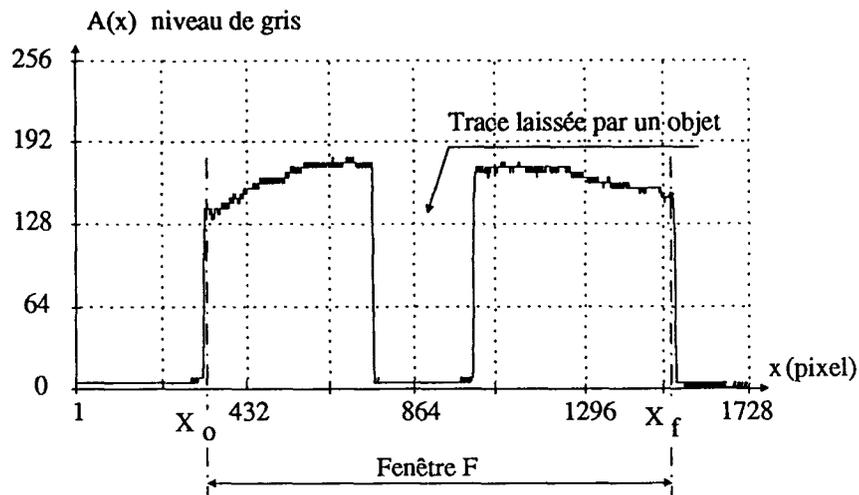


Fig. -3-3 : Image ligne avec objet dans le plan.

Le but de l'étude n'est pas seulement de détecter la présence d'un objet mais aussi de donner la position de cet objet dans le plan de surveillance. Les techniques de mesure de profondeur par focalisation décrites au chapitre précédent peuvent être mise en oeuvre sur les image lignes pour obtenir l'information de position recherchée.

II.2) Focalisation.

La mise au point de l'objectif étant ajustée sur la rampe lumineuse, on peut remarquer qu'un objet proche de la caméra ne donne pas la même image ligne qu'un objet éloigné de celle-ci. La figure 3-3 représente l'image ligne d'un objet proche de la rampe lumineuse tandis que la figure 3-4 représente un objet proche de la caméra.

L'objet donnant l'image de la figure 3-4 se trouve éloigné de la distance de mise au point. Les variations de niveaux de gris dues à la présence de cet objet dans le champ de la caméra donnent une pente d'autant moins accentuée que la défocalisation est importante. En effet, l'épaisseur du contour-image d'un objet varie en fonction de la position de l'objet. Pour mieux distinguer les effets de la focalisation on ne représente sur les images ligne des figures 3-6 à 3-10 qu'une fenêtre de visualisation de 256 pixels sur les 1728 pixels qui constituent l'image ligne (Fig. 3-5).

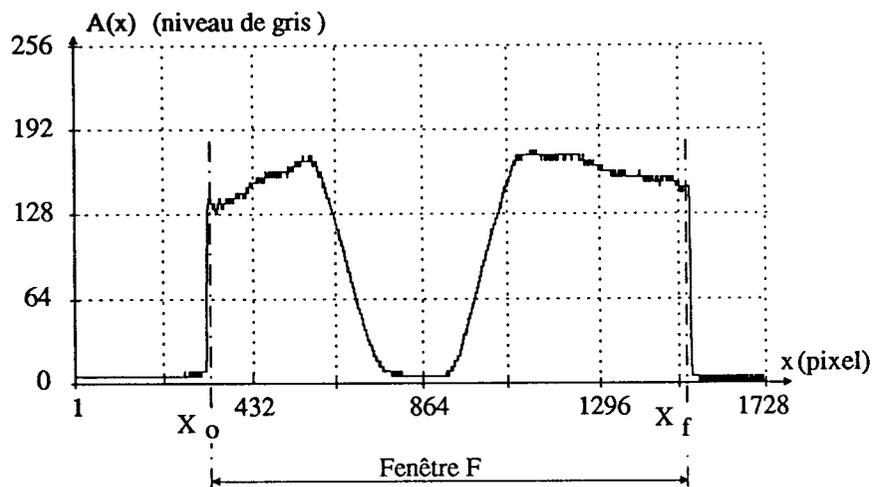


Fig. -3-4: Image ligne d'un objet proche de la caméra.

Ces figures montrent le contour de l'image d'un objet traversant le plan de surveillance et situé à différentes distances d , exprimée en pourcentage de la distance D de mise au point.

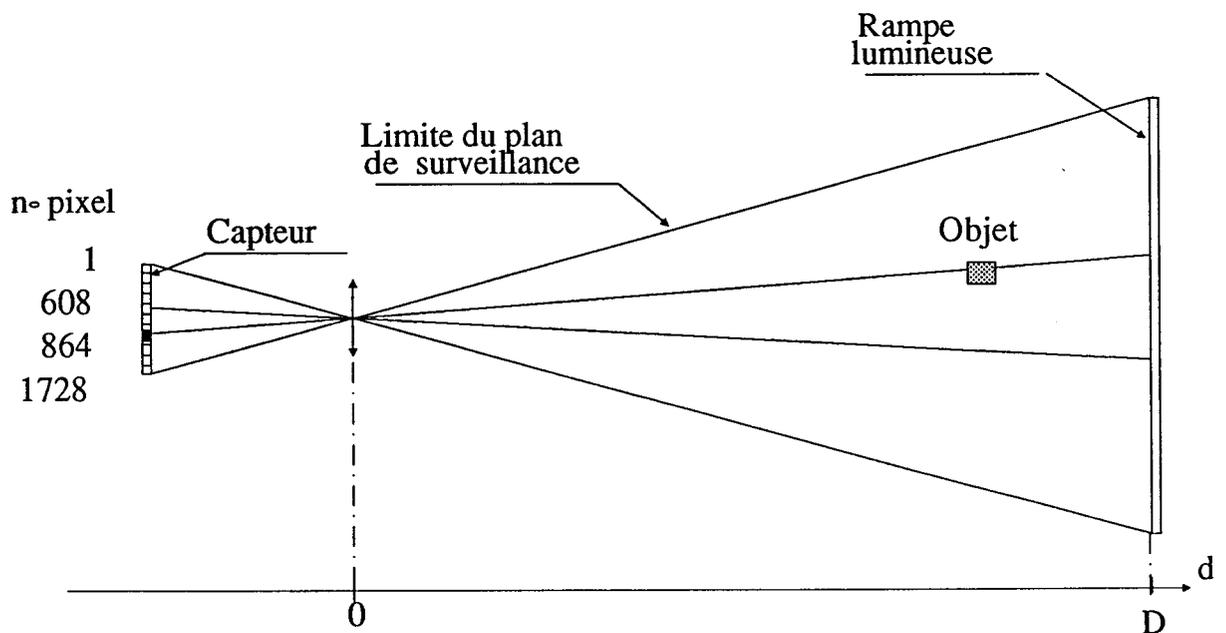


Fig.-3-5-Effet de la focalisation-Montage expérimental

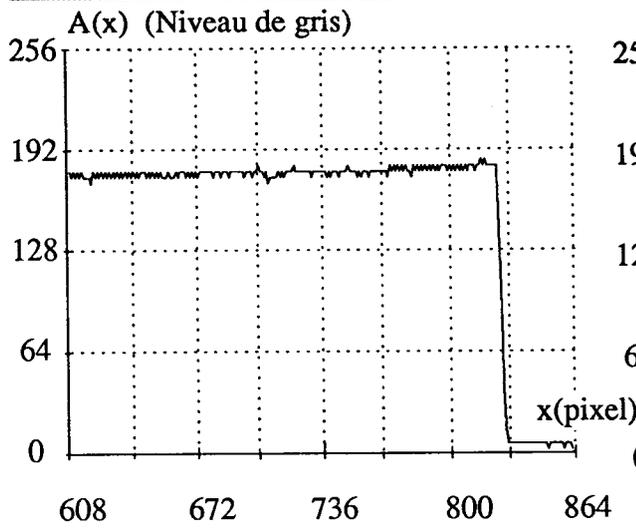


Fig. -3-10-Objet à D.

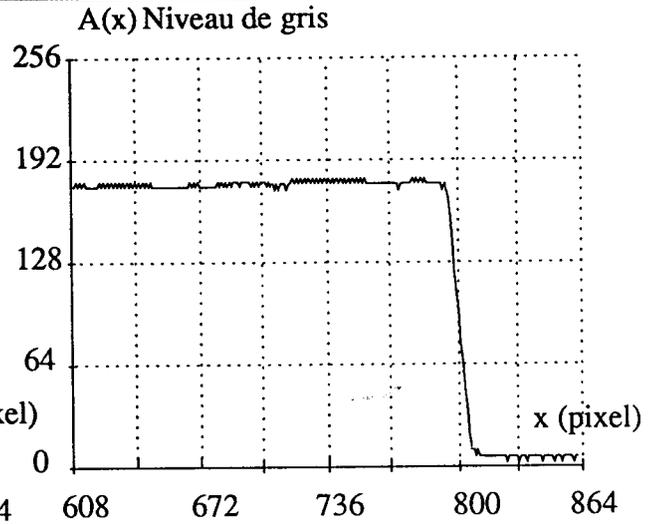


Fig. -3-6-Objet à 0,8 D.

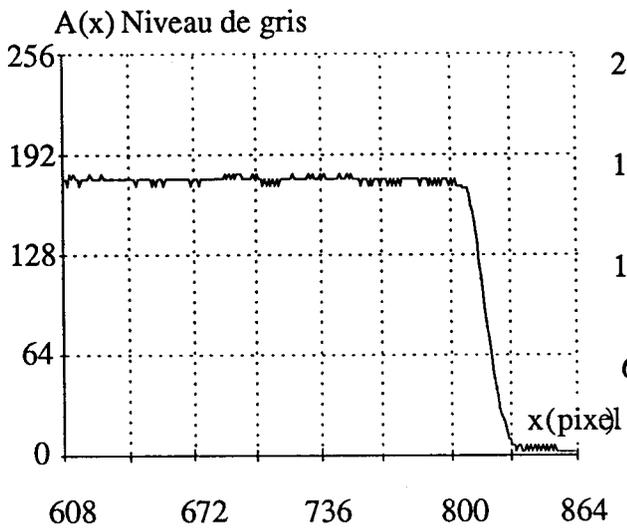


Fig. -3-7-Objet à 0,6 D.

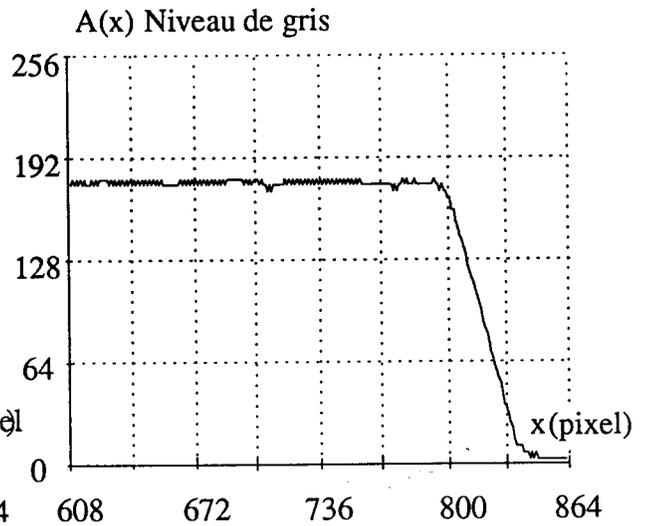


Fig. -3-8-Objet à 0,4 D.

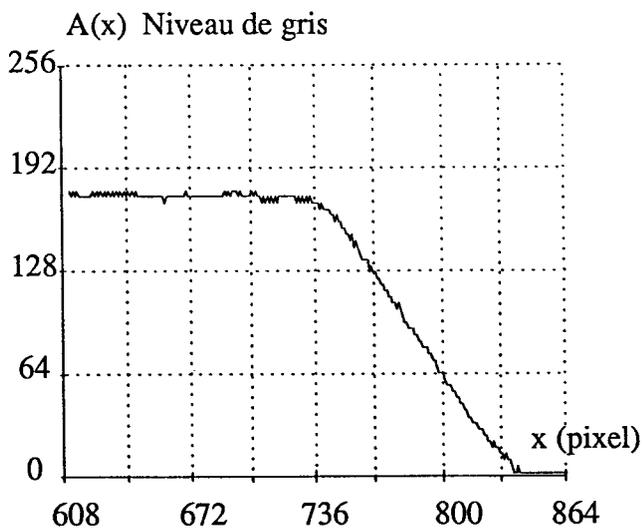


Fig. -3-9-Objet à 0,2 D.

La mesure de la pente $P(x)$ du contour est obtenue par une estimation de la dérivée de la fonction de niveau de gris $A(x)$ au pixel d'abscisse x sous la forme :

$$P(x) = \frac{|A(x+k) - A(x)|}{k} = \frac{\Delta A}{k}$$

pour chaque pixel x telle que $X_0 \leq x \leq X_f - k$ où k est un entier tel que entre $1 \leq k \leq X_f - X_0$ qui représente la taille de l'opérateur gradient unidimensionnel.

La majorité des algorithmes de mesure de distance par focalisation associe le critère de "netteté d'une image" à un opérateur du type Gradient ou Laplacien [LIG 82]. Plus l'intensité du gradient est élevée plus l'image d'un objet est "nette" car ses contours sont abruptes et inversement, plus le gradient est faible moins l'image est focalisée car ses contours sont larges. Le problème de cette technique de mesure se situe dans le choix de la taille k de l'opérateur gradient. Un opérateur de petite taille favorise la mesure de distance des objets proches de la distance de mise au point. Un opérateur de grande taille effectue un filtrage passe-bas sur l'image et diminue ainsi le contraste aux hautes définitions ce qui rend l'image floue; il convient mieux aux mesures de distances d'objets éloignés de la distance de mise au point.

Pour éviter ce problème, nous avons abordé la mesure de netteté en déterminant la largeur k en pixels de l'intervalle sur lequel apparaît la chute de niveau de gris correspondant au passage de la partie éclairée du capteur à celle occultée par l'objet opaque (Fig. 3-11). Ceci n'est possible que si la chute de luminosité ΔA attendue est constante, c'est à dire si l'éclairage ne présente pas d'importantes variations spatiales et temporelles d'intensité en absence d'objet dans le champ de la caméra.

k représente alors l'épaisseur en pixels du contour dans l'image ligne.

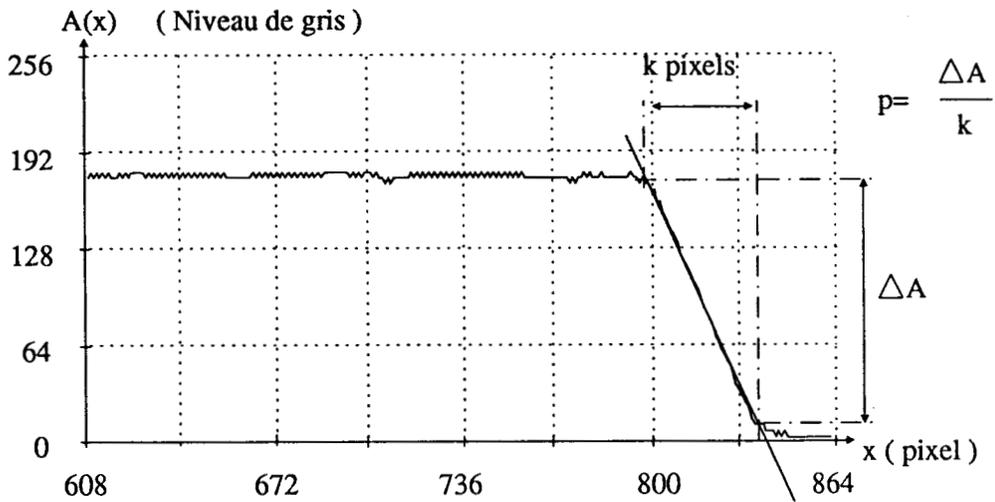


Fig. -3-11- Calcul de la pente.

La pente du contour d'un objet occultant la rampe lumineuse sur une image ligne dépend de paramètres physiques tels que l'intensité lumineuse reçue par chaque pixel, la qualité de l'objectif. Un étalonnage préalable donnant la position en fonction de la pente du contour est donc nécessaire. La courbe de la figure 3-12 représente le relevé de la distance d séparant la caméra d'un objet occultant le plan de surveillance en fonction de la pente $P(x)$ pour ΔA constant.

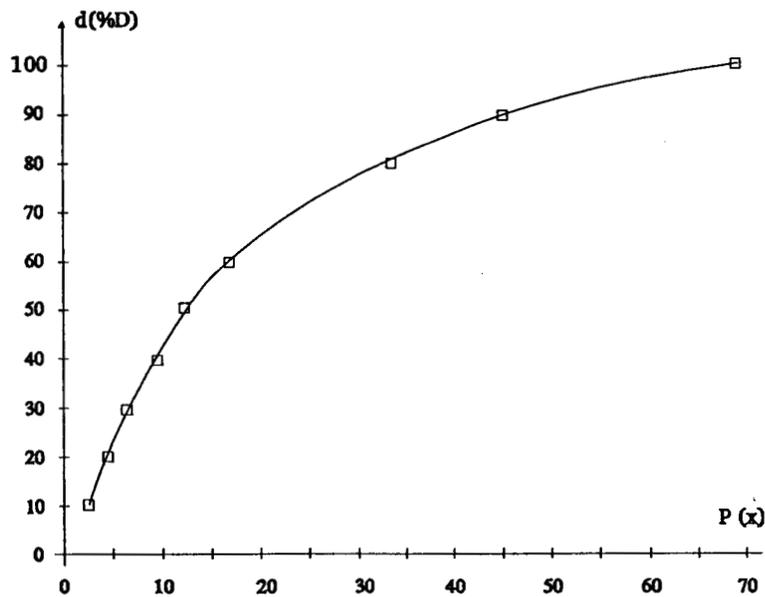


Fig. -3-12- $d = f(P(x))$.

Soulignons que la courbe ci-dessus a été tracée en plaçant les objets le long de l'axe optique de la caméra c'est à dire au centre de l'image. Une seule courbe au centre ne garantit pas une exploitation correcte des mesures pour des objets situés sur les bords de l'image. Le centre de l'image est en général la zone de sensibilité maximale. Le contraste dans les zones de faible sensibilité correspondant au bords de l'image est inférieur au contraste dans les zones de forte sensibilité. Ceci est dû à la non uniformité de l'éclairage dont l'intensité diminue sur les bords de l'image (cf Fig. 3-2).

Cet étalonnage permet de déterminer la position d'un objet détecté grâce à la mesure de la pente $P(x)$ du contour de son image.

On peut toutefois remarquer sur la courbe de la figure 3-12 différentes zones:

- La seconde zone correspond à la partie croissante de la courbe. Dans cette zone la mesure de $P(x)$ conduit à une valeur correcte de la distance. La distance d varie en fonction de p telle que
$$d = \frac{p \cdot K \cdot D}{p \cdot K \cdot D + \Delta A \cdot l_e}$$
 (cf ANNEXE 2). Cette fonction de d dépend de l'ouverture utile de l'objectif $\frac{f}{n}$, de la distance du plan de mise au point D et de la taille d'un élément photosensible l_e .

- La troisième zone, la pente variant peu limite la précision sur la mesure de distance. Cette partie définit la profondeur de champ de l'objectif avant le plan de mise au point (cf ANNEXE 2 pour le calcul de la profondeur de champ). A l'intérieur de cette zone tous les objets paraissent vus avec la même netteté. On sait que la profondeur de champ est d'autant plus grande, que l'objet visé est éloigné de la caméra, que la distance focale est petite et que l'ouverture est petite. Mais il faut également remarquer l'effet de l'échantillonnage spatial sur cette profondeur de champ. En effet, les éléments photosensibles de la caméra n'étant pas des points mais de petites surfaces, l'image d'un point d'un objet à différentes distances peut toujours être projetée sur le même élément. Si les dimensions des éléments augmentent, la profondeur de champ augmente et la mesure de distance devient moins précise (Fig. 3-13). On note $[s]$ et $[r]$ la longueur respective des images des points a et b sur le plan de mise au point, et $[r']$ et $[s']$ la longueur des images des

points a et b sur l'image ligne. Si les longueurs [r'] et [s'] sont égales à la taille d'un élément photosensible alors la distance entre les points a et b est la profondeur de champ car les points a et b seront vus avec la même netteté.

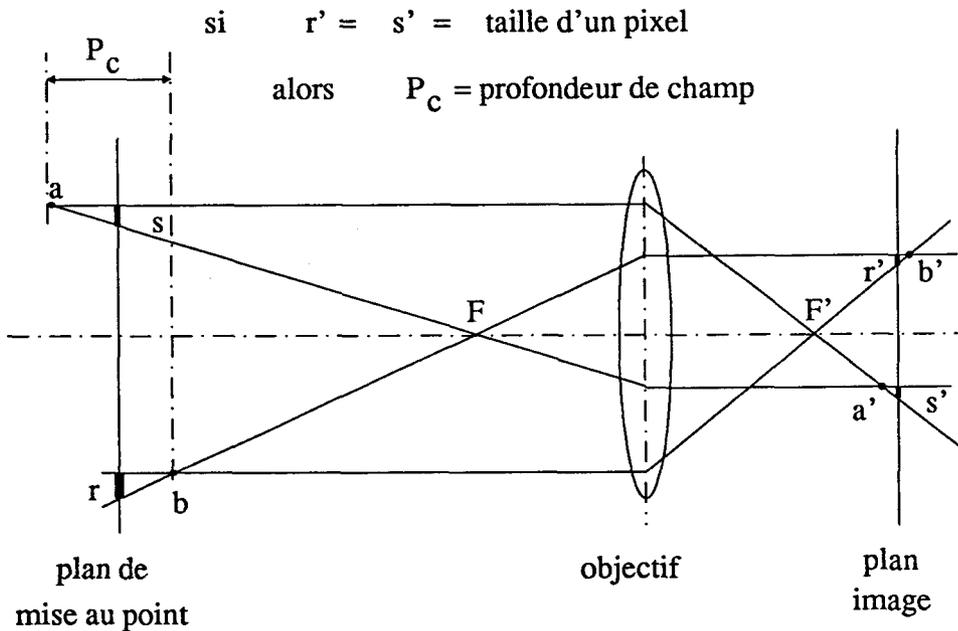


Fig. -3-13:Profondeur de champ.

Le dessin de la figure 3-13 ne tient pas compte de l'ouverture du diaphragme pour la représentation de la profondeur de champ.

Ce système de détection à base d'une caméra linéaire utilisant une technique de mesure de netteté des image lignes donne la position des objets traversant le plan de surveillance. Cependant, la mesure de distance par focalisation est très sensible au bruit de mesure sur les pixels essentiellement dans la zone de profondeur de champ où l'épaisseur k du contour se limite à quelques pixels.

Ce plan de surveillance sera toutefois utilisé comme élément de base du système de détection qui comportera deux plans identiques afin d'accéder à la mesure de profondeur par des techniques de stéréovision. Avant de combiner l'information dans deux images lignes, il convient de déterminer avec précision les traces laissées sur les capteurs par les objets traversant les plans de surveillance. Nous abordons donc maintenant ce problème de segmentation.

II.3) Segmentation des images lignes.

Comme en éclairage diascopique, les objets apparaissent nettement en foncé sur fond clair et comme les variations spatiales d'éclairement restent faibles, on peut tenter de segmenter l'image ligne par un simple seuillage à niveau fixe. Cette technique permet un traitement binaire de l'image puisque l'on n'a alors que deux niveaux de gris, le noir pour les objets et le blanc pour la source lumineuse.

L'avantage de la méthode de segmentation avec un seuil constant est sa simplicité et sa rapidité de mise en oeuvre. En effet, le système de vision linéaire utilisé autorise une implémentation matérielle d'un seuil fixe. La segmentation s'effectue en même temps que l'acquisition des données de la caméra en mémoire ligne. Les adresses des pixels qui franchissent le seuil sont mémorisées dans un buffer. Dans une image ligne, les contours d'un objet opaque traversant le champ de la caméra sont donc définis par deux valeurs correspondant aux adresses des extrémités de la trace de l'objet.

Le problème majeur de cette méthode est la détermination du seuil par rapport aux niveaux de gris correspondant à la rampe lumineuse et ceux correspondant à la présence d'un objet. En effet les calculs ultérieurs des paramètres de taille et de position des objets dépendent beaucoup de la précision de la segmentation.

Il est important de remarquer que les objets de différentes tailles peuvent laisser une trace identique dans l'image ligne. Inversement, des objets de tailles identiques peuvent produire des ombres différentes (Fig. 3-5). Dans le cas d'un objet proche de la rampe lumineuse la valeur du seuil n'est pas critique. Par contre lorsque l'objet se trouve près de la caméra, la taille de l'image de l'objet dépend de la valeur du seuil. En effet l'épaisseur du contour de l'image d'un objet varie avec la position de celui-ci (cf paragraphe II-2) .

Le problème est de trouver le seuil qui respecte la taille de l'objet, quelque soit sa position. Si on considère deux objets identiques situés à des distances différentes de la caméra le long de l'axe optique le plus proche de ces deux objets étant à une distance a du système optique égale à m fois la distance b entre les deux objets alors les tailles respectives de leurs images sont dans le rapport $\frac{m}{m+1}$.

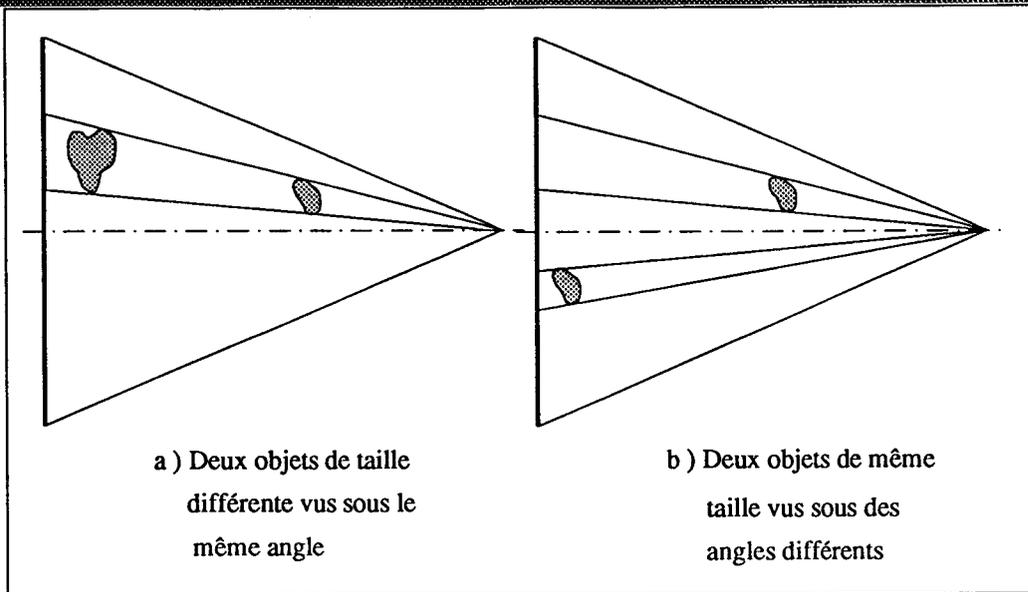


Fig. -3-14: Effet de la projection de différents objets.

Nous allons montrer comment déterminer le seuil en plaçant un objet à deux positions différentes le long de l'axe optique de la caméra. Ces deux positions sont déterminées de telle sorte que les deux images du même objet donnent des tailles dans un rapport 1/2.

La configuration expérimentale est donnée figure 3-15 .

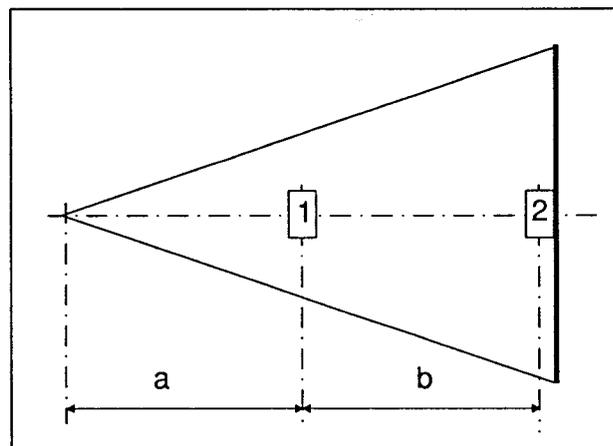


Fig. -3-15: Configuration expérimentale.

Pour que les deux objets-images soient dans le rapport 1/2, on doit avoir $m=1$:

comme $a = mb$, il vient $a = b$.

Dans ces conditions la taille de l'image de l'objet dans la position 1 après seuillage doit être le double de celle dans la position 2 obtenue avec le même seuil (Fig. 3 - 16 et Fig. 3 - 17).

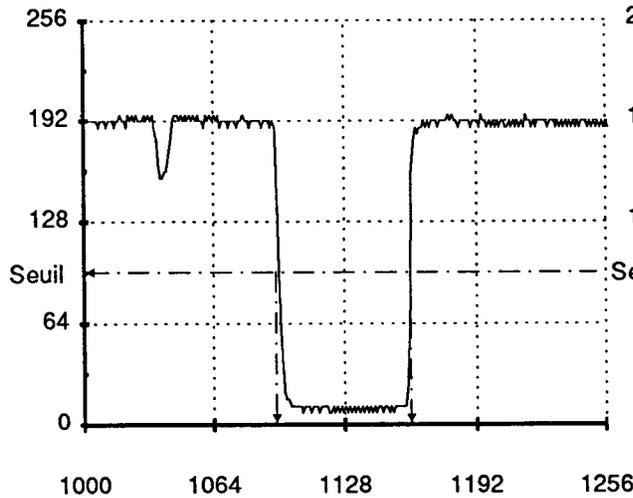


Fig. -3-16: Position 2.

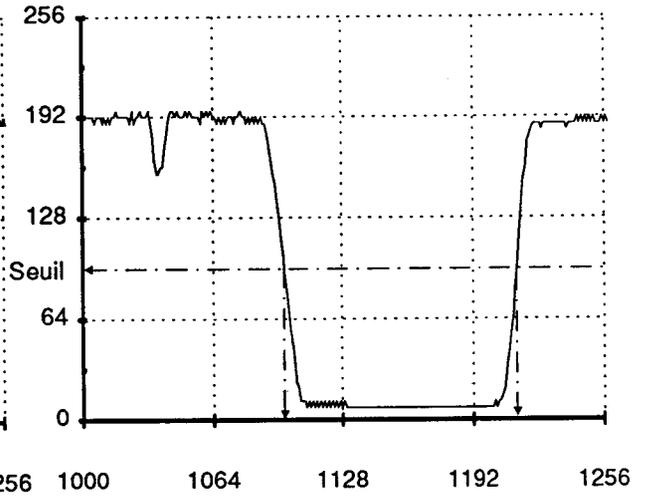


Fig. -3-17: Position 1.

On constate sur les figures 3 - 16 et 3 - 17 que le réglage du seuil à mi-hauteur entre le niveau lumineux de la rampe et le niveau de la partie occultée permet d'obtenir les images binaires attendues.

Pour connaître la taille réelle d'un objet il faut, connaissant la résolution spatiale de l'image numérique (dimension du champ observé/nombre d'éléments photosensibles du capteur), déterminer sa position dans le plan de surveillance. Pour déterminer, en temps réel, la taille et la position des objets pénétrant dans la zone surveillée, on utilise deux plans de surveillance identiques et parallèles.

III) DOUBLE PLAN DE SURVEILLANCE

III.1) Introduction

Afin d'améliorer la perception des objets, on utilise deux plans de base P1 et P2 parallèles et identiques à celui décrit au paragraphe précédent. Deux configurations sont envisageables pour évaluer la position et la taille de n'importe quel objet traversant ces plans, chacune d'elles correspondant à une zone de surveillance de géométrie différente.

La première configuration correspond à une géométrie de stéréovision classique où les caméras observent la même scène. La droite épipolaire en vision linéaire s'identifie alors à l'image ligne elle-même. Le calcul des coordonnées de position s'effectue avec les mêmes équations que celles définies pour la stéréovision matricielle (cf ANNEXE 1).

La deuxième configuration est une version particulière de la vision binoculaire pour laquelle les caméras linéaires observent la même scène mais en se faisant face. Néanmoins le calcul des paramètres de position se fait en utilisant les mêmes équations que dans la configuration précédente, mais les caractéristiques particulières de la géométrie de cette seconde configuration conduisent à des simplifications substantielles.

Il est à noter que la coordonnée selon l'axe Oy qui n'existe pas pour une image ligne ne sera pas calculée. Pour rester cohérent avec les notations utilisées au chapitre II pour le calcul des coordonnées en stéréovision matricielle, on gardera la même disposition des axes. L'axe Ox sera l'axe de l'image ligne, l'axe Oy disparaîtra et l'axe Oz coïncidera avec l'axe optique de la caméra.

III.2) Caméras visant la même rampe lumineuse.

III.2.1) Principe.

Dans la première configuration, présentée figure 3-18, les deux caméras C et C', séparées d'une distance $2.D$, visent la même rampe de tubes lumineux. La distance séparant les deux plans étant faible, on les considère comme confondus dans les calculs géométriques suivants.

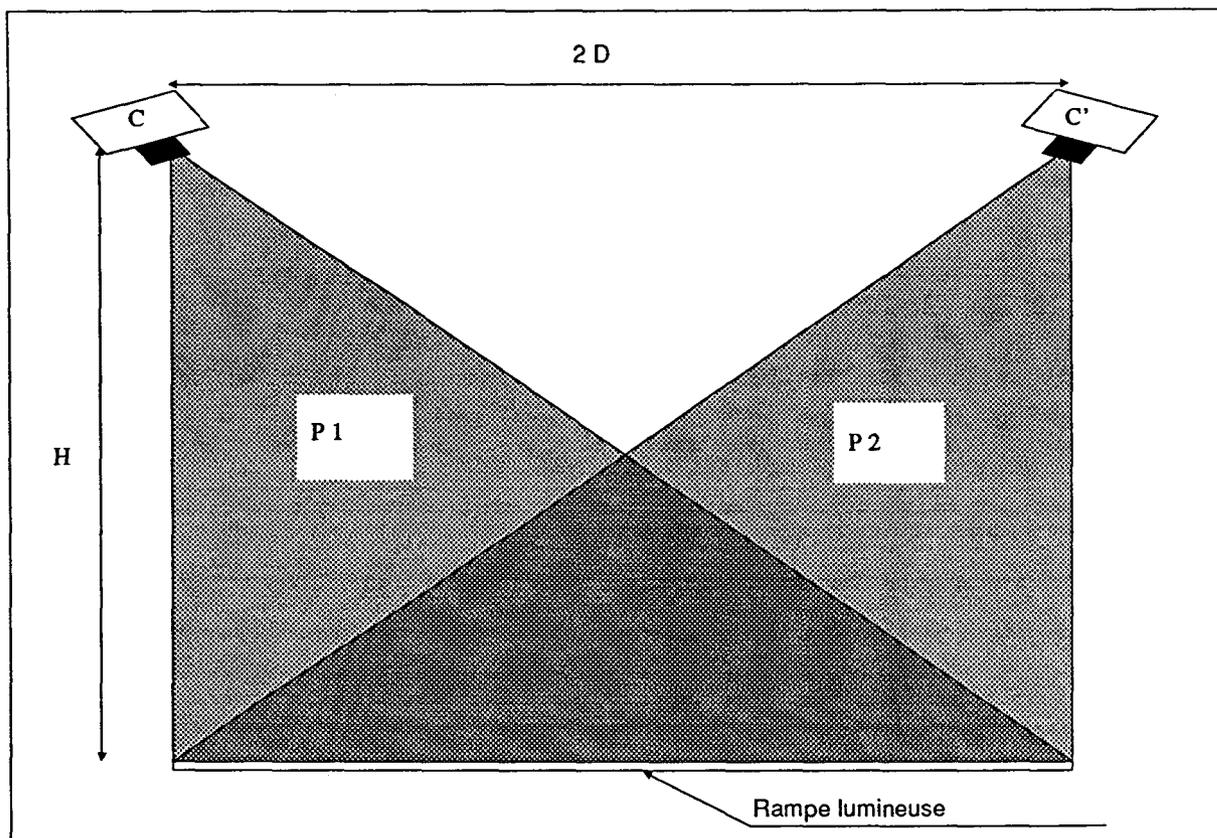


Fig. -3-18- Position relative des deux plans.

III.2.2) Détermination de la position des objets.

Comme pour la stéréovision matricielle nous définissons un repère pour chaque caméra. Dans chaque repère, l'axe des z coïncide avec l'axe optique de la caméra associée. Les images ligne sont parallèles aux axes Ox et $O'x'$. La distance focale est choisie comme unité, $f = f' = 1$. Un objet traversant la barrière de surveillance est inclus dans un quadrilatère $ADBC$ défini par les ombre projetées sur les capteurs des caméras linéaires. Les ombres projetées par ce quadrilatère sur chacune des

caméras nous permettent de calculer les coordonnées (X', Z') du point A par rapport au repère $O' x'z'$ associée à la caméra de droite C'.

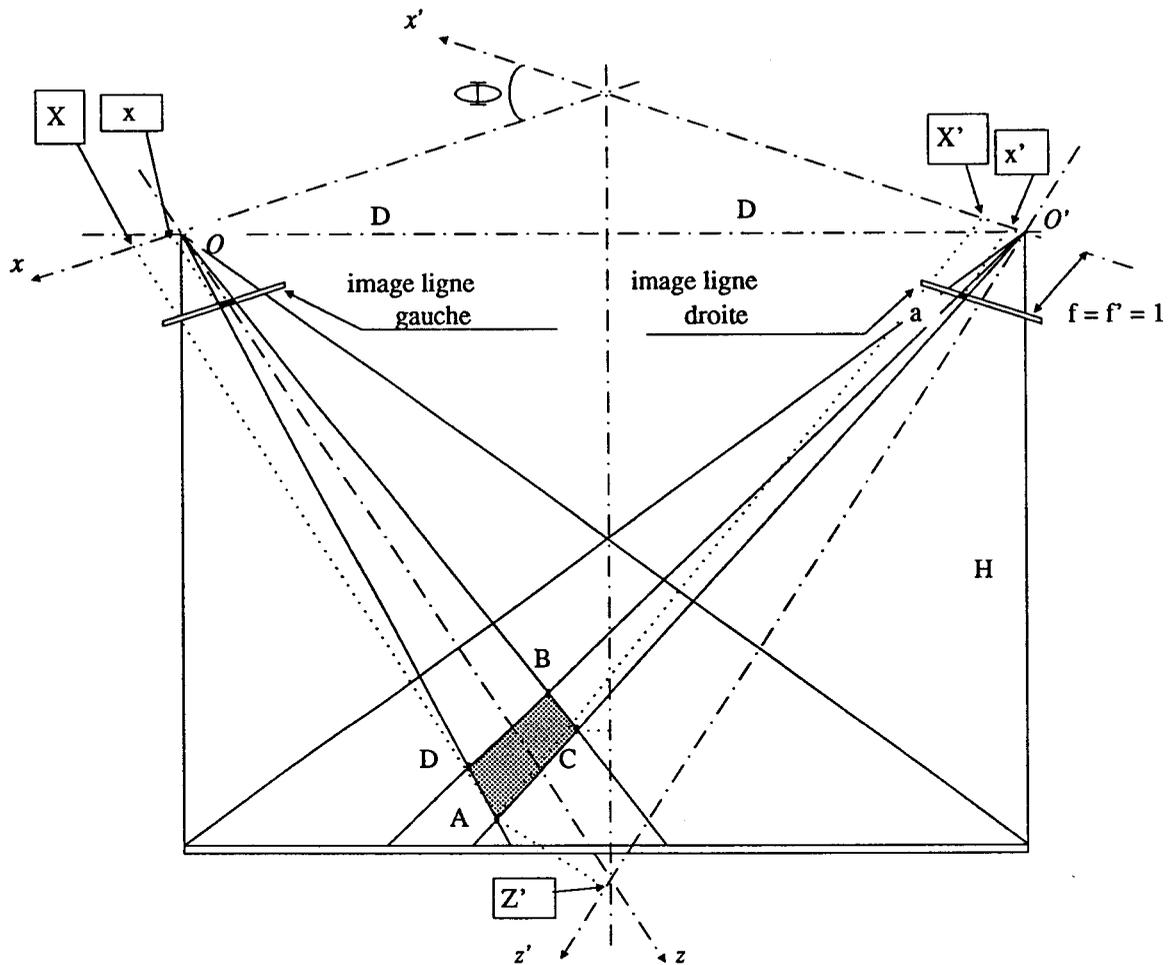


Fig. -3-19:Principe de détermination de la position.

Soient x et x' les abscisses de la projection perspective du point physique A sur les images lignes, recueillies par les deux caméras respectivement de gauche et de droite.

On exprime les coordonnées du point A dans le repère associée à la caméra de droite (X', Z') :

$$\begin{pmatrix} X' \\ Z' \\ 1 \end{pmatrix} = A \cdot \begin{pmatrix} X \\ Z \\ 1 \end{pmatrix} \quad \text{avec} \quad A = \begin{pmatrix} \cos\Phi & \sin\Phi & b_x \\ -\sin\Phi & \cos\Phi & b_z \\ 0 & 0 & 1 \end{pmatrix}$$

b_x et b_z sont les composantes du vecteur de translation $O\vec{O}'$ dans le repère associé à la caméra de gauche et Φ l'angle de rotation de l'axe Ox par rapport à l'axe $O'x'$ dans le plan Oxz . A partir des équations ci-dessus on exprime les coordonnées du point $A (X', Z')$ en fonction de x, x' et des paramètres b_x, b_z et Φ définissant la géométrie du système (cf ANNEXE 1).

$$X' = \frac{\cos \Phi (x' . b_x - x . x' . b_z) - \sin \Phi (x' . b_z + x . x' . b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 - x . x')}$$

$$\text{et: } Z' = \frac{\cos \Phi (b_x - x . b_z) - \sin \Phi (b_z + x . b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 - x . x')}$$

On obtient de la même manière les coordonnées des autres points B, C et D en prenant les projections de ces points sur les images lignes de gauche et de droite.

Une procédure de calibration pour les deux caméras peut s'avérer nécessaire si le système optique, composé d'un objectif (association de lentilles) et d'une caméra, présente certaines distorsions, dues principalement aux aberrations optiques. La calibration d'un système caméra-objectif sert à établir une correspondance entre une scène réelle et sa représentation sous la forme d'une image. En théorie, cette correspondance est une transformation linéaire, mais en pratique l'image peut subir des déformations qui mettent en défaut la propriété de linéarité.

La calibration définit un modèle du système caméra-objectif, c'est à dire qu'elle détermine les paramètres de la transformation entre scène et image. Dans notre cas, l'utilisation de caméras linéaires réduit la scène observée à un espace 2D et sa représentation image à un espace 1D. Le modèle du système caméra-objectif s'exprime sous la forme d'une matrice C qui fait correspondre les points de la scène 2D aux points de l'image ligne 1D obtenus par projection perspective. Soit la matrice C telle que:

$$C = (C_{ij}) \quad 1 \leq i \leq 3 \quad \text{et} \quad 1 \leq j \leq 2 \quad C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \\ C_{31} & C_{32} \end{pmatrix}$$

Chaque association entre un point objet de la scène et un point image fournit une équation à six variables. Six points de l'espace à des positions connues suffisent

pour déterminer les différents termes C_{ij} . Si on suppose que le système optique n'introduit aucune distorsion alors la matrice C devient:

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

D'autre part la précision de positionnement des caméras est aussi un critère de performance du système de mesure 3D par triangulation. Les positions relatives des deux caméras définissent les paramètres de translation et de rotation pour passer du repère associé à une caméra au repère associé à l'autre caméra. Deux types de calibration sont possibles. La figure 3-19 sera prise comme référence géométrique pour les explications qui suivent.

a) *Calibration mécanique.*

Dans un repère à deux dimensions d'axes Ox et Oz , trois paramètres suffisent pour déterminer complètement la translation et la rotation. On suppose les centres optiques de chaque caméra dans le même plan Oxz ($y=0$). b_x et b_z sont les composantes de la translation et Φ l'angle de rotation de l'axe Ox dans le plan Oxz par rapport à l'axe Ox' . Ces paramètres peuvent être calculés lors de la phase d'installation matérielle de la barrière de surveillance. On mesure la distance séparant les caméras de la rampe lumineuse H , ainsi que la longueur de cette même rampe, correspondant à la fenêtre de détection sur l'image ligne. Les caméras seront installées de telle sorte que les axes optiques de chacune d'elles passent par le pixel au centre de leur fenêtre de détection respective $\frac{(X_f - X_o)}{2}$, dans ce cas:

$$\Phi = \arctg \frac{2.D}{H} , \quad b_x = 2.D \cdot \cos \frac{\Phi}{2} \quad \text{et} \quad b_z = 4.D^2 \cdot \sin^2 \frac{\Phi}{2}$$

b) *Calibration optique.*

Une méthode efficace consiste à utiliser les caméras elles-mêmes comme instruments de mesure. Les images lignes obtenues par les deux caméras servent de mesure pour la détermination des paramètres de translation et de rotation [FAU 86].

Les coordonnées (x', z') d'un point image a correspondant au point objet A de coordonnées (X, Z) dans le repère $O'x'z'$ sont données par:

$$x' = \frac{X'}{Z'}, \quad z' = 1 \quad \text{si } f' = 1.$$

Ces équations définissent la transformation perspective.

Les coordonnées du point O , centre optique de la caméra de droite sont $X'O, Z'O$ dans le repère $O'x'z'$. On souhaite obtenir les coordonnées (x', z') du point image a en fonction des coordonnées $X'O, Z'O$ du point O et des coordonnées (X, Z) du point A dans le repère Oxz . On exprime d'abord les coordonnées (X', Z') du point A dans le repère $O'x'z'$, on obtient:

$$X' = m_{11}(X - X'O) + m_{12}(Z - Z'O)$$

$$Z' = m_{21}(X - X'O) + m_{22}(Z - Z'O).$$

Les valeurs m_{11} à m_{22} dans les expressions ci-dessus représentent les coefficients de la matrice de rotation déterminée par l'angle Φ , $X'O$ et $Z'O$ représentent les composantes du vecteur de translation b_x et b_z . Finalement, on obtient les coordonnées de a :

$$x' = \frac{X'}{Z'} = \frac{m_{11}(X - X'O) + m_{12}(Z - Z'O)}{m_{21}(X - X'O) + m_{22}(Z - Z'O)} \quad (1)$$

$$z' = 1$$

L'expression (1) est l'équation d'une droite dans le repère Oxz de la caméra de droite. Cette droite passe par le centre optique O' de la caméra de gauche et par le point image a . Le point objet A se trouve également sur cette droite. Ses coordonnées dans le repère Oxz de la caméra de gauche sont:

$$x = \frac{X}{Z} \quad (2)$$

$$z = 1$$

L'expression (2) est l'équation d'une droite passant par le point objet A et le centre optique O de la caméra de gauche. La position du point objet A peut alors être trouvée en cherchant le point d'intersection de ces deux droites. Les paramètres m_{11} à m_{22} se déterminent par la recherche des valeurs des paramètres qui minimisent l'écart entre les points objets observés et les points calculés pour un nombre de points objets connus. Cette méthode a été utilisée pour la calibration de système de triangulation active [POS 82].

III.2.3) Exemple de configuration en station.

Les caméras seront montées dans la partie centrale de la station et viseront la rampe lumineuse placée sous le nez de quai.

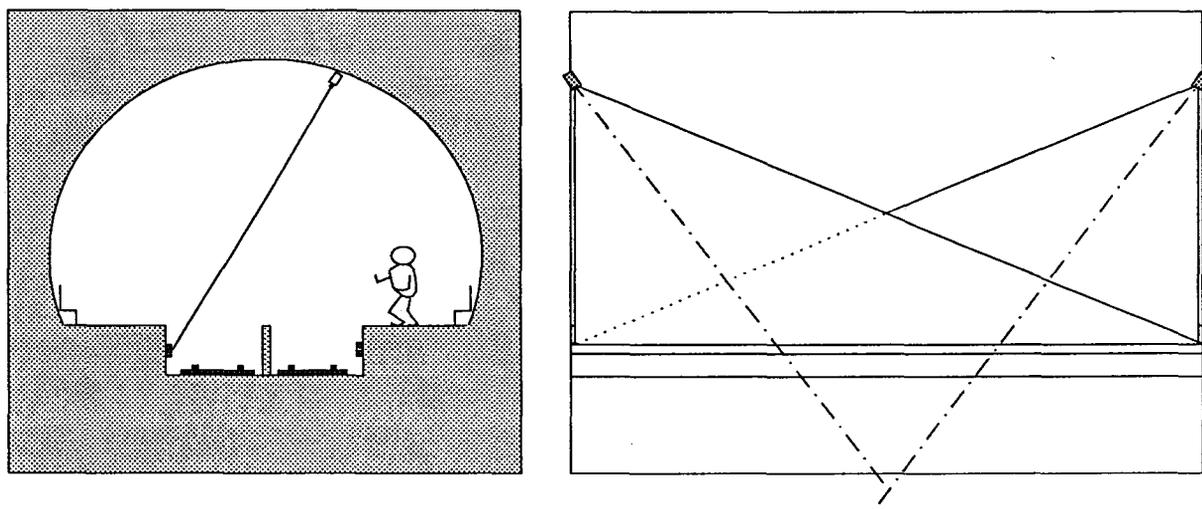


Fig. -3-20-Exemple de système de détection.

III.3) Caméras montées tête-bêche.

III.3.1) Principe.

Dans la deuxième configuration, les caméras sont montées tête-bêche ainsi que les rampes lumineuses (Fig. 3-21).

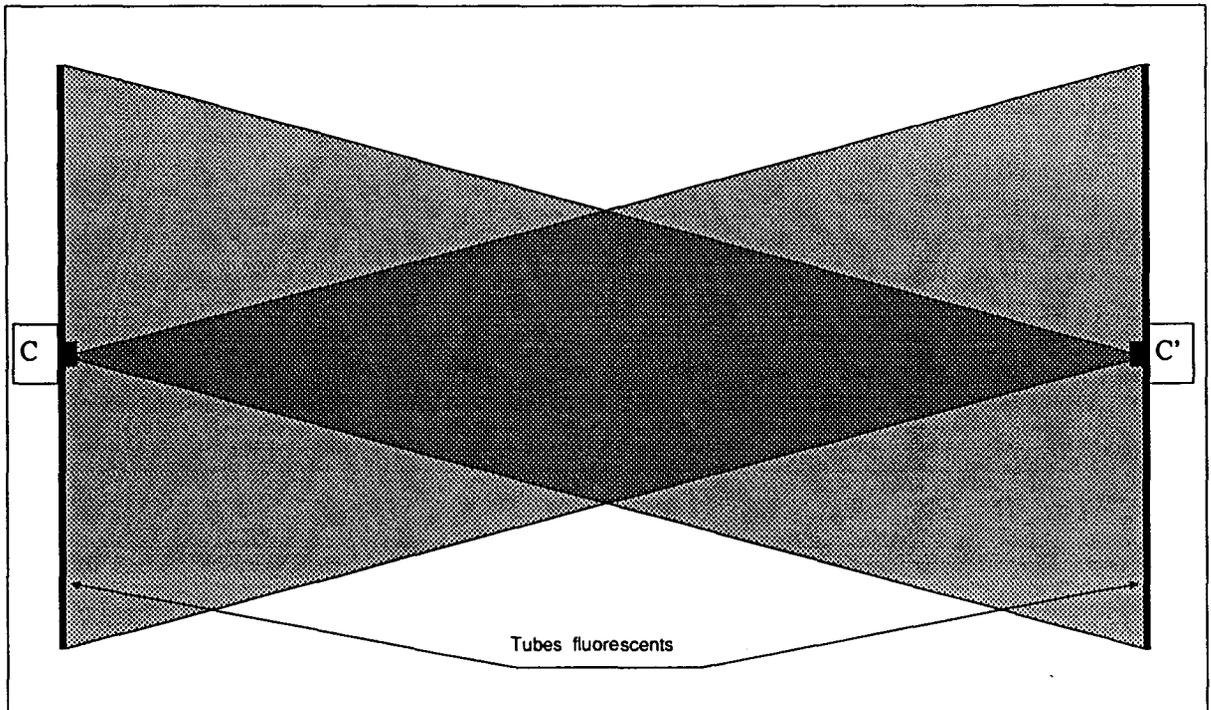


Fig. -3-21:Barrière avec caméras tête-bêche.

III.3.2) Détermination des attributs.

Ce système particulier de stéréovision avec les caméras tête-bêche correspond à un angle de rotation $\Phi = \pi$ et à la composante de translation $b_x = 0$. Dans ce système, les axes optiques de chaque caméra sont confondus, les axes Ox et $O'x'$ sont parallèles et en sens inverse. Comme dans le modèle de stéréovision classique les lignes images sont à une distance focale unité du centre optique de leur caméra respective (Fig. 3-22)

Les équations des coordonnées du point A dans le repère associée à l'image de droite sans aucune simplification due à la géométrie du système, X' et Z' sont:

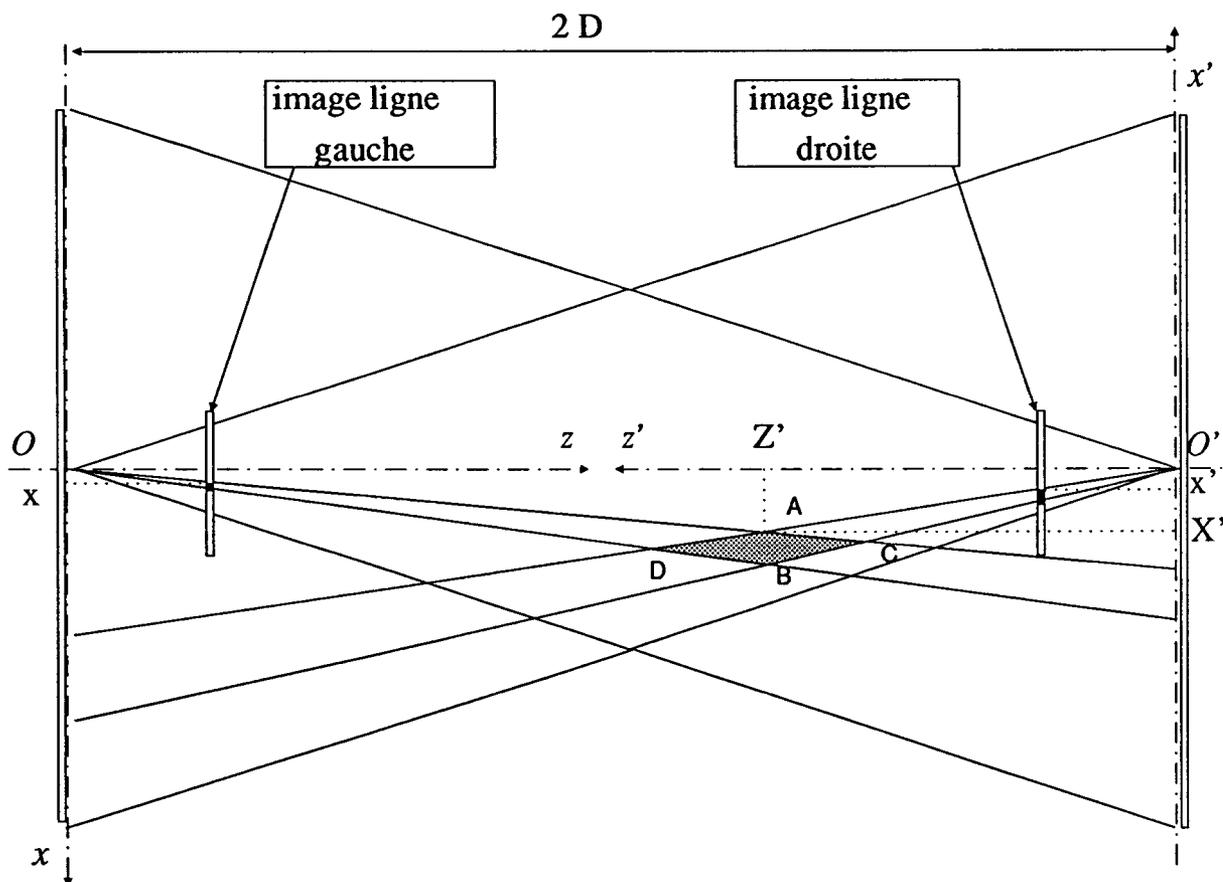


Fig. -3-22: Principe de détermination des attributs.

$$X' = \frac{\cos \Phi (x' \cdot b_x - x \cdot x' \cdot b_z) - \sin \Phi (x' \cdot b_z + x \cdot x' \cdot b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 - x \cdot x')}$$

$$\text{et } Z' = \frac{\cos \Phi (b_x - x \cdot b_z) - \sin \Phi (b_z + x \cdot b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 - x \cdot x')}$$

avec les simplifications $\Phi = \pi$ et $b_x = 0$ ces équations deviennent:

$$X' = - \frac{(x \cdot x' \cdot b_z)}{(x' - x)}$$

$$\text{et } Z' = - \frac{x \cdot b_z}{(x' - x)}$$

Il existe dans cette configuration une indétermination lorsque les abscisses des projections d'un point physique, x et $x' = 0$. Dans ce cas le point physique se trouve sur l'axe optique des caméras $X' = 0$ et Z' est indéterminé.

La phase de calibration mécanique est rigoureusement identique à la configuration précédente. La détermination des paramètres se simplifie considérablement. En effet il ne reste qu'un seul paramètre b_z et celui ci représente la distance séparant les deux caméras. La calibration optique reste néanmoins toujours utile.

III.3.3) Exemple de configuration en station.

Les caméras et les rampes lumineuses seront montées sur le bord du quai.

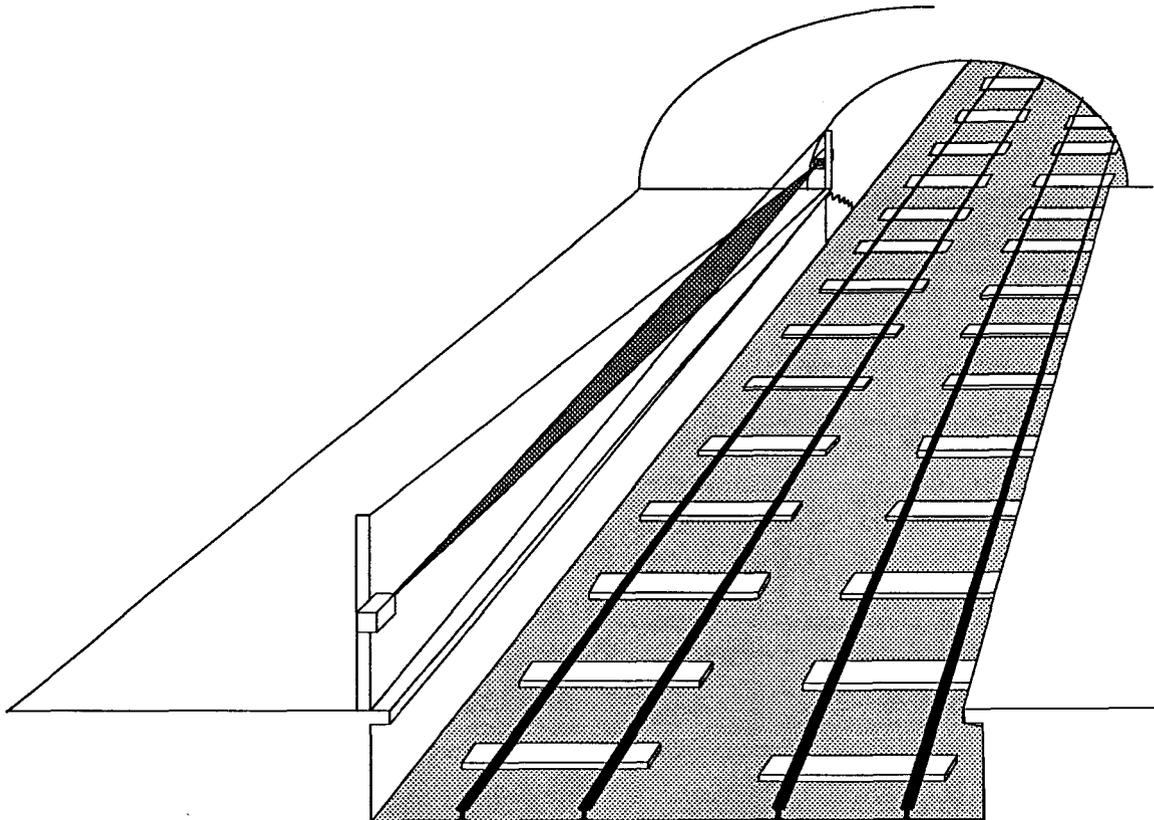


Fig. -3-23-Exemple de système de détection.

IV) CONCLUSION

Quelle que soit la configuration envisagée, un objet traversant la barrière de surveillance est inclus dans un quadrilatère ADBC défini par les ombres projetées sur les capteurs des caméras linéaires (Fig. 3-22) . Le point M de coordonnées (X'_M, Z'_M) telles que,

$$X'_M = \frac{X'_A + X'_B}{2} \quad \text{et} \quad Z'_M = \frac{Z'_A + Z'_B}{2}$$

défini une position moyenne de l'objet dans le quadrilatère.

En posant $T = | X'_A - X'_B |$ on détermine la taille de l'objet.

L'ensemble des équations déterminées précédemment permet de calculer les trois paramètres X'_M , Z'_M et T à partir des traces laissées par un objet sur les deux capteurs linéaires.

Un objet traversant la barrière de surveillance est ainsi caractérisé à l'instant de l'acquisition des images ligne par trois paramètres: sa position M , de coordonnées (X'_M, Z'_M) et sa taille T .

Toutefois si plusieurs objets traversent le plan de surveillance en même temps il peut y avoir ambiguïté d'appariement. La figure 3-24 représente la configuration à caméras tête-bêche avec deux objets dans le plan de surveillance. Selon les mises en correspondance des pixels A_1, A_2, A_3 , et A_4 avec les pixels B_1, B_2, B_3 et B_4 on peut obtenir soit les quadrilatères W et X soit les quadrilatères Y et Z .

$$(A_1, B_1) \text{ et } (A_2, B_2) \rightarrow W \quad (A_3, B_3) \text{ et } (A_4, B_4) \rightarrow X$$

$$(A_1, B_3) \text{ et } (A_2, B_4) \rightarrow Z \quad (A_3, B_1) \text{ et } (A_4, B_2) \rightarrow Y$$

Il est très difficile de mettre en correspondance deux pixels se caractérisant uniquement par leur niveau de gris, en général identique dans notre cas. Il est cependant possible de lever l'indétermination en supposant que deux objets ne

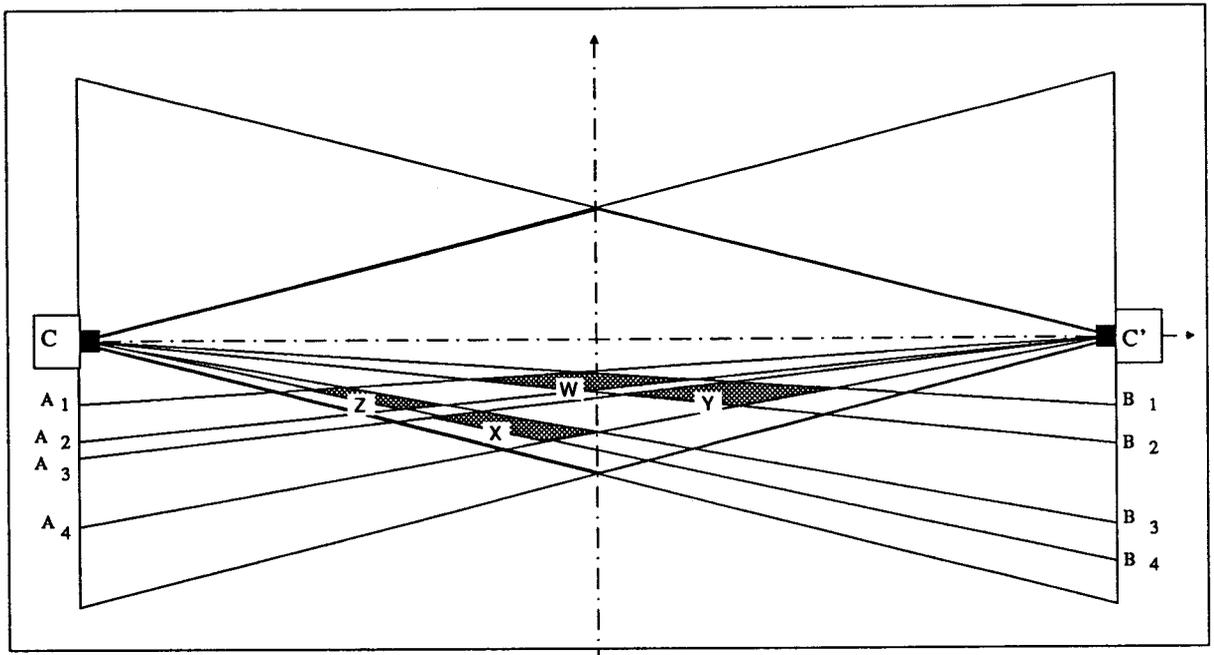


Fig. -3-24: Ambiguïté d'appariement.

peuvent pénétrer dans le plan de surveillance simultanément. Ceci nécessite un échantillonnage temporel de la scène rapide, ce qui est une hypothèse tout à fait raisonnable pour un système de vision linéaire. L'analyse de séquences d'images lignes successives recueillies simultanément par les deux caméras permet de suivre un objet dans le plan $O'x'z'$, dès son intrusion dans le plan de surveillance.

CHAPITRE IV

- ▣ DECLENCHEMENT D'ALARME PAR ANALYSE
DES SEQUENCES DE COUPLES D'IMAGES LIGNE**

I) INTRODUCTION

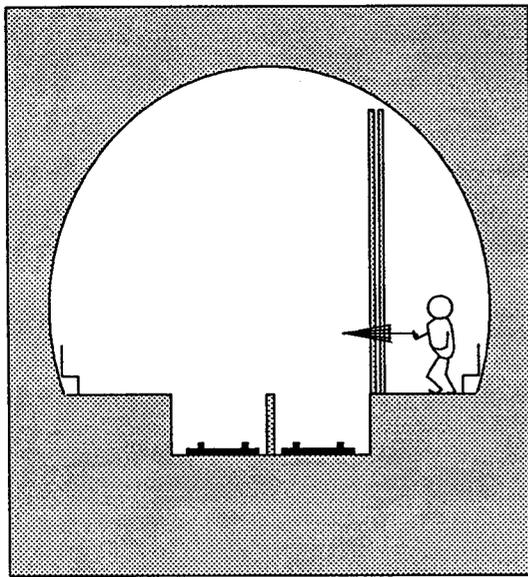
Dans le chapitre précédent nous avons détaillé le principe du système de détection. Nous avons également présenté les traitements dits de "bas niveau" concernant l'acquisition des images ligne et leur segmentation. A partir de la longueur de chacune des deux traces laissées par un objet présent dans le plan de surveillance sur chacune des deux images ligne, nous avons calculé, la position de cet objet par rapport à un repère associé aux deux caméras par triangulation optique. Nous en avons ensuite déduit, la taille T de l'objet traversant la barrière de surveillance.

Dans ce chapitre, nous considérons les traitements dits de "haut niveau". Ils concernent l'intégration d'informations temporelles obtenues par analyse de séquence d'images ligne. La présence momentanée d'un objet de taille importante où d'une personne au travers de la barrière optique ne signifie pas automatiquement qu'il y ait intrusion sur les voies du métro.

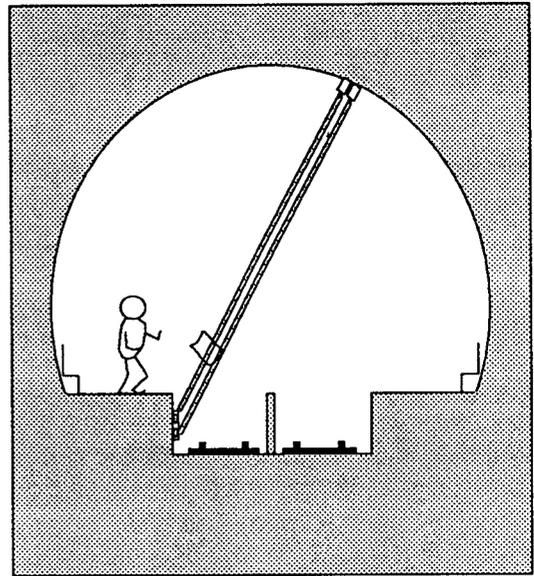
En effet, dans une configuration du type "caméras tête-bêche" disposées le long du nez de quai, une personne peut circuler au bord du quai, ou maintenir un objet au travers du plan de surveillance, sans toutefois créer une intrusion sur les voies (Fig. 4 - 1 - a).

Dans une configuration avec les caméras montées sous la voûte de la station et les rampes de tubes fluorescent sous le nez de quai, un journal tombant sur les voies et occultant le plan de surveillance sera perçu comme un objet de taille importante (Fig. 4 - 1 - b). Mais, en déterminant la vitesse à laquelle il traversera le plan, il sera facilement différencié d'une personne dont la chute serait beaucoup plus rapide. L'analyse de l'évolution temporelle des occultations de la barrière de surveillance peut donc apporter des informations supplémentaires très utiles pour l'interprétation des phénomènes observés par les caméras.

Pour accéder à ces informations, les deux plans élémentaires formant la barrière optique sont volontairement espacés de quelques centimètres. Cette modification entraîne un décalage temporel de l'occultation de ces plans élémentaires



a - pas d'intrusion sur les voies



b - intrusion d'un journal

Fig. -4-1-Exemples de passage

lors du passage d'un objet. Si un objet commence par couper le plan P1 à l'instant t_1 et le plan P2 à l'instant t_2 , la comparaison de ces deux instants permet de déterminer le sens de passage de l'objet. De plus, l'intervalle de temps entre chaque occultation, c'est à dire la valeur absolue $|t_1-t_2|$ de la différence entre les instants où l'objet commence à traverser chacun des plans, permet d'évaluer la vitesse de passage.

La figure 4 - 2 schématise des exemples d'évolution temporelle d'un objet à travers le double plan de surveillance. Chaque plan élémentaire est représenté par une barre verticale. Les instants t_n représentent les instants d'acquisition des deux images ligne. Sur l'exemple a de la figure 4 - 2 on voit un objet traversant complètement le double plan, alors que sur les exemples b et c l'objet, qui vient d'occulter le double plan, ne le traverse pas entièrement puisqu'il revient à sa position initiale.

Ces exemples du comportement d'un objet à travers le double plan de surveillance montrent la nécessité de développer une stratégie d'analyse des séries d'occultations. On débouche sur un concept de reconnaissance de séquences de positions successives d'un objet au travers de la barrière de surveillance. Chaque séquence est exprimée comme un agencement structuré de positions élémentaires. Il est clair que ce n'est pas le nombre relatif de positions qui va importer, mais les

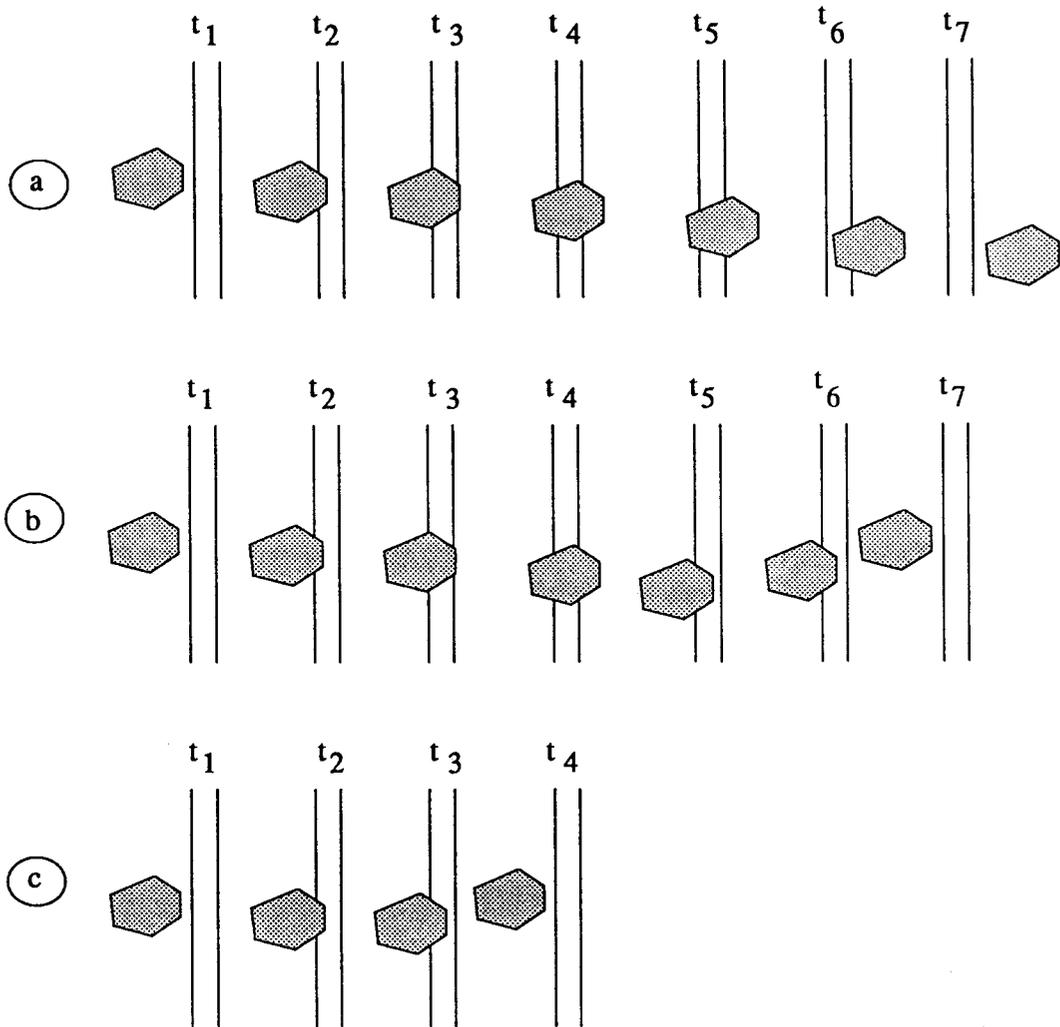


Fig-4-2- Exemples de l'évolution temporelle d'un objet

caractéristiques de leur enchaînement temporel. La reconnaissance se fait en cherchant à traduire la succession de positions dans un espace de représentation adapté. Pour cela nous avons choisi de décrire la structure générale d'une séquence par un automate à états finis.

Cet automate est complété par des traitements associés à chaque état. Le passage dans certains états déclenche sous certaines conditions les procédures d'alarme.

Si nous reprenons les exemples de la figure 4 - 2, dans le cas **a** le déclenchement de la procédure d'alarme est conditionné par la taille de l'objet, par contre en **b** et **c** aucune procédure d'alarme ne sera déclenchée quelque soit la taille de l'objet.

II) ANALYSE SPATIO-TEMPORELLE.

Les caméras linéaires permettent de faire des acquisitions d'images ligne en un temps relativement court (toutes les millisecondes avec le matériel utilisé), ce qui donne un échantillonnage temporel très rapide. A chaque acquisition d'un nouveau couple d'images ligne, on réitère les calculs définis au chapitre III.

Soit $I(t_n)$ avec $n = 1, 2, 3, \dots, N$, une séquence de couples d'images ligne acquises simultanément par les deux caméras aux instants t_1, t_2, \dots, t_N et soient $X(t_n), Z(t_n)$ et $T(t_n)$ les valeurs des trois paramètres définissant la position et la taille de l'intrusion à l'instant t_n .

La séquence $(X(t_n), Z(t_n))$, avec $n = 1, 2, 3 \dots, N$, indique le mouvement en deux dimensions de l'objet dans le plan de surveillance.

Dans le contexte de la détection d'intrusions, la barrière est placée de telle sorte que le mouvement des objets pouvant la traverser lui soit sensiblement perpendiculaire. En connaissant la distance Δy séparant les deux plans élémentaires P1 et P2 et en repérant les instants d'apparition et de disparition de la trace d'un objet dans les images ligne, il devient possible d'estimer la vitesse dans la direction perpendiculaire au double plan de surveillance. On utilise cette information sur la vitesse pour déterminer la profondeur de l'objet traversant le plan de surveillance dans la direction O_z .

La figure 4-3, illustre la manière dont un objet occulte successivement les deux plans de élémentaires P1 et P2 constituant la barrière de surveillance.

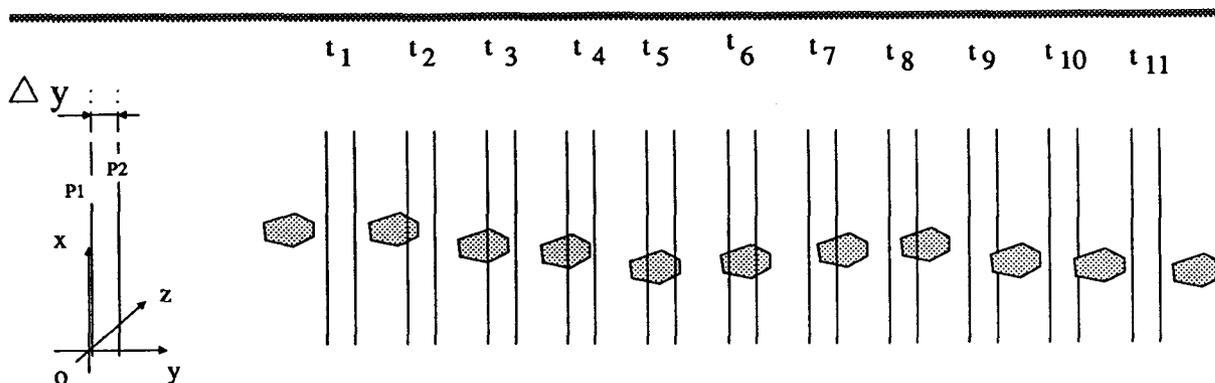


Figure-4-3-

La différence $(t_n - t_{n-1}) = T_p$ correspond à la période d'échantillonnage de la scène par les deux caméras. Cette période d'échantillonnage est déterminée par le temps de pose, identique pour chaque caméra, tous les calculs étant effectués en temps masqué, durant le temps de pose nécessaire à l'acquisition des deux images ligne suivantes. Dans l'exemple présenté, l'objet traverse le double plan en sept acquisitions de couples d'images. Connaissant la distance Δy séparant les deux plans et le temps de pose T_p , il est aisé d'estimer la vitesse de l'objet dans la direction perpendiculaire au plan de surveillance. En pratique, on définit une vitesse d'entrée V_e d'un objet dans la barrière de surveillance et une vitesse de sortie V_s de l'objet de la barrière de surveillance.

La vitesse d'entrée se calcule lorsque l'on a obtenu une séquence du type:

- aucune occultation,
- occultation d'un plan puis
- occultation des deux plans.

Dans l'exemple de la figure 4-3 cette séquence est obtenue entre les instants t_1 et t_5 . Connaissant Δy et T_p , on en déduit une valeur approchée de la vitesse d'entrée:

$$V_e = \frac{\Delta y}{(N+1) \cdot T_p} = \frac{\Delta y}{5 \cdot T_p} \text{ avec } N \text{ le nombre d'intervalles.}$$

La vitesse de sortie V_s se calcule lorsque l'on a obtenu une séquence du type:

- occultation des deux plans,

- occultation d'un plan puis
- aucune occultation.

Dans l'exemple de la figure 4-3 cette séquence est obtenue entre les instants t_6 et t_{11} , ce qui permet d'estimer V_s sous la forme:

$$V_s = \frac{\Delta y}{(N+1) \cdot T_p} = \frac{\Delta y}{6 \cdot T_p} \text{ avec } N \text{ le nombre d'intervalles.}$$

Il importe de noter qu'il existe une vitesse maximale au delà de laquelle un objet traversant la barrière ne peut être détecté. Cette vitesse maximale est fonction du temps de pose et du seuil de détection défini au chapitre précédent. La quantité de lumière E reçue par les éléments des capteurs linéaires est proportionnelle à leur durée d'exposition à la lumière des rampes lumineuse. Sans occultation, cette quantité de lumière est donc proportionnelle au temps de pose T_p . Soit

$$E = \int_{T_p} e \cdot dt = [e \cdot t]_{T_p} = e \cdot T_p .$$

Pour détecter le passage d'un objet, on seuille le niveau de gris de l'image ligne à la moitié de sa valeur maximale. Ceci signifie qu'on ne détectera le passage de l'objet que s'il occulte la rampe lumineuse durant au moins la moitié du temps de pose.

Cette remarque permet d'estimer la vitesse maximale dont peut être animé un objet pour être détecté:

$$V_{\max} = \frac{L}{T_p / 2} \text{ avec } L \text{ profondeur de l'objet.}$$

Par exemple pour un objet de 20 cm de profondeur et un temps de pose de 5ms, on obtient une vitesse de 80 m/s, soit 288 km/h. Au dessus de cette vitesse l'objet ne sera pas détecté.

D'autre part, le temps de pose des caméras limite la précision de mesure de la vitesse. Lorsque l'on passe directement d'un état des deux plans libre aux deux plans occultés par un objet on ne peut ni obtenir la vitesse de passage de l'objet avec

précision ni connaître le sens de passage. Dans ce cas, on ne dispose pas de toutes les informations nécessaires à la prise de décision. Cette remarque donne une limitation sur la vitesse de passage d'un objet plus sévère que la précédente.

On peut cependant affirmer que sa vitesse V est dépendante du rapport $\frac{\Delta y}{T_p}$

L'analyse du comportement de l'objet au travers de la barrière de surveillance peut être réalisé en suivant l'évolution temporelle des deux images ligne. En effet le comportement de l'objet se déduit de l'analyse des deux images ligne.

Toutefois, il est important de remarquer que, sur la séquence de la figure 4-3, on ne peut déterminer la position $X(t_n)$, $Z(t_n)$ de l'objet que lorsqu'il occulte les deux plans simultanément. Par contre, le suivi de l'objet au travers de la barrière de surveillance peut se faire dès qu'il est présent dans au moins un des plans.

Afin de suivre la progression de l'objet à travers la barrière, il convient de développer une stratégie permettant de repérer l'enchaînement des différentes positions illustrées sur la figure 4-3.

III) STRATEGIE.

III.1) Représentation de l'information.

Nous sommes en présence d'une séquence de paires d'images ligne dont on désire extraire une représentation de l'information aussi bien adaptée que possible à une exploitation immédiate et rapide, en vue de la surveillance de l'interface qui va en découler.

On choisit une représentation de l'information simple qui permet de décrire l'état des plans élémentaires P1 et P2. A chaque plan élémentaire est associée une variable binaire qui indique la présence ou l'absence d'un objet dans le plan. Si le plan P_i est traversé par un objet, la variable a_i correspondante est forcée au niveau 1, dans le cas contraire elle est laissée à 0.

$$a_i = \begin{cases} 0 & \text{si absence d'un objet dans le plan } P_i \\ 1 & \text{si présence d'un objet dans le plan } P_i \end{cases}$$

A chaque acquisition, l'état des deux variables a₁ et a₂ indique la présence ou l'absence d'un objet, respectivement dans les plans P1 et P2. On dispose ainsi, pour représenter l'information contenue dans une séquence d'images ligne, d'une suite de couples (a₁, a₂) qui caractérisent la position d'un objet par rapport aux deux plans P1 et P2. Cette représentation de l'information par des variables binaires apporte une réduction importante du volume de données à traiter.

La figure 4 - 4 indique les positions des plans P1 et P2 dans la configuration où les caméras sont montées tête-bêche.

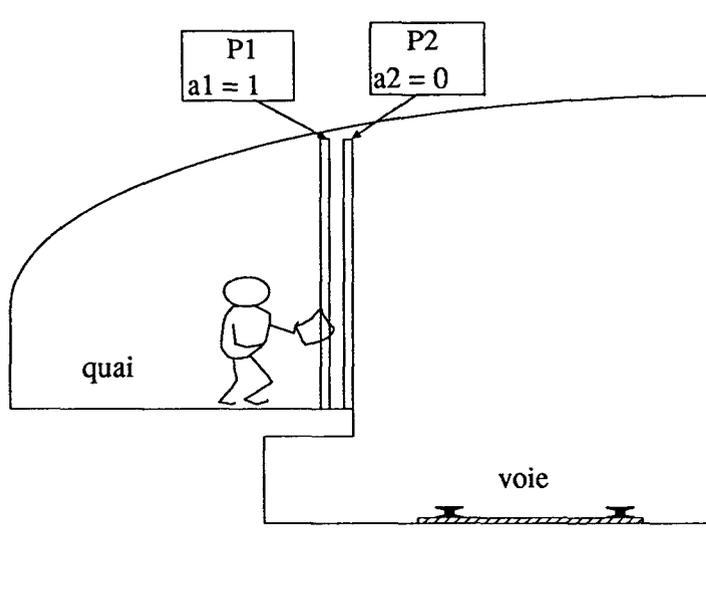


Fig. -4-4-Exemple de représentation des variables a_i .

III.2) Modélisation par un automate.

III.2.1) Définition générale d'un automate.

Les automates sont des outils mathématiques dont le rôle est d'engendrer des ensembles de phrases munies d'une structure commune [HOP 79]. Le principe de leur application à la reconnaissance structurelle est donc simple. Un automate à états fini consiste en un ensemble fini d'états et un ensemble de transitions d'un état à un autre (éventuellement le même) qui surviennent lors de l'entrée de symboles choisis dans un alphabet Σ . Il existe un état appelé état initial, généralement noté q_0 , et des états de sortie ou de "réussite".

Un automate est représenté par un graphe orienté appelé diagramme de transition, dont les sommets sont les états de l'automate et les arcs sont les transitions. Les arcs entre les noeuds sont étiquetés par des lettres de l'alphabet Σ .

L'élément de base d'un automate est donc la figure suivante.

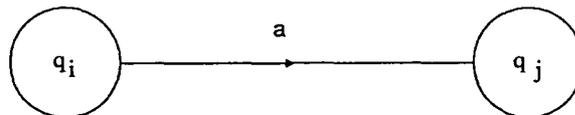


Fig-4-5-

Cela se traduit par: on peut aller de l'état q_i à l'état q_j par la lettre a de l'alphabet Σ .

Nous notons formellement un automate à états fini par un quintuplet $(Q, \Sigma, \delta, q_0, F)$:

Q est un ensemble fini d'états,

Σ est un alphabet fini de symboles d'entrée,

δ est la fonction de transition qui lie $Q \times \Sigma$ à Q , c'est à dire que $\delta(q, a)$ est l'état d'arrivée obtenu par la transition a à partir de l'état q ,

q_0 représente l'état initial et appartient à l'ensemble Q et

F est l'ensemble des états de sortie ou de "réussite" et est inclus dans Q .

On dit qu'un automate à états fini accepte une chaîne x si la séquence de transitions correspondant aux symboles de la chaîne x , partant de l'état initial aboutit à un état de réussite. Plus formellement, une chaîne x est acceptée par l'automate $M = (Q, \Sigma, \delta, q_0, F)$ si $\delta(q_0, x) = p$ tel que $p \in F$. Le langage accepté par M , désigné par $L(M)$, est l'ensemble $\{x \mid \delta(q_0, x) \in F\}$.

III.2.2) Reconnaissance de séquences de transitions par automate.

a) Définitions.

Nous avons pris pour l'application à la reconnaissance du comportement d'un objet dans le double plan de surveillance les mêmes notations que pour la définition générale d'un automate. Avant d'explicitier en détail le principe de la reconnaissance de séquence par l'automate à états fini, il est utile de définir le vocabulaire utilisé.

La **position** d'un objet par rapport aux deux plans P1 et P2 est définie par les variables a_1 et a_2 . L'ensemble des positions constitue l'ensemble fini d'états Q .

L'**alphabet** Σ est composé de l'ensemble des positions, c'est à dire des combinaisons possibles du couple de variables (a_1, a_2) . Ces variables représentent l'état des plans élémentaires P1 et P2.

On a donc:

$$\Sigma = \{ (0,0), (0,1), (1,1), (1,0) \}.$$

δ est la **fonction de transition**. Elle correspond à l'évolution de l'objet d'une position à une autre (éventuellement la même).

q_0 est l'**état initial** et représente l'absence d'objet dans le double plan de surveillance. Il appartient à l'ensemble Q .

On entend par **comportement** (c) une séquence de transitions entre positions successives qui reflète l'évolution d'un objet dans le double plan de surveillance. Le comportement est l'équivalent de la chaîne dans la définition générale d'un automate.

On définit de plus:

- un **comportement normal** (cn): séquence de transitions n'entraînant aucune procédure particulière.
- un **comportement anormal** (ca): séquence de transitions entraînant le déclenchement conditionnel d'une procédure d'alarme.

L'ensemble des états atteints lors d'un comportement anormal (ca) définit l'ensemble des états de sortie ou de "réussite" . On le note F . On aura donc:

$$\delta (q_0, ca) = p \text{ tel que } p \in F.$$

p représente donc un état de l'ensemble F des états entraînant un déclenchement conditionnel d'une alarme.

On obtient finalement l'automate de reconnaissance de comportement d'un objet ou d'une personne évoluant dans le double plan de surveillance et on le note:

$$R = (Q, \Sigma, \delta, q_0, F).$$

b) *Reconnaissance d'une intrusion.*

Reprenons l'exemple de la figure 4-3 qui représente le passage d'un objet dans le sens du quai vers la voie, et construisons l'automate dont le but sera de reconnaître ce comportement anormal. Cette intrusion est caractérisée par une séquence composée d'une entrée dans la barrière de surveillance par le plan P1, d'une présence simultanée dans les deux plans P1 et P2 suivi d'une sortie par le plan P2.

- La phase d'entrée dans le plan P1 est indiquée par la position (1,0) ou par une suite de positions (1,0) précédé d'une position (0,0).

- La phase de présence simultanée dans les deux plans est indiquée par une position (1,1) ou par une suite de positions (1,1).
- La phase de sortie par le plan P2 est indiquée par une position (0,1) ou par une suite de positions (0,1) précédée d'une position (1,1).

A partir de ces observations, on peut représenter cette intrusion par l'automate de la figure 4-6 .

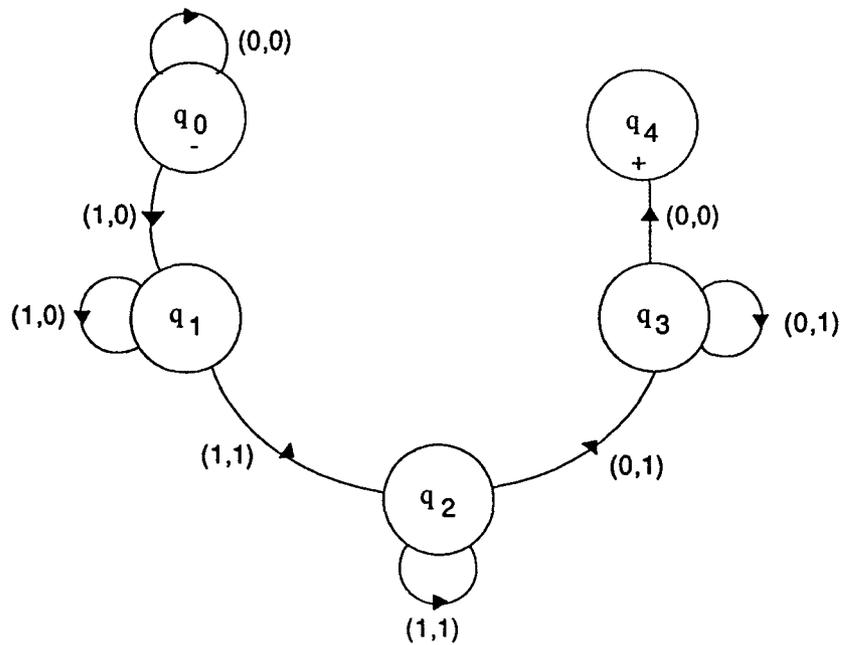


Fig. -4-6- Automate de reconnaissance d'intrusion .

Son interprétation est la suivante:

- l'état q_0 est l'état initial (noté par un signe -). On y reste tant qu'aucun objet n'entre par quelque coté que se soit dans la barrière de surveillance.
- l'état q_1 est celui qui indique l'entrée d'un objet dans la barrière de surveillance par le plan P1.
- l'état q_2 est celui qui indique la présence d'un objet dans les deux plans élémentaires P1 et P2.

d) l'état q_3 est celui qui indique la sortie d'un objet de la barrière de surveillance par le plan P2.

e) l'état q_4 indique la fin de passage. C'est l'état de réussite noté par un signe +. Dans cet état, on a reconnu un comportement anormal qui pourra déclencher une procédure d'alarme.

On obtient l'automate qui reconnaît la séquence " quai ---> voie ", noté $M = (Q, \Sigma, \delta, q_0, F)$ avec:

$$Q = \{q_0, q_1, q_2, q_3, q_4\} \text{ et } F = \{q_4\} .$$

Ce comportement anormal (ca) est décrit par la séquence minimum:

$$ca = (0, 0)(1, 0)(1, 1)(0, 1)(0, 0) \text{ tel que } \delta(q_0, ca) = q_4$$

$$\delta(q_0, (0, 0)) = q_0, \delta(q_0, (1, 0)) = q_1, \delta(q_1, (1, 1)) = q_2, \\ \delta(q_2, (0, 1)) = q_3 \text{ et } \delta(q_3, (0, 0)) = q_4$$

Lorsque l'on se trouve dans l'état d'alarme q_4 l'automate a reconnu une situation potentiellement dangereuse. Cependant, le déclenchement de la procédure d'alarme n'est pas envisagé uniquement en fonction de la réalisation de la séquence. Il dépend aussi des critères de taille et de position calculés directement sur les images ligne lors des traitements de bas niveau. Ces traitements ne sont pas effectués seulement lorsque l'on se trouve dans l'état d'alarme potentielle mais également lors de la progression de l'objet en mouvement à travers la barrière de surveillance. Cette remarque conduit à associer à chaque étape un traitement des informations recueillies en fonction de l'état correspondant. Lorsque l'on arrive dans l'état de réussite on possède ainsi tout les éléments nécessaire à la prise de décision.

Une limitation de l'automate à états fini tel que nous l'avons défini est que sa sortie est limitée à un signal binaire: "alarme" ou "normal". On définit alors la stratégie de décision par un automate à états fini avec sortie. Il existe deux approches distinctes, dans la première on associe la sortie avec l'état (Machine de Moore), dans la seconde on associe la sortie avec la transition (Machine de Mealy). Comme nous choisissons d'effectuer les traitements en fonction de l'état, on utilise la Machine de Moore. Notons que par rapport à un processus industriel utilisant le même type de

modèle, les traitements associées aux états de notre machine n'influent pas sur le comportement du processus, c'est à dire sur le comportement de l'objet ou de la personne. Les sorties de la machine ne font pas évoluer l'automate.

III.2.3) Automate à états fini avec sortie.

a) Définition générale d'une Machine de Moore.

Une Machine de Moore est un sextuplet $(Q, \Sigma, \Delta, \delta, \lambda, q_0)$ où Q, Σ, δ et q_0 ont les mêmes définitions que pour un automate à états fini. Δ représente les sorties et λ est la relation entre Q et Δ qui donne la sortie associée à chaque état. La sortie de l'automate en réponse aux entrées a_1, a_2, \dots, a_n , avec $n \geq 0$, est la séquence d'états $\lambda(q_0), \lambda(q_1), \dots, \lambda(q_n)$ où q_0, q_1, \dots, q_n telle que $\delta(q_{i-1}, a_i) = q_i$ pour $1 \leq i \leq n$.

b) Sorties de l'automate de reconnaissance d'intrusion.

On reprend l'automate M représenté par la figure 4-6 pour associer des sorties à chaque état. Ces sorties dépendent, pour la plupart, des traitements de bas niveau effectués au chapitre III. Après chaque acquisition des images ligne, c'est à dire à chaque période d'échantillonnage T_p , on obtient les variables d'entrée qui font évoluer l'automate d'un état à un autre (éventuellement le même).

Sur la figure 4-7 les traitements de sortie sont notés z_i et sont représentés à l'intérieur des cercles qui contiennent les indications des états q_i .

Les traitements sont:

z_0 : Initialisation du compteur d'acquisition N à 0 ($N \leftarrow 0$) et $S(t_n) = 0$.

z_1 : Incrémentation du compteur d'acquisition ($N \leftarrow N+1$).

z_2 : Calcul de la position $(X(t_n), Z(t_n))$, de la taille $T \leftarrow T(t_n)$ si $T(t_n) > T$, puis approximation de la vitesse d'entrée $V_e = \frac{\Delta y}{(N+1) \cdot T_p}$ de l'objet. Mémorisation de ces paramètres. Initialisation du compteur d'acquisition ($N \leftarrow 0$).

$$z_3 = z_1 .$$

z_4 : Approximation de la vitesse de sortie $V_s = \frac{\Delta y}{(N+1) \cdot T_p}$, puis comparaison des paramètres de vitesse et de taille respectivement à des seuils de vitesse et de taille. Ces seuils sont déterminés expérimentalement et permettent de différencier les objets légers tombant lentement et les petits objets des personnes.

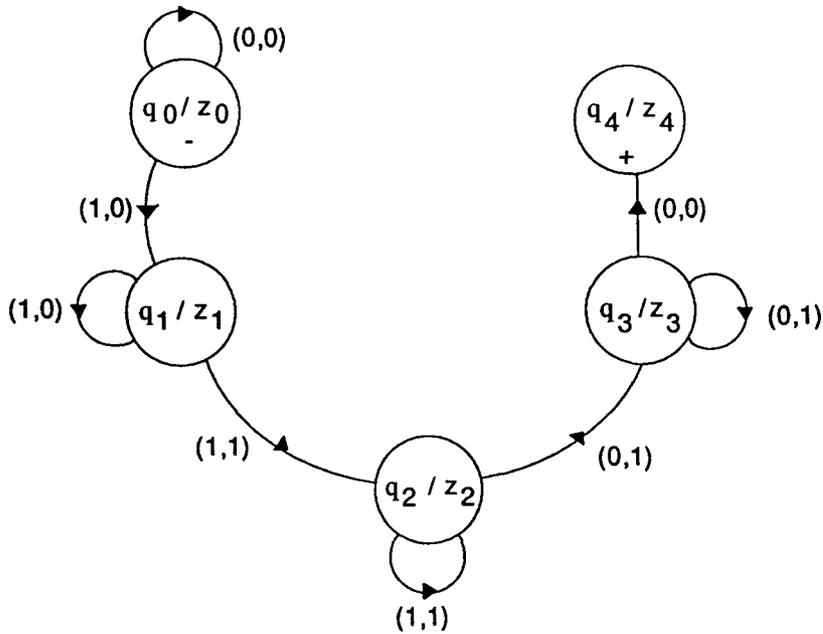


Fig. -4-7-Machine de Moore.

Lorsqu'on se trouve dans l'état q_4 , l'automate a reconnu un comportement anormal. La procédure d'alarme ne se déclenche que s'il y a dépassement des seuils. La conduite de la rame est alors fonction de la position de l'objet ayant pénétré sur les voies.

Après le déclenchement d'une procédure d'alarme, la remise en fonction du système de détection se fera manuellement. L'automate se positionnera alors dans l'état initial.

La reconnaissance du comportement d'un objet traversant la barrière de surveillance par un automate à états fini permet d'affiner les prises de décision au niveau du déclenchement des alarmes. Jusqu'ici nous n'avons étudié qu'un comportement anormal entraînant une procédure d'alarme conditionnelle. Nous allons maintenant nous intéresser aux autres comportements possibles d'objets ou de personnes à travers le double plan de surveillance.



III.3) Analyse de différents types de séquences.

Lors de la conception de l'automate M nous avons négligé le fait qu'une personne peut faire marche arrière ou rester présente dans les deux plans P1 et P2. Deux états intermédiaires q_5 et q_6 s'ajoutent alors à ceux décrits précédemment.

Ces états notés q_5 et q_6 diffèrent des états q_2 et q_1 par les traitements des sorties z_5 et z_6 associées. L'automate M devient celui de la figure 4 - 8 .

z_0 : Initialisation du compteur d'acquisition N à 0 ($N \leftarrow 0$) et de la taille ($S(t_n) \leftarrow 0$).

z_1 : Incrémentation du compteur d'acquisition ($N \leftarrow N+1$).

z_2 : Calcul de la position $(X(t_n), Z(t_n))$, de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$, puis approximation de la vitesse d'entrée $V_e = \frac{\Delta y}{(N+1) \cdot T_p}$ de l'objet. Mémorisation de ces paramètres. Initialisation du compteur d'acquisition ($N \leftarrow 0$).

$z_3 = z_1$.

z_4 : Approximation de la vitesse de sortie $V_s = \frac{\Delta y}{(N+1) \cdot T_p}$.
Déclenchement conditionnel d'une procédure d'alarme.

z_5 : Calcul de la position $(X(t_n), Z(t_n))$, de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$. Initialisation du compteur d'acquisition ($N \leftarrow 0$)

z_6 : Initialisation du compteur d'acquisition ($N \leftarrow 0$).

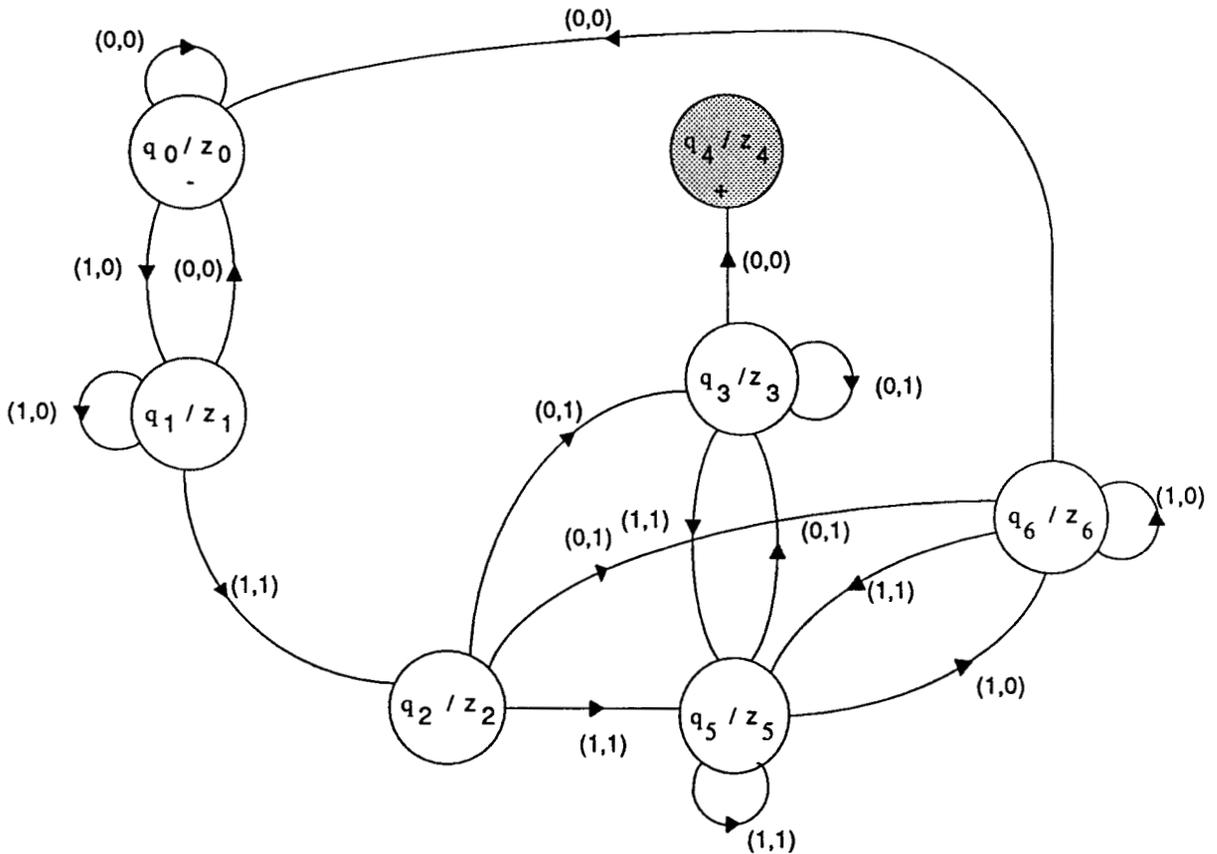


Fig. -4-8- Différents types de séquences.

La période d'échantillonnage, ou temps de pose, est un paramètre important qui conditionne l'analyse du passage d'un objet. Nous allons maintenant nous intéresser aux cas limites de détection.

III.4) Analyse des limites de la procédure de détection.

Nous avons analysé le fonctionnement du double plan de surveillance dans le cas idéal d'un objet de profondeur relativement importante par rapport à Δy et animé d'une vitesse suffisamment lente pour que son passage puisse se décomposer en:

- une phase d'entrée dans le double plan de surveillance
- une phase de présence dans le double plan de surveillance.
- une phase de sortie du double plan de surveillance.

En fait, selon la profondeur et la vitesse de l'objet, on peut se trouver en présence de séquences incomplète, dans lesquelles peuvent manquer l'une des trois phases précédentes.

III.4.1) Absence de phase d'entrée ou de sortie.

Si le temps de pose a une valeur élevée, on peut passer entre deux acquisitions successives, d'une situation où aucun plan n'est occulté à une situation où les deux plans sont occultés par un objet et inversement. Ces cas de figure correspondent à des séquences de transitions (0,0) (1,1) ou (1,1) (0,0) pour lesquelles les phases d'entrée et de sortie ne sont pas détectées. Pour obtenir de telles séquences il faut que la vitesse de passage de l'objet soit:

$$V > \frac{\Delta y}{T_p} .$$

Il y a alors indétermination du sens de passage. On ne peut savoir si l'objet a traversé la barrière des quais vers les voies ou inversement. Ce type de séquence est reconnu par l'automate de la figure 4 - 8 où l'état q_7 est l'état associé à ce comportement sans phase d'entrée ni de sortie.

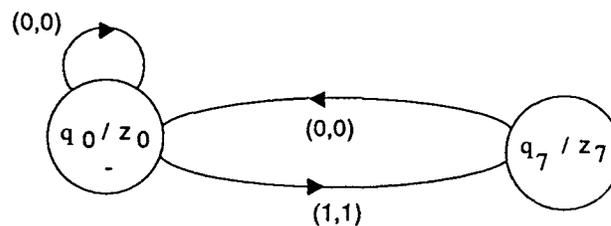


Fig-4-9-

Si $c = (0,0)(1,1)(0,0)$ alors:

$$\delta(q_0, (0,0)) = q_0 \quad \delta(q_0, (1,1)) = q_7 \quad \text{et} \quad \delta(q_7, (0,0)) = q_0 .$$

Les sorties sont:

z_0 : Initialisation du compteur d'acquisition N à 0 ($N \leftarrow 0$) et de la taille ($T(t_n) \leftarrow 0$).

z_7 : Calcul de la position $(X(t_n), Z(t_n))$, de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$ puis de la vitesse approchée $V = \frac{\Delta y}{T_p}$ de l'objet. Mémorisation de ces paramètres.

Dans le cas d'un tel comportement, on n'obtient pas l'information concernant le sens de passage de l'objet. Il est donc important que le temps de pose soit le plus faible possible pour limiter les passages dans l'état q_7 .

III.4.2) Absence de phase de présence.

D'autres séquences de transitions peuvent correspondre au schéma de la séquence minimale $c=(0,0)(1,0)(0,1)(0,0)$. Ce type de comportement peut être le fait, soit d'un objet passant rapidement soit d'un objet ayant une profondeur inférieure à la distance Δy séparant les deux plans élémentaires. Pour les objets de profondeur L , supérieure à Δy , la vitesse de passage devra être supérieure à $2 \frac{L - \Delta y}{T_p}$ pour qu'il n'y ait pas de phase de présence. Dans ces conditions un objet ne pourra être présent dans les deux plans élémentaires durant le temps minimum nécessaire à sa détection, c'est à dire $T_p/2$.

L'automate qui reconnaît le type de séquence sans phase de présence dans les deux plans de surveillance est donné figure 4 - 10 .

Si $c = (0,0)(1,0)(0,1)(0,0)$ alors:

$$\delta(q_0, (0, 0)) = q_0, \delta(q_0, (1, 0)) = q_1, \delta(q_1, (0, 1)) = q_{10} \text{ et} \\ \delta(q_{10}, (0, 0)) = q_{11}.$$

Les sorties $\lambda(q_0)\lambda(q_1)\lambda(q_{10})\lambda(q_{11})$ obtenues avec cette séquence, ne donnent aucune information sur la position et la taille de l'objet. Les sorties sont:

z_0 : Initialisation du compteur d'acquisition N à 0 ($N \leftarrow 0$) et de la taille ($T(t_n) \leftarrow 0$).

z_1 : Incrémentation du compteur d'acquisition ($N \leftarrow N+1$).

z_8 : Initialisation du compteur d'acquisition ($N \leftarrow 0$).

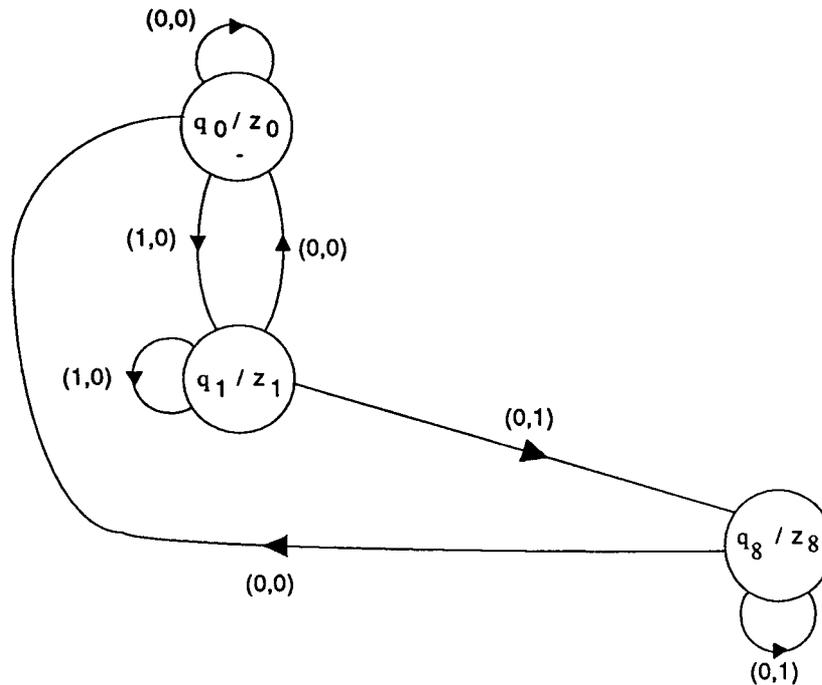


Fig. -4-10-

L'étude de ces cas limites, qui ne considère pas les objets de profondeur inférieure Δy , supposés sans danger pour la conduite de la rame, indique la limite supérieure de la vitesse au delà de laquelle il n'est pas possible d'obtenir les paramètres nécessaire à une prise de décision fiable, $V_{lim} = 2 \frac{L - \Delta y}{T_p}$.

Si on prend un objet de profondeur $L = 10$ cm, avec $\Delta y = 5$ cm et un temps de pose $T_p = 5$ ms on obtient comme vitesse limite 36 km/h. Cette vitesse correspond effectivement à la vitesse maximale que peut atteindre un coureur de haut niveau, ce qui assure la sécurité des passagers.

Remarque: Nous n' avons pas considéré les séquences de transitions dues au passage dans le sens voie vers le quai. Une personne quittant la voie pour regagner le quai constitue un comportement anormal. Ces séquences de transitions sont fort improbables dans une station souterraine s'il n'y a pas eu d'intrusion préalable. Dans le cas d'une station à l'air libre, des objets ou animaux volants peuvent, par exemple, traverser la barrière à contre sens. Un automate plus complet prenant en compte ces phénomènes, peut être développé selon les mêmes principes.

III.5) Automate final.

L'automate représenté par la figure 4 - 8 reflète tout les comportements possibles de l'évolution d'un objet ou d'une personne à travers la barrière de surveillance.

Les sorties de l'automate sont:

z_0 : Initialisation du compteur d'acquisition N à 0 ($N \leftarrow 0$) et de la taille ($T(t_n) \leftarrow 0$).

z_1 : Incrémentation du compteur d'acquisition ($N \leftarrow N+1$).

z_2 : Calcul de la position ($X(t_n), Z(t_n)$), de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$, puis approximation de la vitesse d'entrée $V_e = \frac{\Delta y}{(N+1) \cdot T_p}$ de l'objet. Mémorisation de ces paramètres. Initialisation du compteur d'acquisition ($N \leftarrow 0$).

$z_3 = z_1$.

z_4 : Approximation de la vitesse de sortie $V_s = \frac{\Delta y}{(N+1) \cdot T_p}$. Déclenchement conditionnel d'une procédure d'alarme.

z_5 : Calcul de la position ($X(t_n), Z(t_n)$), de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$. Initialisation du compteur d'acquisition ($N \leftarrow 0$)

z_6 : Initialisation du compteur d'acquisition ($N \leftarrow 0$).

z_7 : Calcul de la position ($X(t_n), Z(t_n)$), de la taille T avec, $T(t_n) \leftarrow T$ si $T > T(t_n)$ puis de la vitesse approchée $V = \frac{\Delta y}{T_p}$ de l'objet. Mémorisation de ces paramètres.

z_8 : Initialisation du compteur d'acquisition ($N \leftarrow 0$).

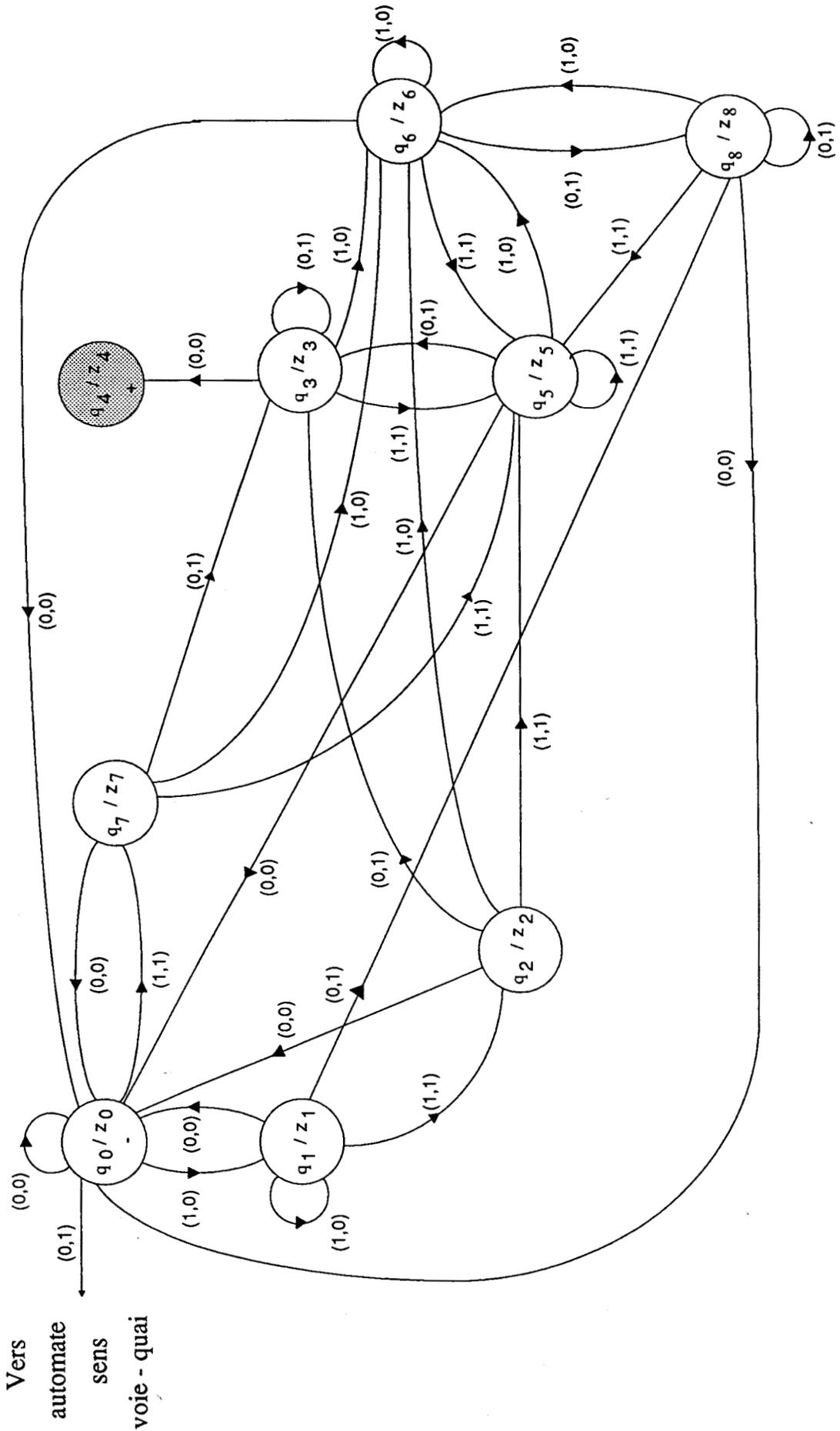


Fig - 4 - 11 - Automate de reconnaissance de séquences.

En conclusion, on peut dire que le temps de pose T_p et l'écartement Δy des deux plans de surveillance constituent le couple de paramètres ajustables le plus important. C'est lui qui conditionne le succès de la reconnaissance de séquences de transitions par l'automate. Un temps de pose faible garantit la détection du passage des objets rapides. De plus, la vitesse limite à laquelle les objets pourront être détectés est d'autant plus élevée que Δy est grand. Cependant l'écartement Δy doit rester relativement faible si on désire détecter des objets de petite taille. On se trouve donc face à un choix qui doit aboutir à un compromis entre la limite inférieure de la taille des objets détectés et la limite supérieure de leur vitesse.

Le chapitre suivant concerne la description matérielle du système de vision linéaire utilisé et de la technique logicielle mise en oeuvre. Nous présentons également une carte réalisée spécifiquement pour cette application afin de respecter les contraintes de la vision stéréoscopique linéaire. D'autre part la contrainte sur le temps de pose sera constamment présente lors de la conception logicielle. Enfin des essais effectués au laboratoire concluront le chapitre.

CHAPITRE V

▣ IMPLANTATION TEMPS REEL ET RESULTATS EXPERIMENTAUX

I) INTRODUCTION.

Dans ce chapitre nous exposons les problèmes d'implantation liés aux contraintes de temps d'exécution des procédures. Celles ci imposent une architecture matérielle spécifique sur laquelle les algorithmes pourront travailler en temps réel. D'autre part la réalisation d'une carte d'acquisitions simultanées d'images ligne s'est avérée nécessaire pour répondre aux besoins des techniques de stéréovision utilisées pour obtenir la position des objets.

Des simulations sur une maquette réalisée en laboratoire sont présentées à la fin de ce chapitre pour démontrer les performances du système de détection proposé.

II) IMPLANTATION.

II.1) Aspect matériel.

II.1.1) Machine de vision linéaire

Nous avons choisi le système de vision linéaire Visionix ,de Digital Design, pour l'implantation des procédures de détection d'intrusions. Ce choix a été motivé par les grandes possibilités de traitement offertes par l'architecture informatique de ce système qui s'organise autour du bus VME. La configuration de base comporte:

a) *Une unité centrale: CPU.*

Elle est constituée d'un processeur MC68000 de Motorola et d'une interface maître du bus VME.

b) Une unité caméra ou interface intelligente caméra: carte IICAM.

La carte IICAM est équipée du même processeur MC68000 et d'une interface esclave sur le bus VME. Les performances de ce système de vision sont obtenues en grande partie grâce à cette interface "intelligente". Cette unité possède quatre principales fonctions:

- Le contrôle de la caméra, c'est à dire l'envoi des commandes et réception des données.

- La mémorisation des images ligne, qui s'effectue dans deux zones mémoires fonctionnant en bascule, l'une est en mode lecture pendant que l'autre est en mode écriture. On peut ainsi effectuer un traitement sur l'image précédente pendant l'acquisition de l'image suivante.

- Les traitements câblés des images, à savoir une opération de seuillage et une opération de soustraction avec une image de référence,

- Le traitement logiciel par le microprocesseur MC68000.

c) une unité de visualisation.

Cette unité permet la visualisation des images ligne acquises par la caméra.

La caméra CYCLOPE du système de vision, est pourvue d'une barrette de 2048 pixels et se connecte sur l'interface caméra IICAM et sur l'unité de visualisation.

II.1.2) Configuration de l'application.

Cette configuration n'est pas figée. En effet, il est possible, en fonction des spécificités d'une application d'ajouter des cartes sur le bus VME. La détection d'intrusions par vision linéaire binoculaire nécessite l'utilisation d'images saisies simultanément par deux caméras. Une possibilité consisterait à utiliser le Multiplexeur conçu par la société Digital Design qui permet de faire l'acquisition des données à partir d'une caméra choisie parmi quatre caméras connectées au système.

Cependant, l'acquisition reste séquentielle ce qui est en contradiction avec les contraintes de la vision stéréoscopique. En effet, la vision stéréoscopique nécessite le traitement d'images acquises simultanément. Une autre possibilité peut être envisagée en utilisant autant d'interface caméra que de caméras. Cette architecture présenterait l'avantage de disposer d'un microprocesseur par caméra, ce qui offrirait une plus grande possibilité de traitements logiciels.

Cependant pour des raisons d'ordre économique, cette éventualité ne sera pas exploitée. Pour cette raison nous avons développé au laboratoire une carte d'acquisitions simultanées pour quatre caméras linéaires qui répond ainsi aux besoins de la vision stéréoscopique. Cette carte, que nous appelons SIAM (Simultaneous Input Acquisition Module), se connecte, d'une part sur le bus VME et d'autre part sur la carte IICAM par un bus de données "image". Nous avons aussi pour les mêmes raisons d'ordre économiques dédié la carte SIAM aux caméras linéaires Optronix distribuée par la société Micronix. Le capteur linéaire de ces caméras de marque Thompson, possède 1728 pixels. La configuration matérielle du système utilisé est représentée par le synoptique de la figure 5- 1 .

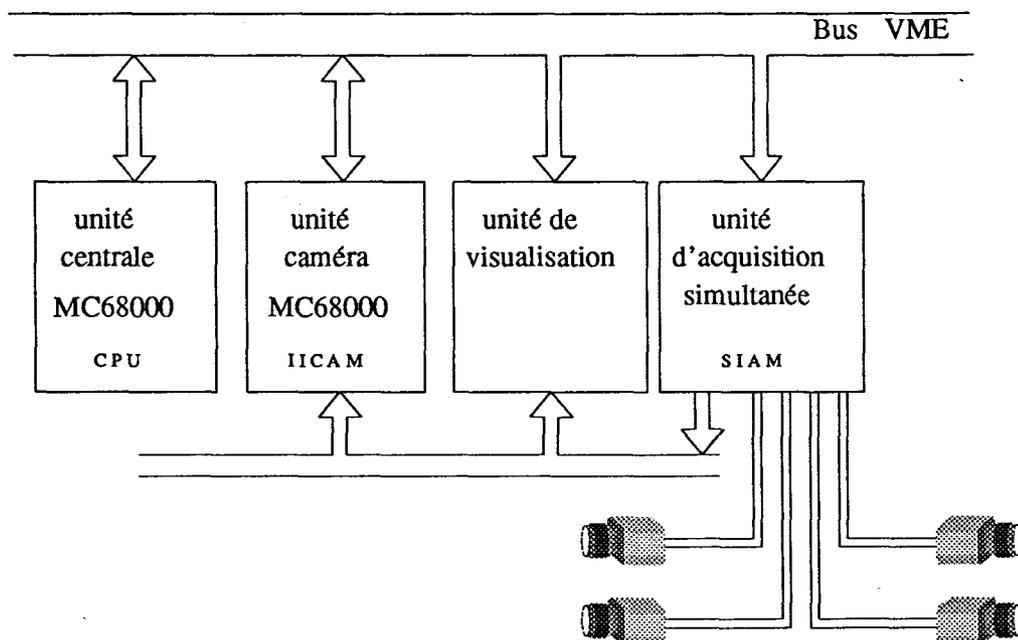


Fig-5-1- Architecture matérielle du système de vision

II.1.3) Unité d'acquisition simultanée d'images ligne: carte SIAM.

Cette unité permet d'interfacer plusieurs caméras Micronix sur le système de vision Visionix. Cette unité est réalisée sur une carte au format Double Europe comprenant une interface esclave sur le bus VME. Les connexions des caméras à la carte SIAM s'effectue par des connecteurs dédiés à chaque caméra qui transmettent les signaux de commandes et de données. Un bus "image", véhiculant les mêmes signaux, relie la carte SIAM à la carte caméra et à l'unité de visualisation.

Le schéma synoptique de la carte est représenté sur la figure 5 - 2 . La photo 1 donne l'implantation des composants sur la carte SIAM.

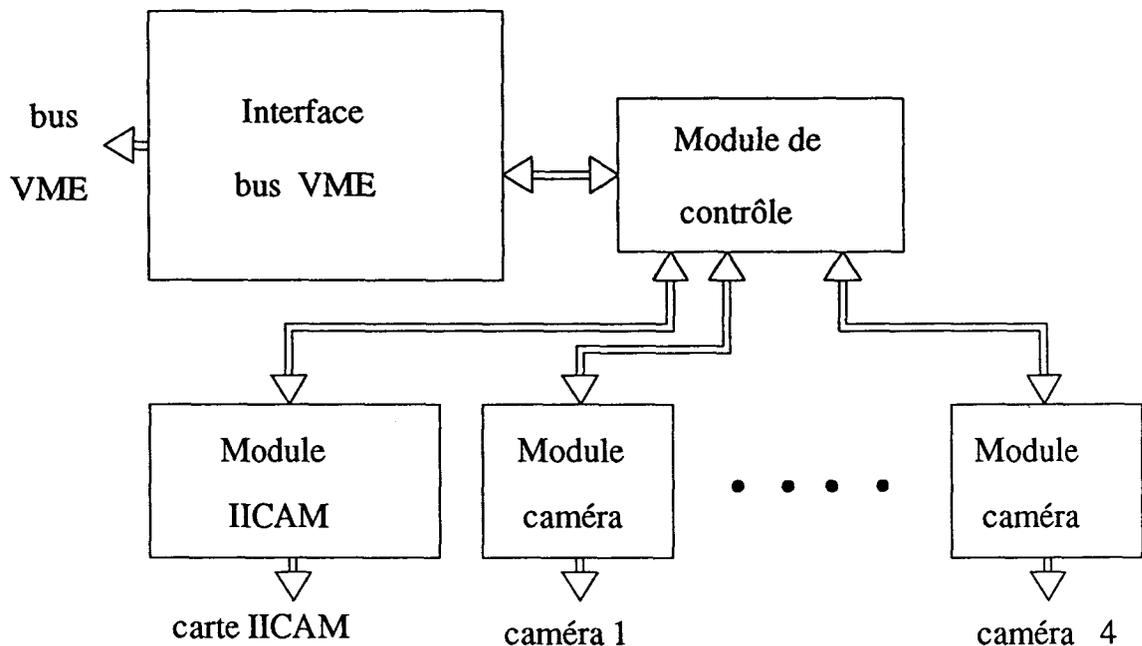


Fig. -5-2- Synoptique : carte d'acquisition simultanée

Chaque module caméra s'articule autour d'un composant EPLD 1810. Chacun de ces composants gère tous les signaux de commandes de la caméra ainsi que les signaux internes du module qui réalise aussi l'interface avec la caméra Micronix.

Le module de contrôle de la carte gère le fonctionnement global de la carte. Le contrôle de la carte, s'effectue par l'intermédiaire de deux registres. Le premier, accessible en écriture, est le registre de commande qui permet de configurer la carte

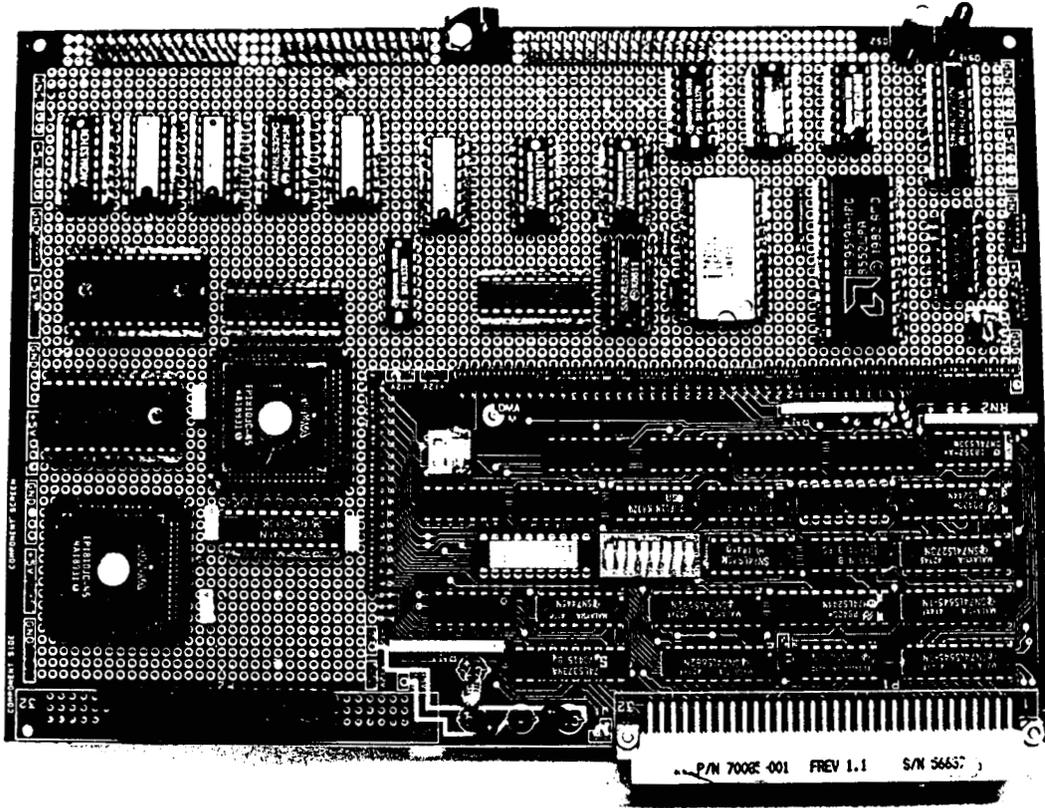


PHOTO-1- Photo de la carte SIAM

et de choisir le type d'opération à effectuer. Le deuxième, accessible en lecture, est le registre d'état.

Le registre de commande de la carte permet de configurer l'unité d'acquisition selon l'un des quatre modes de fonctionnement suivants:

- mode transparent.
- mode acquisitions simultanées.
- mode restitution.
- mode lecture par le bus VME.

a) Mode transparent.

Dans ce mode de fonctionnement la carte restitue directement vers la carte IICAM les données issues d'une des caméras. Il permet la visualisation des images ligne acquises par l'une des caméras. Ce mode est nécessaire pour la mise au point optique de chaque caméra.

b) Mode acquisitions simultanées.

Ce mode de fonctionnement constitue la principale fonction de la carte. Il permet de faire l'acquisition, en mémoire image de plusieurs images ligne acquises simultanément par plusieurs caméras. Chaque module de contrôle caméra possède sa propre mémoire image.

c) Mode restitution.

Le mode restitution effectue la restitution séquentielle de chaque image stockée en mémoire vers la carte IICAM dont on peut ainsi utiliser toute les ressources logicielles et matérielles. La restitution s'effectue par une demande d'acquisition émise par la carte IICAM vers la carte SIAM.

d) Mode lecture par le bus VME.

Ce mode sert uniquement à vérifier le bon fonctionnement de la carte en testant une caméra à la fois. La lecture, par le bus VME, de la mémoire image du module caméra sélectionné, s'effectue toujours à la même adresse. Cette adresse s'incrémente automatiquement à chaque lecture d'une donnée.

II.2) Aspects logiciels.

II.2.1) Contrainte temps réel.

La solution idéale pour effectuer l'ensemble des traitements nécessaires à la détection des intrusions en temps réel consiste à réaliser les calculs en temps masqué. Cette technique revient à faire les calculs associés à un couple d'images ligne acquises simultanément pendant le temps d'intégration nécessaire à l'acquisition du couple d'images ligne suivante. Si on échantillonne la scène avec une période d'échantillonnage T_p , le suivi temps réel des objets présents dans la scène consiste à traiter les données des caméras entre deux prises d'images successives. Cette procédure implique, soit une limitation du temps de calcul à la période d'échantillonnage, soit une augmentation de la période d'échantillonnage. Or la période d'échantillonnage de la scène, c'est à dire le temps de pose T_p des caméras, doit être suffisamment rapide pour détecter tous les objets susceptibles de traverser la barrière de surveillance.

Nous avons mentionné au chapitre précédent que plus le temps de pose est faible, meilleure est la détection. Il s'agit donc de trouver une solution matérielle et logicielle qui permette de diminuer le plus possible les temps de traitements afin que le temps de pose soit le plus faible possible.

Le temps d'exécution des procédures de traitement est lié à la configuration matérielle du système. Une solution avec un processeur unique n'apportant pas des performances de traitement suffisantes, nous avons été conduits à choisir une solution multiprocesseur, utilisant ainsi au mieux les possibilités matérielles du système.

II.2.2) Traitement multiprocesseur.

Le fait que les calculs soient divisibles en plusieurs parties indépendantes avec un flux limité de communications entre modules de traitement, nous permet d'envisager un traitement multiprocesseur. Le travail simultané de deux processeurs permet en principe, d'effectuer deux fois plus de traitements:

- le premier processeur, sur l'unité caméra (carte IICAM), effectue la partie acquisition des données, détection de contours, mise en forme des données et mémorisation. Le logiciel est écrit en assembleur MC68000.

- le deuxième processeur, sur l'unité centrale (carte CPU), pilote l'automate décrit dans le chapitre précédent. Les traitements en sortie de l'automate ainsi que le déclenchement d'alarmes sont également réalisés par cette unité centrale. Le logiciel est écrit en langage C.

Les traitements s'effectuent au fur et à mesure que les images ligne sont acquises par les caméras. Il est impératif que les procédures effectuées sur chaque processeur soient inférieures au temps de pose. Ces procédures ont une durée variable selon les données acquises par les caméras, l'état de l'automate et les traitements à exécuter. Le respect de cette contrainte passe par le choix d'un temps de pose supérieur au temps de calcul supposé le plus long.

Les acquisitions des images sont déclenchées toutes les 5ms par un "séquenceur" présent sur la carte d'acquisition simultanée. Le temps moyen de l'exécution de la procédure la plus longue qui est confiée à la CPU est de l'ordre de 3,5 ms. Mais ce temps peut se trouver accru selon l'état dans lequel se trouve l'automate. Les procédures implantées sur la carte IICAM demandent un temps de traitement constant égale à 1,5 ms auquel il faut néanmoins ajouter le temps de transfert des images ligne provenant de la carte SIAM, soit environ 1ms. Avec un temps de pose de 5 ms on dispose donc d'une marge de sécurité importante.

Le principe de communication des données inter processeurs se fait par l'intermédiaire d'une zone mémoire adressable par chaque processeur. Cette zone mémoire se trouve sur la carte IICAM. La figure 5-3 représente le synoptique de fonctionnement des procédures sur chaque processeur en fonction du temps.

T_a représente le temps d'acquisition des images des caméras 1 et 2.

$$T_a = (\text{Nbre pixels}) * (\text{Horloge pixel}) = (1728 * 0,55 \mu s) = 0,96 \text{ m s}$$

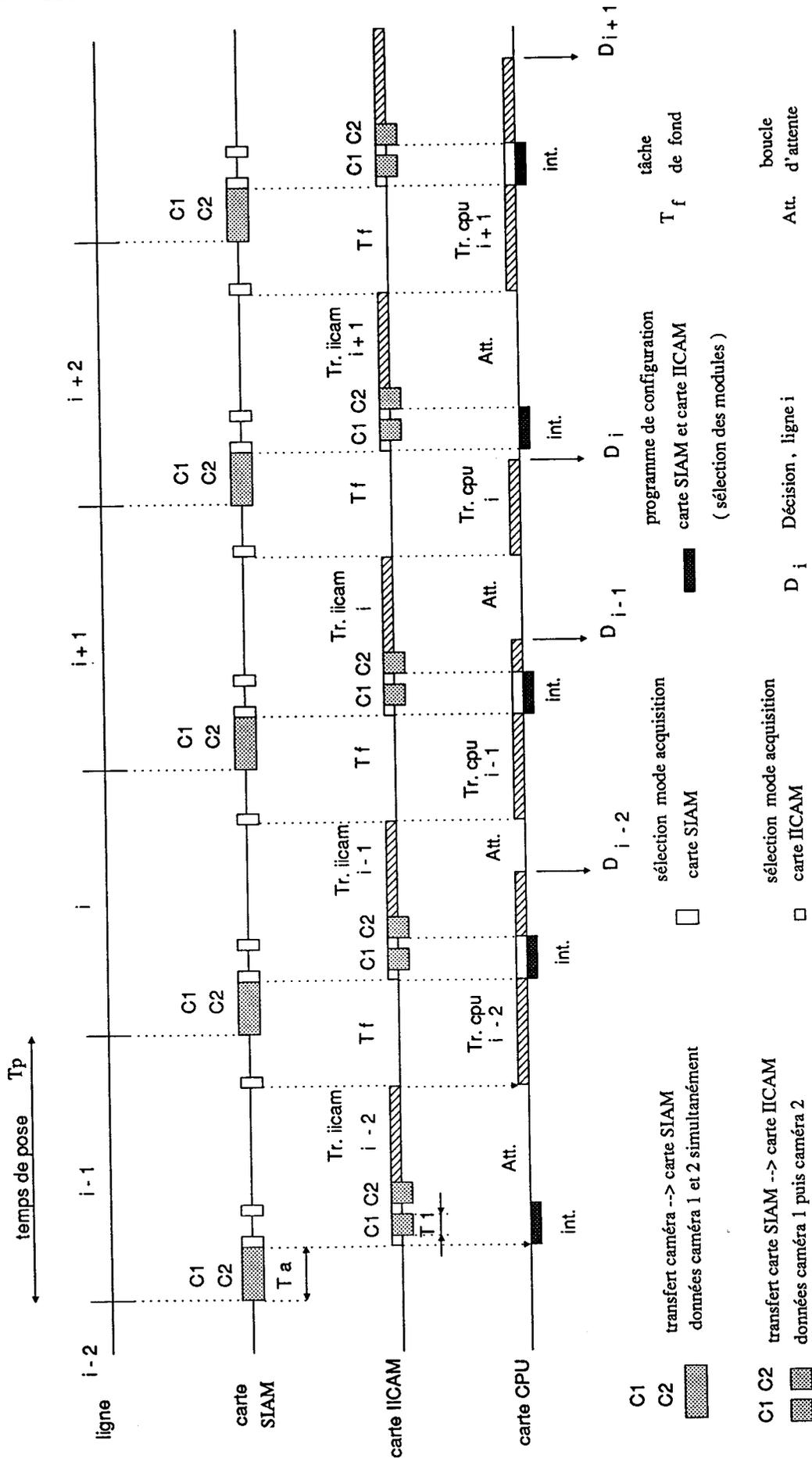


fig. - 5 - 3 Diagramme temporel des traitements

T 1 représente le temps de transfert des données de la caméra 1 de l'unité d'acquisition simultanée vers l'unité IICAM. Ce transfert se fait à la fréquence de 4 MHz.

$$T_1 = 1728 * \frac{1}{4.10^{-6}} = 0,432 \text{ ms}$$

Le traitement correspondant à un couple d'images ligne sur la carte CPU (Tr. cpu) ne démarre qu'après la fin d'exécution du traitement de ces mêmes images sur la carte IICAM (Tr. iicam), mais il peut par contre se terminer pendant le traitement (Tr. iicam) des images ligne suivantes. Un programme d'interruption implanté sur la carte CPU (int.) permet de synchroniser les cartes d'acquisition simultanée et IICAM pour le transfert des données. L'unité d'acquisition simultanée permet de gérer des interruptions. Une interruption, générée à chaque fin d'acquisition d'images ligne, lance le programme d'interruption (int.) de la carte CPU. Le programme présent sur la carte IICAM est activé par le programme d'interruption de la carte CPU.

La validation du procédé de détection, des traitements logiciels et du fonctionnement de la carte SIAM passe par une phase d'expérimentation exposée au paragraphe suivant.

III) RESULTATS EXPERIMENTAUX.

Nous exposons, dans cette partie, les résultats obtenus, avec notre méthode de détection d'intrusions, pour différents essais. Pour démontrer l'efficacité et les performances de la méthode, nous utilisons la machine de vision multiprocesseur, décrite précédemment, équipée de deux caméras linéaires, de 1728 pixels chacune. La plate-forme expérimentale montée au Centre d'Automatique de Lille correspond à la configuration avec les caméras et les rampes lumineuses montées tête-bêche. Chaque caméra équipée d'un objectif de distance focale 50mm, est placée à 2,00m de la rampe lumineuse visée. Celle-ci est constituée d'un tube fluorescent de 0,80m utile, alimenté en haute fréquence (20 khz). Cette alimentation permet de rendre

l'acquisition des images ligne indépendante de la fréquence du secteur. Les deux plans élémentaires P1 et P2 définis chacun par la caméra et la rampe lumineuse visée, sont séparés de $\Delta y = 4 \text{ cm}$. Le dessin de la figure 5 - 4 représente la géométrie de la plate-forme expérimentale avec le repère de l'espace tridimensionnel associé.

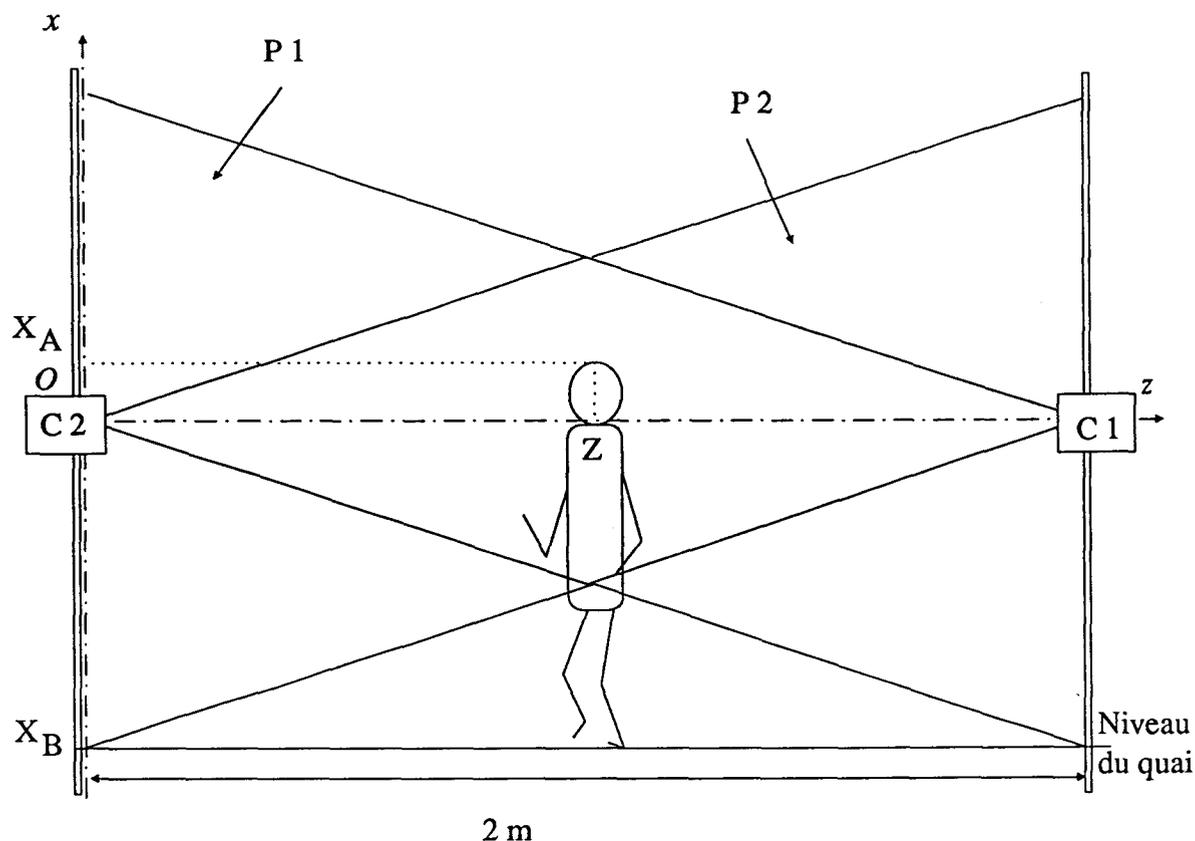


Fig-5-4- Dessin de la plate-forme expérimentale

Les essais réalisés, avec cette plate-forme de dimension réduite par rapport à une installation en vraie grandeur, ont été effectués avec un objet représentant la silhouette d'un personnage. Compte tenu des limitations de la plate-forme expérimentale les dimensions de cette silhouette sont de 32 cm de haut sur environ 6 cm de large. L'axe optique des caméras est situé à 30 cm au dessus du bas des champs visualisés. Pour les essais, nous avons simulés, avec notre silhouette, le passage d'un piéton à travers la barrière de surveillance. Le bas de la silhouette est donc situé, durant le mouvement à l'ordonnée correspondant au bas des champs visualisés, c'est à dire, au niveau du quai. Dans ces conditions, on ne détecte que la

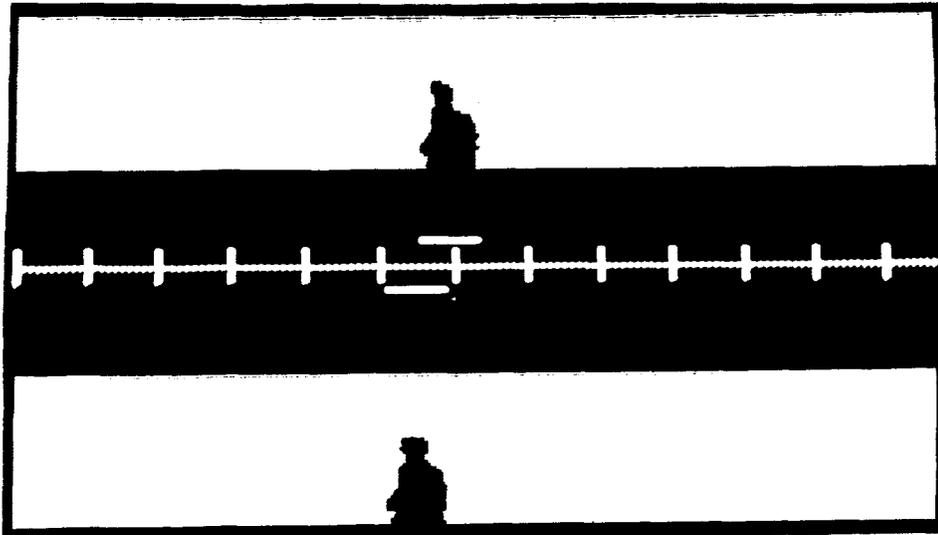
partie haute de la silhouette. Le calcul de l'ordonnée extrême X_A est suffisant pour déterminer la taille de notre silhouette.

Chacune des figures suivantes représente une séquence de 256 couples d'images ligne consécutives. La séquence en dessous des graduations correspond aux images ligne de la caméra du plan P1 et la séquence au dessus des graduations correspond aux images ligne de la caméra du plan P2. La longueur des traits blancs, en dessous et au dessus des graduations, représente le nombre d'acquisition où la silhouette est présente dans le plan de surveillance correspondant. Pour permettre l'affichage des images ligne sur le moniteur vidéo nous avons été contraints d'utiliser un temps de pose de 20 ms. Lorsque nous n'effectuons pas cet affichage, nous pouvons utiliser un temps de pose de 5ms.

a) *Effet de la vitesse de passage.*

Dans un premier temps, nous évaluons les performances de l'algorithme de suivi par automate sur 3 séquences de couples d'images ligne pour différentes vitesses de passage. Afin d'obtenir des images de taille comparable sur les caméras, on passe la silhouette au centre de la barrière de surveillance de telle sorte que $Z=1$ m.

La première séquence présentée (séquence 1), correspond le passage de notre silhouette à travers les deux plans de surveillance. Il s'agit d'une intrusion sur les voies devant conduire à un déclenchement conditionnel d'une alarme. Le tableau 1 donne les résultats de la reconstruction de l'objet en fonction de l'état de l'automate. X_A représente l'ordonnée maximale de la silhouette par rapport à l'axe optique des caméras, X_B l'ordonnée minimale étant fixe est égale à -30 . La taille $T= |X_A - X_B|$ de notre silhouette est donc égale à $[4 - (-30)]$ soit 34 cm au lieu de 32 cm réel. Cette erreur dans la détermination de la taille T correspond à moins de 10% de la taille totale. Par contre, la position longitudinale Z de la silhouette est donné avec une très bonne précision .



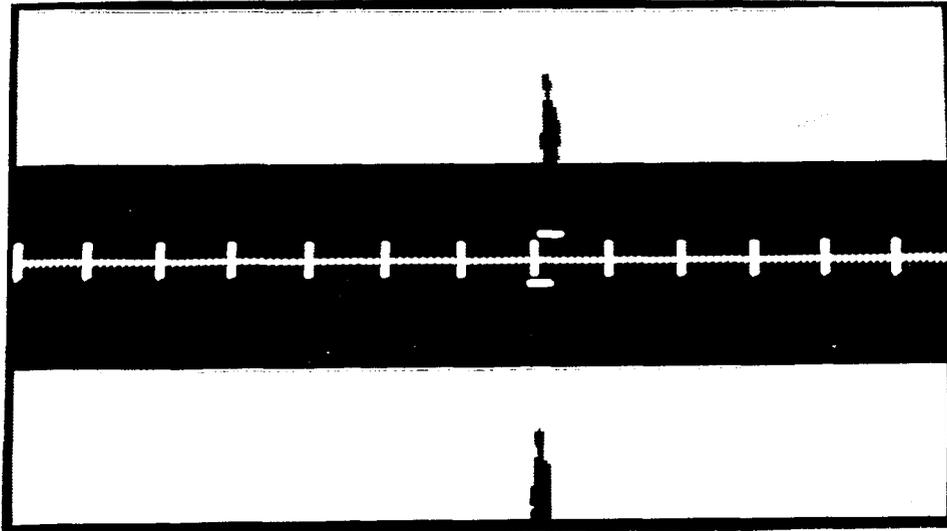
Séquence 1

état de l'automate	N	X _A	X _B	Z	Ve	Vs
q0	0	--	--	--	--	--
q1	1	--	--	--	--	--
q1	2	--	--	--	--	--
:	:		--			
q1	8	--	--	--	--	--
q1	9	--	--	--	--	--
q2	0	4,0 cm	-30,0 cm	100 cm	0,2 m/s	--
q5	0	4,0 cm	:	100 cm	:	--
:	:	:	:	:	:	--
q5	0	4,0 cm	:	100 cm	:	--
q3	1		:		:	--
q3	2		:		:	--
:	:		:		:	--
q3	8		:		:	--
q3	9		:		:	--
q4					:	0,2 m/s
Paramètres utilisés pour le déclenchement de l'alarme	Taille T		Distance Z		Vitesse moyenne	
	34,0 cm		100 cm		0,2 m/s	

Tableau 1: Suivi de la silhouette dans une séquence d'images.

La deuxième séquence proposée (séquence 2), représente le passage de la même silhouette à travers les deux plans de surveillance avec une vitesse plus élevée que la précédente. Le tableau 2 donne les résultats en fonction de l'état de

l'automate. On constate que seules les vitesses d'entrée et de sortie diffèrent du tableau 1.



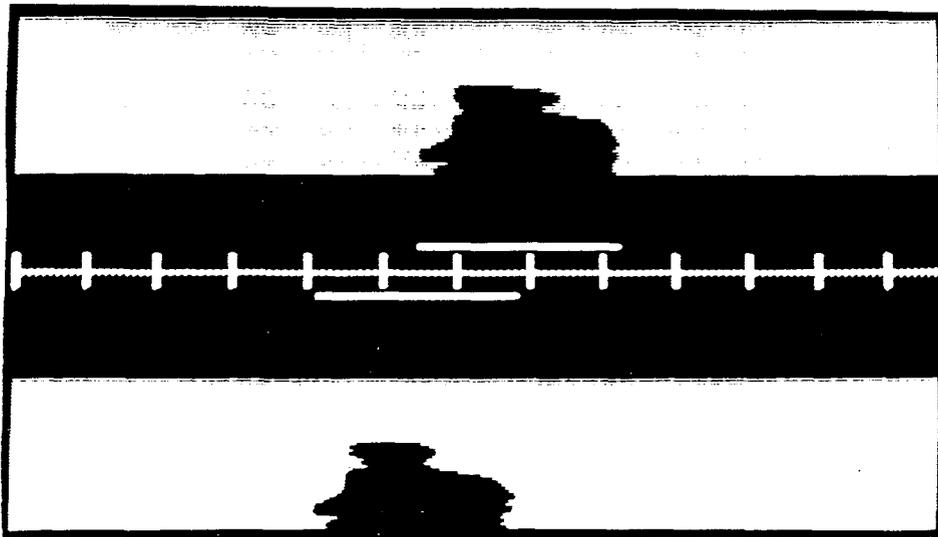
Séquence 2

état de l'automate	N	X _A	X _B	Z	V _e	V _s
q ₀	0	--	--	--	--	--
q ₁	1	--	--	--	--	--
q ₁	2	--	--	--	--	--
q ₁	3	--	--	--	--	--
q ₁	4	--	--	--	--	--
q ₂	0	4,0 cm	-30,0 cm	100 cm	0,4 m/s	--
q ₅	0	4,0 cm	:	100 cm	:	--
:		:	:	:	:	--
q ₅	0	4,0 cm	:	100 cm	:	--
q ₃	1		:		:	--
q ₃	2		:		:	--
q ₃	3		:		:	--
q ₄						0,5 m/s
Paramètres utilisés pour le déclenchement de l'alarme	Taille T		Distance Z		Vitesse moyenne	
	34,0 cm		100 cm		0,45 m/s	

Tableau 2: intrusion de la silhouette à vitesse rapide.

La troisième séquence d'images lignes (séquence 3) représente toujours le même type d'intrusion, seule la vitesse de passage change. On voit sur les séquences 1, 2 et 3 que les formes de la silhouette sur les images respectives sont fonction de sa vitesse de passage.

Les différences de forme des images de la silhouette sur les séquences 1,2 et 3 résultent de la fréquence d'acquisition des images lignes. Le nombre d'images lignes ou la silhouette est présente est d'autant plus grande que la traversée des plans de surveillance est lente. Ce nombre se matérialise par la longueur du trait blanc du plan correspondant.



Séquence 3

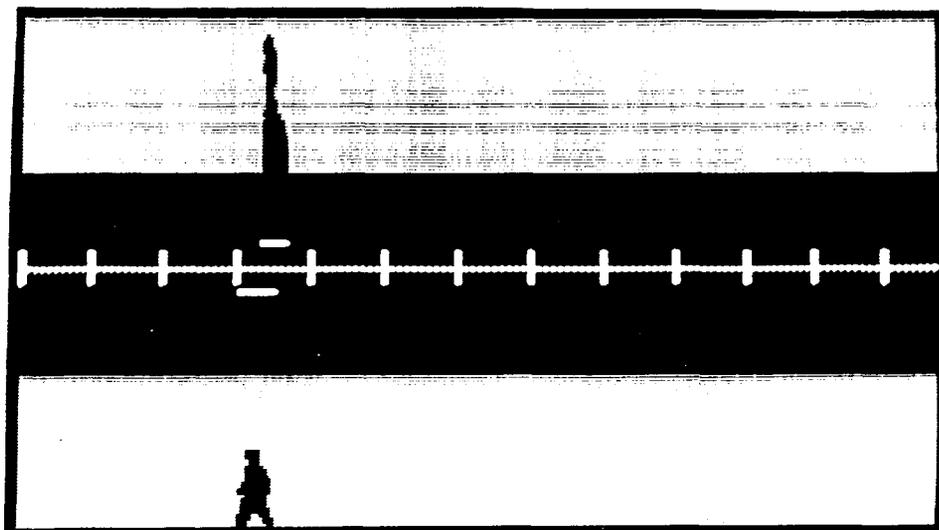
état de l'automate	N	X _A	X _B	Z	V _e	V _s
q0	0	--	--	--	--	--
q1	1	--	--	--	--	--
:	:					
q1	28	--	--	--	--	--
q2	0	4,0 cm	-30,0 cm	100 cm	0,06m/s	--
q5	0	4,0 cm	:	100 cm	:	--
:	:	:	:	:	:	--
q5	0	4,0 cm	:	100 cm	:	--
q3	1	:	:	:	:	--
:	:	:	:	:	:	--
q3	28	:	:	:	:	--
q4						0,06m/s
Paramètres utilisés pour le déclenchement de l'alarme	Taille T		Distance Z		Vitesse moyenne	
	34,0 cm		100 cm		0,06m/s	

Tableau 3: intrusion de la silhouette à vitesse lente.

b) *Effet de la position longitudinale.*

Dans un deuxième temps nous vérifions l'exactitude du calcul de la position de l'intrusion le long du quai pour deux séquences d'images ligne.

La première séquence (séquence 4) montre le passage de la silhouette près de la caméra relative au plan P2. La distance réelle est de Z=40 cm dans le repère associé au montage. Le tableau 4 donne les résultats du suivi de l'objet à travers le double plan de surveillance.



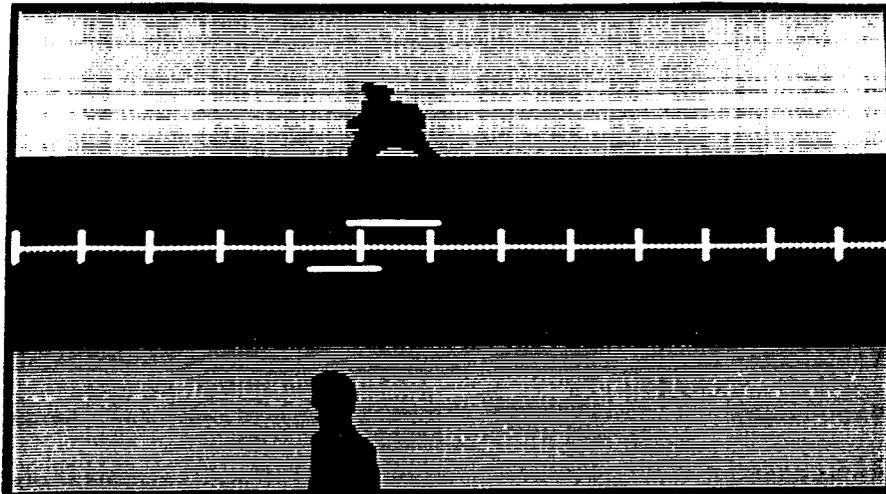
Séquence 4

état de l'automate	N	X _A	X _B	Z	V _e	V _s
q0	0	--	--	--	--	--
q1	1	--	--	--	--	--
q1	2	--	--	--	--	--
:	:	--	--	--	--	--
q1	6	--	--	--	--	--
q2	0	3,2cm	-30,0 cm	40 cm	0,28 m/s	--
q5	0	3,2cm	:	40 cm	:	--
:	:	:	:	:	:	--
q5	0	3,2cm	:	40 cm	:	--
q3	1	:	:	:	:	--
q3	2	:	:	:	:	--
q3	3	:	:	:	:	--
q4					:	0,5 m/s
Paramètres utilisés pour le déclenchement de l'alarme	Taille T		Distance Z		Vitesse moyenne	
	33,2cm		40cm		0,39 m/s	

Tableau 4: intrusion - distance 40 cm.

On aperçoit sur cette séquence la différence de hauteur de chacune des formes de la silhouette sur les images ligne. Cette différence est à la base du calcul de la position de l'objet.

La deuxième séquence (séquence 5) est quant à elle l'inverse de la séquence précédente. Elle correspond au passage de la silhouette près de la caméra relative au plan P1. La distance réelle mesurée est de $Z=160$ cm. La distance calculée est 170 cm.

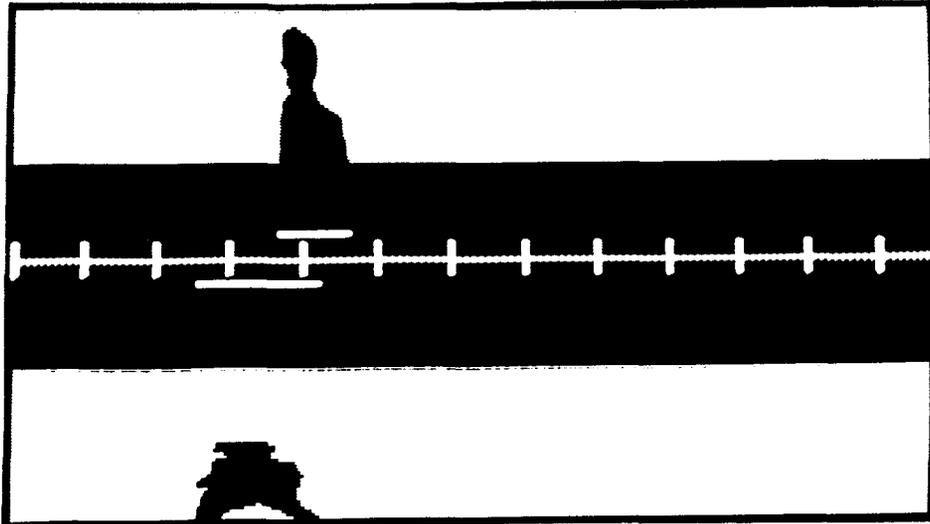


Séquence 5

état de l'automate	N	X _A	X _B	Z	V _e	V _s
q0	0	--	--	--	--	--
q1	1	--	--	--	--	--
:	:	--	--	--	--	--
q1	11	--	--	--	--	--
q2	0	3,4cm	-30,0 cm	170 cm	0,16m/s	--
q5	0	3,4cm	:	170 cm	:	--
:	:	:	:	:	:	--
q5	0	3,4cm	:	170 cm	:	--
q3	1	:	:	:	:	--
:	:	:	:	:	:	--
q3	17	:	:	:	:	--
q4					:	0,11m/s
Paramètres utilisés pour le déclenchement de l'alarme	Taille T		Distance Z		Vitesse moyenne	
	33,4cm		170 cm		0,135m/s	

Tableau 5: intrusion - distance 160 cm.

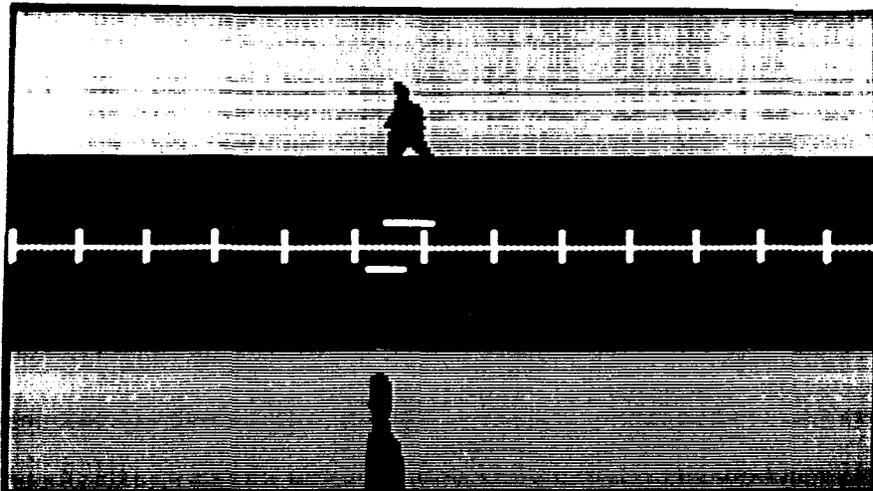
D'autres essais ont été effectués pour différentes vitesses de passage et à différentes positions. Les séquences 6,7,8 sont le reflet de ces essais. Pour la séquence 8 nous avons pris une silhouette de taille T =22 cm.



Séquence 6

Paramètres utilisés pour le déclenchement de l'alarme	Taille T	Distance Z	Vitesse moyenne
	33,4cm	34cm	0,12m/s

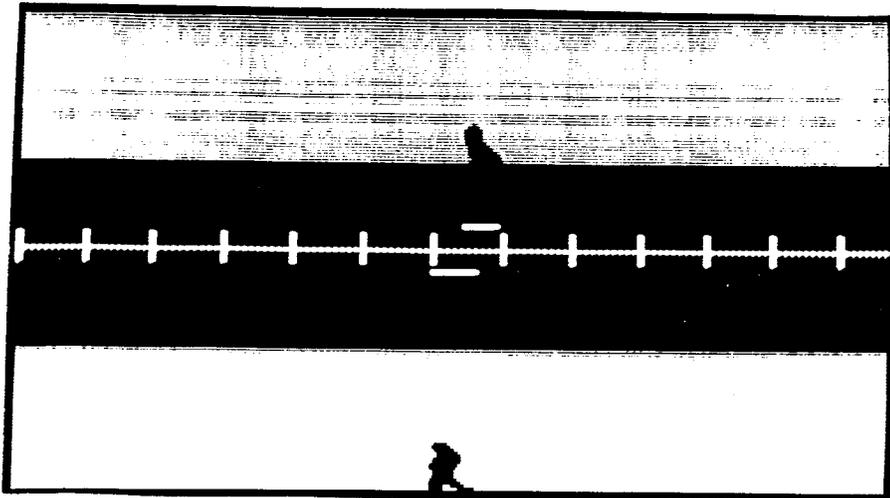
Tableau 6: intrusion - distance 40 cm .



Séquence 7

Paramètres utilisés pour le déclenchement de l'alarme	Taille T	Distance Z	Vitesse moyenne
	31,5 cm	170 cm	0,26m/s

Tableau 7: intrusion - Distance 160 cm .

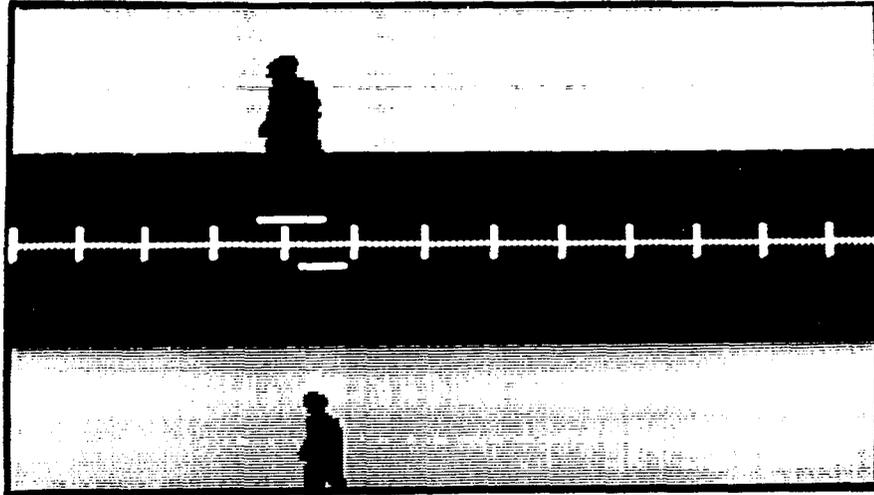


Séquence 8

Paramètres utilisés pour le déclenchement de l'alarme	Taille T	Distance Z	Vitesse moyenne
	24 cm	40 cm	0,23m/s

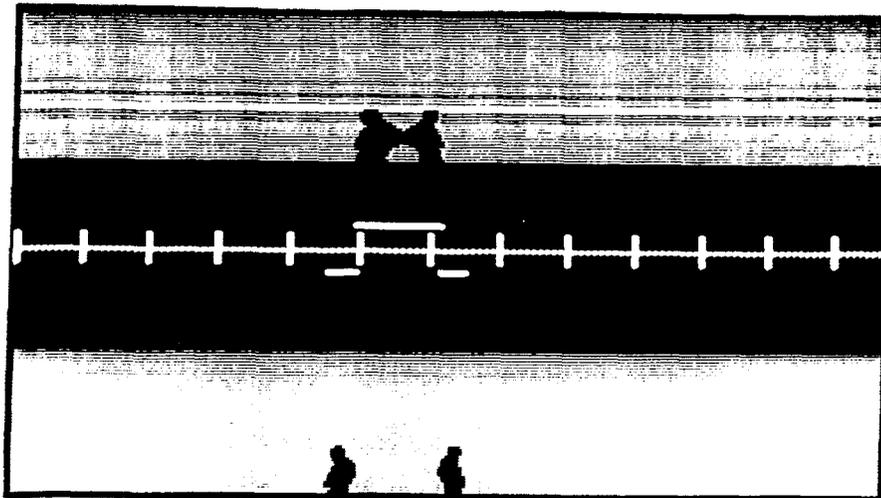
Tableau 8: intrusion - Taille 22 cm- Distance 40 cm.

La séquence d'images ligne (séquence 9) représente le passage de notre silhouette dans le sens de la voie vers le quai. L'image de la silhouette apparaît d'abord par le plan P2, ce qui est tout à fait conforme au résultat attendu. Le calcul de la vitesse et de la position pour certain état des plans P1 et P2 est identique au cas du passage d'un objet dans le sens du quai vers la voie.



Séquence 9

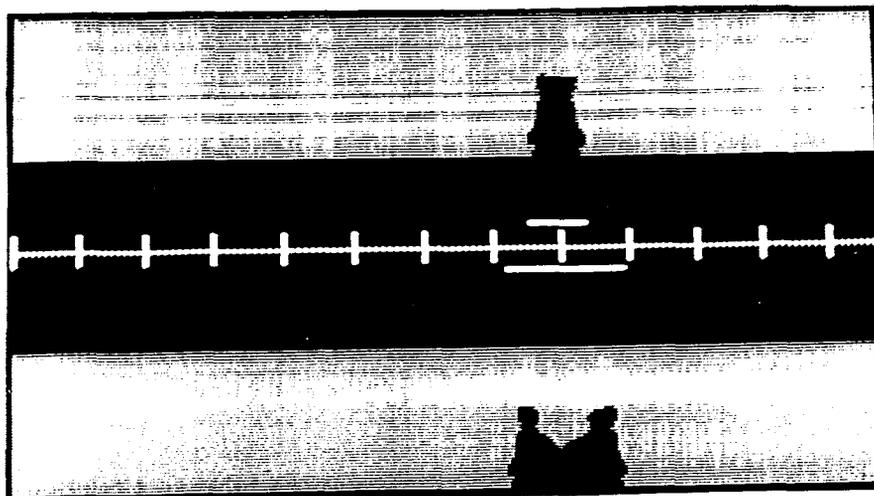
Enfin les dernières séquences (séquences 10 et 11) d'images ligne proposées correspondent à un aller retour de la silhouette. Le tableau 4 donne le résultat des calculs effectués lors de son passage en fonction des états de l'automate pour la première séquence (séquence 10). Pour cette séquence nous avons pris une silhouette de taille plus petite.



Séquence 10

état de l'automate	N	X _A	X _B	Z	Ve	Vs
q0	0	--	--	--	--	--
q1	1	--	--	--	--	--
q1	2	--	--	--	--	--
:	:		--			
q1	7	--	--	--	--	--
q1	8	--	--	--	--	--
q2	0	-7,7 cm	-30,0 cm	110 cm	0,22m/s	--
q5	0	-7,7 cm	:	110 cm	:	--
q3	1	:	:	:	:	--
q3	2		:		:	--
:			:		:	--
q3	22		:		:	--
q5	0	-7,7 cm	:	110 cm	:	--
q5	0		:		:	--
q6	0		:		:	--
Paramètres calculés	Taille T			Distance Z	Vitesse moyenne	
	22,7 cm			110 cm	0,22 m/s	

Tableau 9: silhouette effectuant un aller retour.



Séquence 11

Paramètres calculés	Taille T	Distance Z	Vitesse moyenne
	32 cm	100 cm	0,2 m/s

Tableau 10: silhouette effectuant un aller retour.

La silhouette traversant le double plan de surveillance a toujours été détectée quelle que soit sa vitesse et sa position. Un simple critère de taille suffit pour distinguer les petits objets des êtres humains.

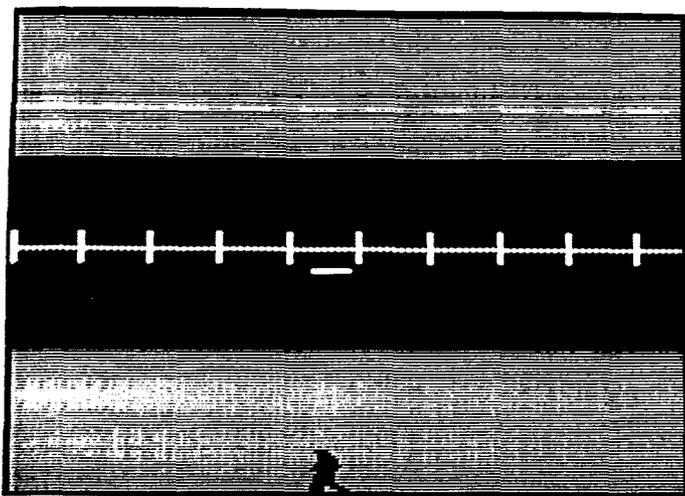
Dans cette thèse nous avons proposé une solution simplifiée à la détermination des paramètres de mouvement d'objet dans une scène à trois dimensions. L'utilisation de caméras linéaires à la place des traditionnelles caméras matricielles permet une détermination temps réel de la position, de la taille et du mouvement d'un objet. Cette nouvelle méthode semble être très attractive pour la détection des intrusions sur les voies d'un métro.

CONCLUSION GENERALE

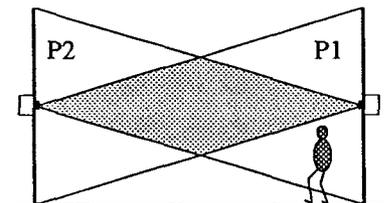
Dans cette thèse, nous avons abordé le problème de la détection d'intrusions sur les voies d'un métro. La difficulté était de concevoir un dispositif capable de différencier des intrusions parasites sans importance, des chutes de passagers sur les voies. Pour ce faire, et compte tenu des aspects temps réel nécessaires aux déclenchement des alarmes, nous avons opté pour un système de stéréovision à base de caméras linéaires. Une barrière constituée de deux plans de surveillance parallèles permet de déterminer, en temps réel, c'est à dire à la vitesse d'acquisition des images, les paramètres géométriques qui définissent l'intrusion et d'obtenir une approximation de la vitesse de pénétration de l'intrus dans la zone surveillée.

Ces paramètres permettent, en analysant les séquences de paires d'images ligne, de différencier les différents types d'intrusions. Une plate forme expérimentale, a été installée pour valider la méthode. Elle a nécessité la réalisation d'une carte spécialisée permettant l'acquisition simultanée de quatre images ligne provenant de quatre caméras.

Cette réalisation permet d'envisager la mise au point d'une barrière plus complète équipée de quatre caméras, afin d'augmenter la taille de la zone surveillée. En effet, avec deux caméras, le champ de détection utile consiste en l'intersection des deux champs triangulaires P1 et P2 (cf chapitre IV). La partie utile du double plan de surveillance obtenu à partir de ces deux caméras se réduit à un losange. Il n'est pas possible de calculer les paramètres de vitesse, de position et de taille si l'objet passe à travers la barrière de surveillance sans couper ce losange (séquence 1).

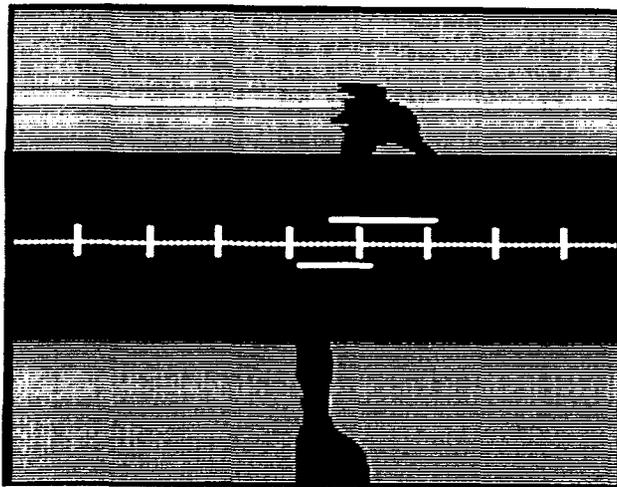


Séquence 1

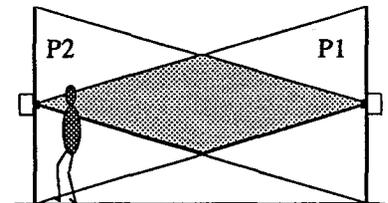


L'intrus ne coupe que le plan P1. Il ne laisse aucune trace sur les images du plan P2.

D'autre part lorsqu'un objet ou une personne passe près d'une des caméras, on obtient une obstruction complète d'un des champs visuels. La séquence d'images ligne présentée (séquence 2), ne permet pas de retrouver les paramètres géométriques définissant l'intrusion.



Séquence 2



Les séquences d'images ligne ci-dessus démontrent les limites du champ surveillée par un système à deux caméras.

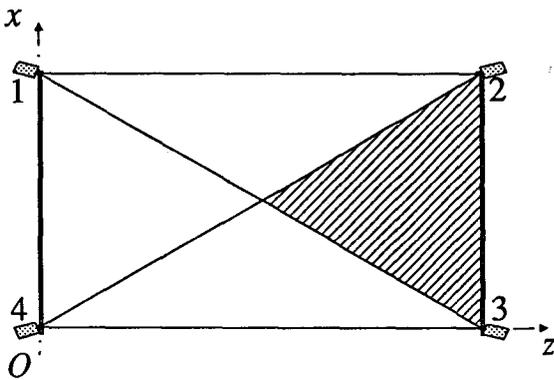
Le passage de l'objet dans la première séquence (séquence 1) n'est visible que dans un seul des plans de la barrière de surveillance. Aucun calcul concernant sa vitesse, sa position et sa taille ne peut être effectué.

La deuxième séquence correspond au passage de l'objet près de l'une des caméras (séquence 2). L'objet obstrue complètement le champ visuel correspondant. Dans ces conditions, la position et la taille ne peuvent être calculées, seule une approximation de la vitesse de passage peut être envisagée.

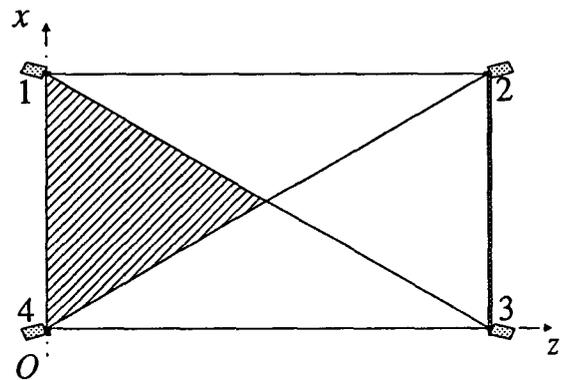
La carte SIAM développée au Centre d'Automatique de Lille a été conçue pour faire l'acquisition simultanée de quatre images ligne à partir de quatre caméras . Ceci permet d'envisager une barrière de surveillance à quatre caméras dont la disposition permet d'obtenir une couverture complète de la barrière (Fig. -1). On obtient deux

Les couples de caméras, (1, 3) d'une part et (2, 4)
d'autre part sont dans le même plan

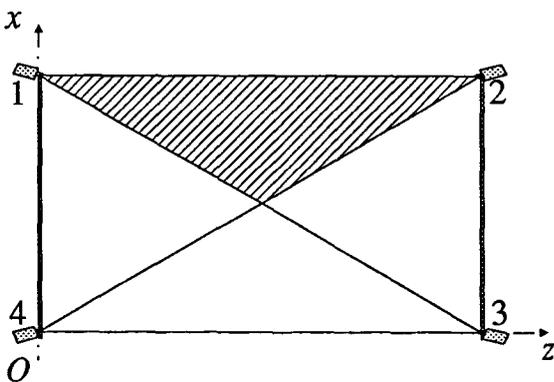
Zones surveillées par:



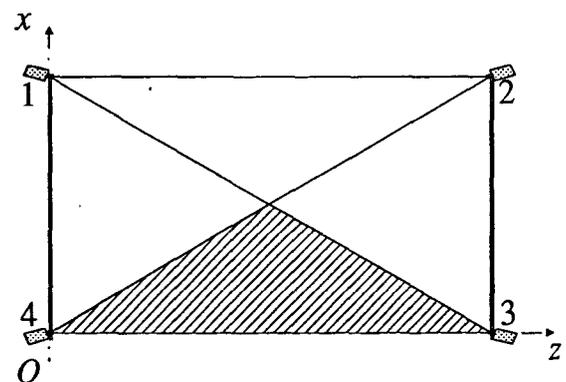
- les caméras 1 et 4



- les caméras 2 et 3



- les caméras 1 et 2



- les caméras 3 et 4

Fig.-1-Configuration à quatre caméras

plans de surveillance rectangulaires parallèles. Le premier est matérialisé par les champs des caméras 1 et 3, et le second par les champs des caméras 2 et 4.

Ainsi, quelle que soit la position d'un objet traversant le rectangle dont les quatres sommets sont les quatre centres optiques des caméras, celui ci sera vu par au moins un couple de caméras situées dans deux plans différents.

L'analyse des traces laissées par l'objet sur les couples d'images ligne est en tout point identique à celle présentée au chapitre III. Les techniques de calcul employées pour le système de détection d'intrusions à deux caméras restent donc valables pour cette configuration à quatre caméras.

Les résultats obtenus par les couples de caméras (1,4), (2,3), (1,2) et (3,4) couvrent tout le plan de surveillance (Fig - 1). La seule difficulté consiste à reconstituer un objet évoluant à cheval sur différentes zones de détection associés à différents couples de caméras pour retrouver les paramètres de position et de taille.

Par contre, l'automate développé au chapitre IV peut être utilisé sans modification majeure si on considère l'aspect temporel des intrusions à travers les deux plans rectangulaires parallèles distants de Δy formés par les caméras 1 et 3 d'une part et les caméras 2 et 4 d'autre part.

Ce système de détection basé sur le principe de la stéréovision linéaire pourrait être utilisé pour d'autres applications, par exemple, pour assurer la sécurité des aires de travail des robots ou des zones de passage de chariots filoguidés. Il remplacerait avantageusement les systèmes traditionnels à barrière optique qui déclenchent systématiquement sans aucune analyse des intrusions. De manière générale, on pourra envisager l'utilisation de la barrière optique, pour tout ce qui concerne la protection de zones bien définies dans un but soit de sécurité, soit de surveillance. l'évitement d'obstacles dans le cas de machines autonomes se déplaçant dans un univers bien défini.

BIBLIOGRAPHIE

- AGG 87** J.K. Aggarwal. [pp 50]
" Analysis of a sequence of image using point and line correspondences "
Proc. of IEEE International Conference on Robotics and Automation,
pp 1275-1280, April 1987.
- ADI 85** G. Adiv. [pp 53]
" Determining three-dimensional motion and structure from optical flow
generated by several moving objects "
PAMI Vol. 7, n° 4, pp 384-401, July 1985.
- AUB 89** D. Aubert. [pp 34]
" Mise en correspondance d'indices images en résolutions multiples "
Thèse de Doctorat 3^{ème} cycle, LIFIA, Grenoble, Janvier 1989.
- AYA 85** N. Ayache and B. Faverjon. [pp 35]
" Fast stereo matching of edge segments using prediction and verification
of hypotheses "
Proc. CVPR, pp 662-664, june1985.
- AYA 87** N. Ayache and O.D. Faugeras. [pp 56]
" Building, registrating and fusing noisy visual maps "
Proc. 1st ICCV, pp 73-82, june1985.
- BAJ 76** R. Bajcsy and L. Lieberman. [pp 42]
" Texture gradient as a depht cue "
CGIP Vol 5, pp 52-67, 1976.
- BAK 81** H. H. Baker and T. O. Binford. [pp 34]
" Depth from edge and intensity based stereo "
Proc . 7th IJCAI, pp 631-636, Vancouver, Canada, August 1981.
- BAR 80** S. T. Barnard and W. B. Thomson. [pp 32]
" Disparity analysis of images "
PAMI Vol. 2, n° 4, pp 333-340, July 1980.
-

- BER 86** F. Bergholm. [pp 43]
" Edge focusing "
Proc. of the 8th ICPR, pp 597-600, Paris, France, 1986.
- BOI 85** L. Boissier. [pp 26]
" Modélisation de l'environnement et localisation du robot mobile
HILARE par télémétrie laser "
Thèse de Doctorat 3^{ème} cycle, LAAS,1985.
- BRA 89** J. C. Brachet. [pp 23]
" Fall detection system in a metro line "
Control, Computers, Communication, Congrès AFCET, pp 143-147,
Paris, Septembre 1989.
- BRI 83** N.J. Bridwell and T. Huang. [pp 27]
" A discrete spatial representation for lateral motion stereo "
CVGIP Vol. 21, pp 33-57, 1983.
- BUR 82** P. J. Burt, C. Yen and X. Xu. [pp 32]
" Local correlation measures for motion analysis: a comparative study "
IEEE Proc. of PRIP, Las Vegas, USA, 1982.
- CAS 84** S. Castan and J. Shen. [pp 32]
"A stereo algorithm taking into account the perspective distorsions "
Proc. 7st Int. Conf. Pattern Recognition, Montreal, Canada, pp 441-443,
1984.
- CHO 87** T.-C. Chou and K. Kanantani. [pp 54]
" Recovering 3-D rigid motions without correspondance "
Proc. 1st Int. Conf. Computer Vision, London, England, pp 203-214, June
1987.
- X **COQ 87** J.-P. Coquerez and A. Gagalowicz. [pp 35]
" Mise en correspondance de régions dans une paire d'images stéréo "
Proc. Colloque Image du MARI-CESTA, pp 165-173, Mai 1987.
-

- CRO 84** J.L. Crowley and R.M. Stern. [pp 34]
" Fast computation of the Difference Of Low Pass transformation "
PAMI, march 1984.
- DAV 83** L.S. Davis, L. Janos and S.M.Dunn. [pp 42]
" Efficient recovery of shape from texture "
PAMI Vol 5, PP485-492, 1983.
- DRE 82** L. Dreschler et H. H. Nagel. [pp 47]
" Volumetric model and 3D trajectory of a moving car derived from
monocular TV frame sequence of a street scene "
CVGIP Vol 20, pp 199-228,1982.
- FAU 86** O. Faugeras et G. Toscani. [pp 75]
" The calibration problem for stereo "
Conference on CVPR, Miami-beach, USA, pp 15-20, June 1986.
- GER 86** A. Gerhard, H. Platzer, J. Steurer and R. Lenz. [pp 38]
" Depth extraction by stereo triples and a fast correspondance estimation
algorithm "
Proceedings of the 8th ICPR, pp 512-515, Paris, France, 1986.
- GRI 81** W. Grimson. [pp 34]
" An implementation of a computational theory of visual surface
interpolation "
CVGIP Vol 22, pp 39-69, 1983.
- GRO 87** P. Grossman. [pp 27]
" Depht from focus "
PRL Vol 5, pp 63-69, january 1987.
- GUR 86** E. Gurewitz, I. Dinstein and S. Sarusi. [pp 38]
" More on the benefit of a third eye for machine stereo perception "
Proceedings of the 8th ICPR, pp 966-968, Paris, France, 1986.
-

- HOP 79** J. E. Hopcroft J. D.Ullman. [pp 92]
" Introduction to automata theory languages, and computation " March,
1979.
- HOR 75** B. K. P. Horn. [pp 39]
" Obtaining shape from shading information "The psychology of
computer vision, Ed. New York, Mc Graw-Hill, pp 115-155, 1975.
- HOR 77** B. K. P. Horn. [pp 27]
" Understanding image intensities "
AI Vol 8, pp 201-231, 1977.
- HOR 81** B. K. P. Horn and B.G. Schunck. [pp 52]
" Determining optical flow "
AI Vol 17, pp 185-203, 1981.
- IKE 81** K. Ikeuchi. and B.K.P. Horn. [pp 41]
" Numerical shape from shading and occluding boundaries "
AI Vol. 17, pp 141-144, 1981.
- ITO 86** M. Ito and A. Ishii. [pp 38]
" Three view stereo analysis "
PAMI Vol 8, n° 4, pp 524-531, 1986.
- JAI 81** R. Jain. [pp 48]
" Extraction of motion information from peripheral processes "
PAMI Vol 3, n°5, pp 489-563, september 1981.
- JAR 76** R.A. Jarvis. [pp 44]
" Focus optimisation criteria for computer image processing "
Microscope 24, PP 163, 1976.
-

- JAR 83** R.A. Jarvis. [pp 26]
" A laser time of flight range scanner for robotic vision "
PAMI Vol 5, pp 505-512, 1983.
- JEN 86** M. Jenkin and J. K. Tsotsos. [pp 47]
" Applying temporal constraints to the dynamic stereo problem "
CVGIP Vol 33, pp 16-32, 1986.
- KAN 89** K. Kanatani and T. Chou. [pp 27]
" Shape from texture: General principle "
AI Vol. 38, pp 1-48, 1989.
- LAW 83** D. T. Lawton. [pp 54]
" Processing translational motion sequences "
CVGIP Vol. 17, pp 116-144, 1983.
- LEW 77** Lewis. [pp 26]
" A scanning laser range finder for robotic vehicle "
Proc. of the 5th IJCAI, pp 762-768, 1977.
- LIG 82** G. Lighthart and F. Groen.[pp 64]
" A comparison of different autofocus algorithms "
Proc. of ICPR, pp 597-600, 1982.
- LIU 86** Y.C. Liu and T.S. Huang. [pp 50]
" Estimation of rigid body motion using straight line correspondences:
Further results "
Proc. of the 8th ICPR, pp 306-309, Paris, France, 1986.
- MAR 79** D. Marr and T. Poggio. [pp 34]
" A computational theory of human stereo vision "
Proc. Roy. Soc. London Vol. B204, pp 301-328, 1979.
-

-
- MAR 80** D. Marr and E. Hildreth. [pp 34]
" Theory of edge detection "
Proc. Roy. Soc. London Vol B207, pp 187-217, 1980.
- ✕ **MED 85** G. Medioni and R. Nevatia [pp 35].
" Segment-based stereo matching "
CVGIP Vol. 31, pp 2-18, 1985.
- MIT 86** A. Mitiche, S. Seida and J.K. Aggarwal. [pp 50]
" Interpretation of structure and motion using straight line correspondences "
Proc. of the 8th ICPR, pp 1110-1112, Paris, France, 1986.
- ✕ **MON 87** O. Monga and B. Wrobel. [pp 36]
" Segmentation d'images : vers une méthodologie "
Traitement du Signal, pp 169-193, 1987.
- MOR 77** H.P. Moravec. [pp 32]
" Towards automatic visual obstacle avoidance "
IJCAI 5th, Cambridge, August 1977.
- MOR 81** H.P. Moravec. [pp 32]
" visual obstacle avoidance "
Proc. IJCAI, pp 785-790, 1981.
- NAG 86 a** H.H. Nagel. [pp 27]
" Images Sequence - Ten (octal) Years - From phenomenology towards a theoretical foundation "
Proc. of the 8th ICPR, pp 512-515, Paris, France, 1986.
- NAG 86 b** H.H. Nagel and W. Enkelman. [pp 53]
" An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences "
PAMI Vol 8, pp 565-593, September 1986.
-

- NAG 87** H.H. Nagel. [pp 53]
" On the estimation of optical flow: Relations between different approaches and some new results "
AI Vol. 33, pp 299-324, 1987.
- NEV 76** R. Nevatia. [pp 38]
" Depth measurement by motion stereo "
CGIP Vol. 5, pp 203-214, 1976.
- NIT 88** D. Nitzan. [pp 26]
" Three-dimensional vision structure for robot applications "
PAMI Vol 10, n°3, pp 291-309, May 1988.
- OHT 85** Y. Ohta and T. Kanade. [pp 34]
" Stereo by Intra- and inter-scanline search using dynamic programming "
PAMI Vol. 7, n°2, pp 139-154, March 1985.
- OHT 86** Y. Ohta, M. Watanabe and K. Ikeda. [pp 38]
" Improving depth map by right-angled trinocular stereo "
Proceedings of the 8th ICPR, pp 519-521, Paris, France, 1986.
- PIS 87** U. Pisu. [pp 23]
" Présentation succincte des systèmes de détection utilisables pour l'interface quai-voie "
RATP, Direction des Equipements Electriques, Service des Etudes, 1987.
- POS 82** J. L. Posdamer. and D. Altschuler [pp 77]
" Surface measurement by space-encoded projected beam systems "
CGIP Vol 18, pp 1-17, 1982.
- PRA 81** K. Prazdny. [pp 54]
" Determining the instantaneous direction of motion from optical flow generated by curvilinearly moving observer "
CGIP Vol. 17, pp 238-248, 1981.
-

- PRA 83** J.M. Pragger and M. Arbib. [pp 47]
" Computing the optic flow: The MATCH algorithm and prediction "
CVGIP Vol. 24, pp 271-304, 1983.
- RAN 80** S. Ranade and A. Rosenfeld. [pp 47]
" Point pattern matching by relaxation "
PR Vol. 12, pp 269-275, 1980.
- ROA 79** J.W. Roach and J.K. Aggarwal. [pp 49]
" Computer tracking of objects moving in space "
PAMI Vol. 1, n°2, pp 127-135, April 1979.
- ROA 80** J.W. Roach. [pp 49]
" Determining the movement of objects from a sequence of images "
PAMI Vol 2, n°6, pp 554-562, November 1980.
- SCH 86** B. G. Schunck. [pp 53]
" The image flow constraint equation "
CVGIP Vol. 35, pp 20-46, 1986.
- SET 87** S.K. Sethi and R. Jain. [pp 47]
" Finding trajectories of feature points in a monocular images sequence "
PAMI Vol. 9, n°1, pp 56-73 , January 1987.
- SHI 71** Y. Shirai and M. Suwa. [pp 27]
" Recognition of polyhedrons with a range finder "
Proc. 2nd IJCAI, pp 80-87, London, September 1971.
- STE 79** K. A. Stevens. [pp 42]
" Surface perception from local analysis of texture and contour "
Ph. D. dissertation, MIT Cambridge, 1979.
-

- STE 85** P. Stelmaszyk. [pp 48]
" Analyse de scènes dynamiques par recherches des contours en mouvement. Application à la conduite automatique d'un tramway "
Thèse de doctorat 3^{ème} cycle, USTLFA, Novembre 1985.
- SUB 87** M. Subbarao. [pp 53]
" Solution and uniqueness of image flow equations for rigid curved surfaces in motion "
Proc. 1st Int. Conf. Computer Vision, London, pp 687-692, June 1987.
- TSA 81** R. Y. Tsai and T.S. Huang. [pp 49]
" Estimating 3-D motion parameters of rigid planar patch "
IEEE Trans. Speech Signal Processing, Vol ASSP-29, n° 6, pp 1147-1152, December 1981.
- TSA 84** R. Y. Tsai and T. S. Huang. [pp 49]
" Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surface "
PAMI Vol. 6, n° 1, pp 13-26, January 1984.
- TUC 81** J. A. Tucker and M. Stark. [pp 44]
" A focus checking techniques for image analysis systems "
PR Vol 14, 1981.
- ULL 79** S. Ullman. [pp 48]
" The interpretation or structure from motion "
Proc. Roy. Soc. london, Vol B203, pp 405-426, 1979.
- X **VIE 88** C. Vieren. [pp 48]
" Segmentation de scènes dynamiques en temps réel. Application au traitement de séquences d'images pour la surveillance d'un carrefours routiers "
Thèse de doctorat 3^{ème} cycle, USTLFA, Avril 1988.
- WAX 84** A.M. Waxman, B. Kamgar-Parsi and M. Subbarao. [pp 53]
" Closed form solutions to image flow equations "
Proc. 1st Int. Conf. on Artificial Intelligence, Denver, pp 12-23, 1984.
-

- WIL 80** T.D. Williams. [pp 54]
" Depth from camera motion in a real world scene "
PAMI Vol 2, pp 511-516, November 1980.
- WIT 81** A.P. Witkin. [pp 42]
" Recovering surface shape and orientation from texture "
AI Vol. 17, pp 17-47, 1981.
- YAC 83** M. Yachida. [pp 53]
" Determining velocity maps by spatio-temporal neighborhoods from
image sequences "
CVGIP Vol. 21, pp 262-279, 1983.
- YAC 86** M. Yachida, Y. Kitamura and M. kimachi. [pp 38]
" Trinocular vision: New approach for correspondance problem "
Proceedings of the 8th ICPR, pp 1041-1044, Paris, France, 1986.
- YAK 78** Yakismovsky and R. T. Cunnigham. [pp 28]
" A system for extracting three-dimensional measurements from a stereo
pair of TV cameras "
CGIP Vol 7, pp 195-210, 1978.
- YEN 83** B. L. Yen and T.S. Huang. [pp 50]
" Determining 3-D motion/structure of a rigid body over 3 frames using
straight line correspondences "
Proc. of the IEEE Computer Society Conf. on Computer Vision and
Pattern Recognition, Washington DC, pp 267-272, June 1983.

ANNEXES

ANNEXE 1

A partir de l'expression matricielle:

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = A \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad \text{avec } A = \begin{pmatrix} \cos \Phi & 0 & \sin \Phi & b_x \\ 0 & 1 & 0 & 0 \\ -\sin \Phi & 0 & \cos \Phi & b_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

on obtient les équations:

$$(1) \begin{cases} X' = X \cdot \cos \Phi + Z \cdot \sin \Phi + b_x \\ Y' = Y \\ Z' = -X \cdot \sin \Phi + Z \cdot \cos \Phi + b_z \end{cases}$$

Soit p et p' les deux projections du point P sur les plans images de gauche et de droite respectivement. Les coordonnées de p et p' sont:

$$\text{pour } p: \quad x = \frac{X}{Z} \quad \text{et} \quad y = \frac{Y}{Z} \quad \text{car} \quad f = 1$$

$$\text{et pour } p': \quad x' = \frac{X'}{Z'} \quad \text{et} \quad y' = \frac{Y'}{Z'} \quad \text{car} \quad f' = 1.$$

On cherche les coordonnées de P dans le repère associé à la caméra de droite (X', Y', Z') , en fonction des coordonnées des points-image p et p' dans les deux images et des paramètres définissant la géométrie du système, b_x , b_z et Φ .

En observant que, $X = x.Z$ et $Y = y.Z$ et en remplaçant ces expressions dans les équations (1) on obtient:

$$(2) \begin{cases} X' = x.Z.\cos \Phi + Z.\sin \Phi + b_x \\ Y' = y.Z \\ Z' = -x.Z.\sin \Phi + Z.\cos \Phi + b_z \end{cases}$$

On détermine Z en fonction des paramètres de la géométrie du système avec:

$$x' = \frac{X'}{Z'} = \frac{x.Z.\cos \Phi + Z.\sin \Phi + b_x}{-x.Z.\sin \Phi + Z.\cos \Phi + b_z} \quad \text{puis}$$

$$x' = \frac{Z(x.\cos \Phi + \sin \Phi) + b_x}{Z(-x.\sin \Phi + \cos \Phi) + b_z}$$

et finalement on obtient pour Z:

$$Z = \frac{b_x - x'.b_z}{\cos \Phi (x' - x) - \sin \Phi (1 + x.x')} \quad (3)$$

ce qui donne pour les équations (2) avec la valeur de Z:

$$(4) \begin{cases} X' = \frac{\cos \Phi (x'.b_x - x.x'.b_z) - \sin \Phi (x'.b_z + x.x'.b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 + x.x')} \\ Y' = \frac{y(b_x - x'.b_z)}{\cos \Phi (x' - x) - \sin \Phi (1 + x.x')} \\ Z' = \frac{\cos \Phi (b_x - x.b_z) - \sin \Phi (b_z + x.b_x)}{\cos \Phi (x' - x) - \sin \Phi (1 + x.x')} \end{cases}$$

Vérification: Si on suppose les plan-images parallèles et dans le même plan alors (Figure-A-1):

$$\Phi = 0 \quad \text{et} \quad b_z = 0 ,$$

ce qui simplifie les équations (4):

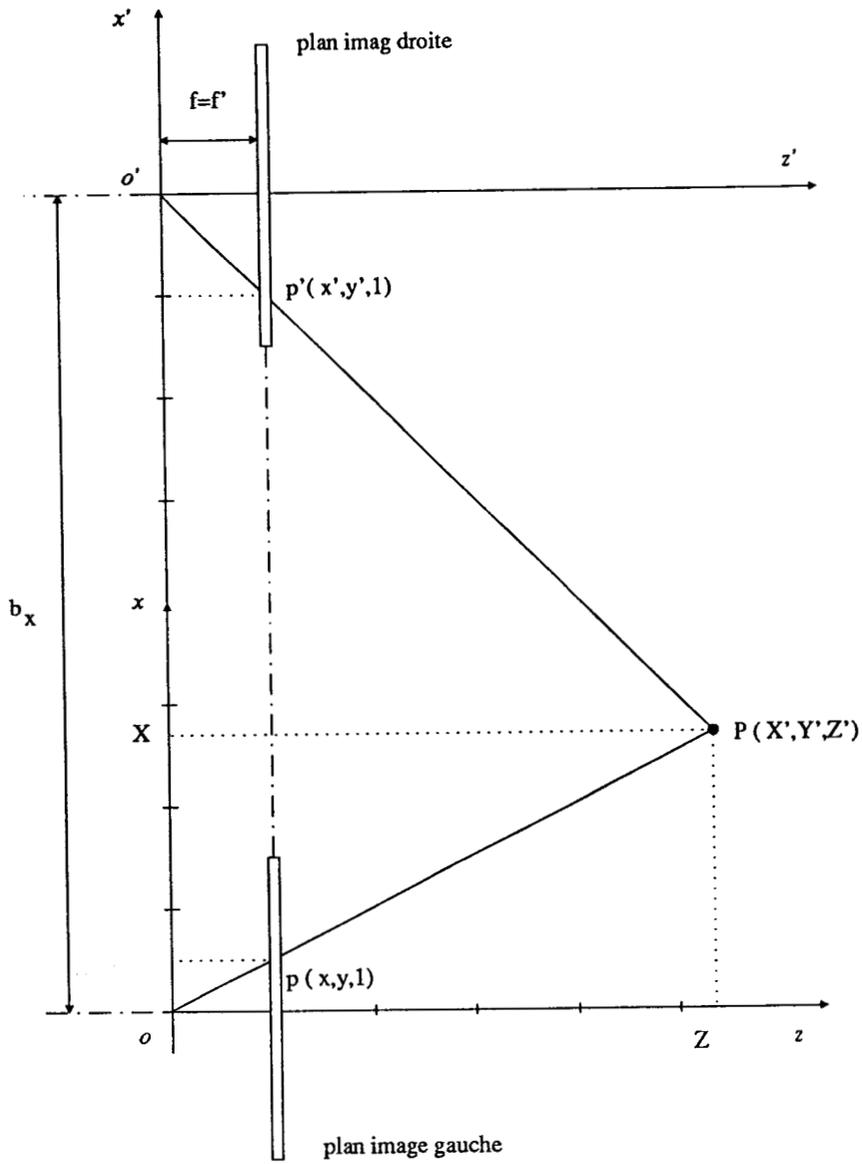


Figure-A-1- Géométrie classique.

$$(5) \begin{cases} X' = \frac{x' \cdot b_x}{(x' - x)} \\ Y' = \frac{y \cdot b_x}{(x' - x)} \\ Z' = \frac{b_x}{(x' - x)} \end{cases}$$

ANNEXE 2

1) Calcul théorique de $d=f(p)$.

On s'intéresse d'abord à la relation liant la distance (d) séparant un point situé dans le plan de surveillance au point O à la taille (l) de l'image du point dans le plan image, $l=f(d)$.

Les objectifs photographiques sont des combinaisons épaisses de plusieurs lentilles centrées sur un même axe, et dont on choisit les formes et les indices pour minimiser les diverses aberrations géométriques et chromatiques. Pour des raisons de simplifications, nous assimilerons cependant à une lentille mince unique, convergente. On rappelle alors les lois de formation d'une image. Dans un système optique supposé *parfait* (figure A-1), on a la relation des lentilles minces de Descartes:

$$\frac{1}{D'} = \frac{1}{f} - \frac{1}{D} \quad \text{ou} \quad \frac{1}{D'} + \frac{1}{D} = \frac{1}{f} \quad (1),$$

D et D' sont appelés distances ultranodales et f distance focale.

Ouverture relative d'un diaphragme:

Quand le diamètre de l'ouverture utile d'un diaphragme est égale à la fraction $\frac{1}{n}$ de la distance focale f, l'ouverture utile $\frac{f}{n}$ est et l'ouverture relative est $\frac{1}{n}$.

Pour permettre de donner à l'ouverture d'un objectif différentes valeurs possibles, les objectifs sont munis d'un diaphragme à ouverture variable par rotation d'une bague extérieure à la monture. Les diamètres des diaphragmes d'un objectif sont en progression géométrique de raison $\sqrt{2}$.

Les valeurs de n sont: 0,5 - 1 - 1,4 - 2 - 2,8 - 4 - 5,6 - 8 -

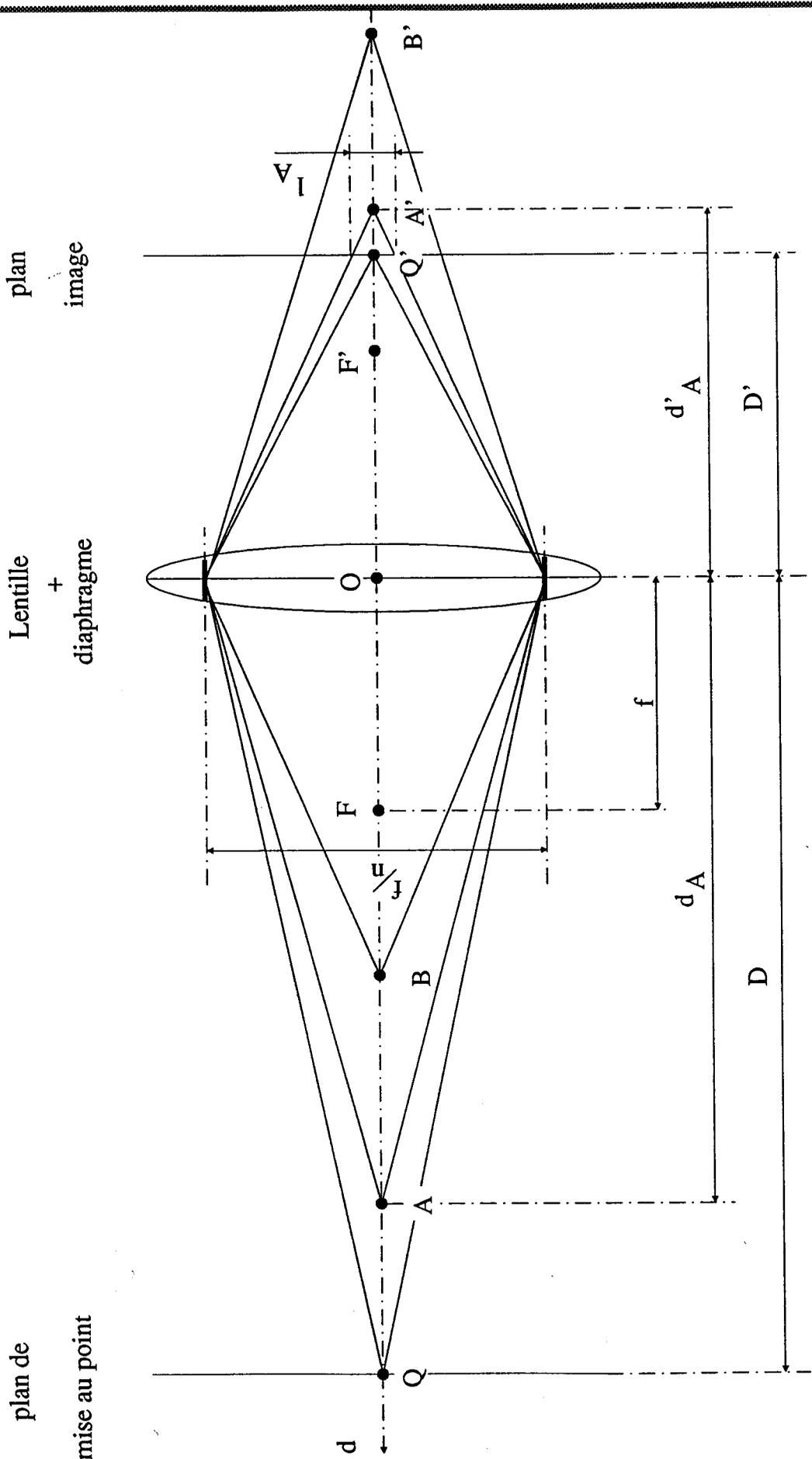


Figure -A-1 - Image de points A et B situés hors du plan de mise au point

Relativement à un objectif d'ouverture $\frac{f}{n}$, soient D et D' les distances ultranodales du plan de mise au point et du plan conjugué des éléments photosensibles, d la distance ultranodale du point objet situé en avant du plan de mise au point et d' la distance ultranodale de l'image nette correspondante.

La longueur l_A de l'image du point A enregistrée dans le plan image est exprimée par:

$$l_A = \frac{f}{n} \frac{d_A' - D'}{d_A} \quad (2)$$

A partir de l'équation (1) exprimée sous une autre forme,

$$\frac{1}{f} = \frac{1}{D'} + \frac{1}{D} = \frac{1}{d_A} + \frac{1}{d_A'}$$

on obtient,

$$d_A' = \frac{f d_A}{d_A - f} \quad \text{et} \quad D' = \frac{f D}{D - f} \quad (3)$$

puis on remplace les expressions exprimées en (3) dans l'équation (2)

$$l_A = \frac{f}{n} \left[\frac{\frac{f d_A}{d_A - f} - \frac{f D}{D - f}}{\frac{f d_A}{d_A - f}} \right]$$

ou, après simplifications,

$$l_A = \frac{f^2}{n(D-f)} \left[\frac{D}{d_A} - 1 \right] \quad \text{ou} \quad l_A = \frac{(\frac{f}{n})^2}{(\frac{D}{n}) - (\frac{f}{n})} \left[\frac{D}{d_A} - 1 \right] \quad (4)$$

et en généralisant, quelque soit le point A, l'équation de $l=f(d)$ se met sous la forme:

$$l = K \cdot \left[\frac{D}{d} - 1 \right] \quad (5)$$

avec $K = \frac{(f/n)^2}{(D/n) - (f/n)}$ constant si la distance du plan de mise au point D et l'ouverture utile $\frac{f}{n}$ sont fixes.

Si A est un point de contour d'un objet alors l'épaisseur du contour image est égale $l_A = l_e \cdot k$.

l_e : taille d'un pixel et k : nombre de pixels du contour image.

Le calcul de la pente donne:

$$p = \frac{\Delta A}{k} \text{ avec } k = \frac{l}{l_e} \text{ et } \Delta A \text{ constant pour un éclairage constant.}$$

en remplaçant l par sa valeur de l'équation (5) on obtient:

$$p = \frac{\Delta A \cdot l_e \cdot d}{K(D - d)}$$

et finalement l'équation $d=f(p)$:

$$d = \frac{p \cdot K \cdot D}{p \cdot K + \Delta A \cdot l_e} \quad (5).$$

2) Calcul de la profondeur de champ.

On cherche à déterminer la distance P_c séparant les plans entre lesquels un point donne une image de taille inférieure à la taille d'un pixel. Cette distance est appelée profondeur de champ.

Soit $L_e = \frac{D}{D'} \cdot l_e$ (6) taille-objet d'un pixel.

L_e est aussi égal à:

$L_e = \frac{f}{n} \cdot \frac{(D - d_b)}{d_b}$ (7) pour une taille-objet d'un pixel situé à une distance d_b avant le plan de mise au point.

d'où on obtient pour d :

$$d_b = \frac{D}{(n/f) \cdot L_e + 1}$$

De la même manière, pour une distance d_f après le plan de mise au point.

$$d_f = \frac{D}{1 - (n/f) \cdot L_e}$$

La profondeur de champ est égal à: $P_c = d_f - d_b$

et la profondeur de champ avant le plan de mise au point: $P_b = D - d_b$.

ANNEXE 3

1) Présentation de la carte SIAM.

La carte permet de réaliser l'acquisition des images de quatre caméras linéaires simultanément et s'intègre parfaitement au système de vision linéaire Visionix.

Cette carte est commandée par un registre de commandes 14 bits et par la carte IICAM du système Visionix. Sa fonction principale est de permettre l'acquisition simultanée des images provenant de quatre caméras linéaires et leur acheminement séquentiel vers la carte IICAM. La carte possède également un mode de réglage de caméra qui relie directement une caméra au choix, à la carte IICAM sans passer par les mémoires RAM de stockage, ainsi qu'un mode de lecture par le bus VME permettant un stockage ou un traitement par la carte maître du bus (carte CPU) du système Visionix.

Il est à noter également que la carte réalise l'interface entre les caméras linéaires de type Micronix et le système de vision linéaire Visionix qui a été conçu pour utiliser des caméras linéaires de type Cyclope.

2) Analyse fonctionnelle.

La réalisation est du type modulaire. Chaque module pilote une caméra de façon autonome et peut fonctionner suivant un des quatre modes ci-dessous:

- acquisition simultanée de une ou plusieurs caméras,
- restitution séquentielle des images ligne vers la carte IICAM,
- lecture des images ligne par le bus VME et
- transparent.

Cette carte a été réalisé autour de composants programmables du type EPLD (Eraseable Programmable Logic Device). Les modules ont été définis de façon à ce qu'ils soient les plus simples possible en utilisant au maximum les possibilités des

EPLD. Chaque module est formé d'une mémoire RAM 2K octets, d'une EP1810, d'un buffer 8 bits et de drivers caméra.

Cette carte est composée de deux parties différentes.

Une partie permettant son fonctionnement global:

- un contrôleur d'interruptions réalisé avec le composant Am 9519,
- un timer 8254,
- une logique de décodage et
- des registres de commande et de contrôle.

La deuxième partie est formée des quatres modules identiques composés chacun:

- une mémoire RAM,
- des buffers,
- des drivers de ligne RS 422 pour la caméra et
- une EPLD 1810.

Chaque module permet le pilotage d'une caméra et partage comme ressources communes les drivers de lignes vers la carte IICAM et le buffer de sortie vers le bus VME. Dans chaque module, l'EPLD gère tous les signaux de commandes externes.

Pour la RAM: A0-A10 , \overline{WE} , \overline{CE} , \overline{OE} .

Pour les drivers caméras: S1 .

Pour le pilotage de la carte IICAM: horloge, start.

Pour le pilotage caméra: EEC.

3) Fonctionnement de la carte.

3.1) sélection d'un module

La sélection des modules de pilotage de caméra se fait par l'intermédiaire du registre de commande, 6 bits y sont affectés:

- 2 bits de mode M0 et M1 permettant de différencier un mode parmi quatre,
- 4 bits de sélection (select), chaque bit étant spécifique à un module.

Deux bits, IN et OUT, des EPLD permettent le chaînage entre modules. Dans le mode acquisition simultanée, une à quatre caméras peuvent être sélectionnées simultanément, tandis que dans les autres modes, une et une seule parmi quatre peut être sélectionnée. On a alors:

M0	M1	SELECTn	IN	OUT	selectEPLDn
1	1	1	0	0	1
1	1	0	0	0	0
x	x	0	0	0	0
x	x	0	1	0	0
x	x	1	0	1	1
x	x	1	1	1	0

$$\text{selectEPLD} = (\text{int} + \overline{\text{select}}) \quad \text{OUT} = ((\overline{\text{M0.M1}}) \text{selectEPLD}) + ((\overline{\text{M0.M1}}) \text{IN})$$

Si plusieurs modules sont sélectionnés dans un mode différent du mode d'acquisition simultanée, ce sera toujours le module n° 1 qui sera sélectionné.

3.2) Sélection du mode de fonctionnement.

Acquisition: mode4 = M0.M1 , Restitution: mode3 = $\overline{\text{M0}} . \overline{\text{M1}}$

Lecture par bus: mode2 = M0. $\overline{\text{M1}}$, Transparent: mode1 = $\overline{\text{M0}} . \text{M1}$.

Le mode3 est le mode dans lequel se retrouve la carte après un reset.

3.3) Le mode transparent.

Le dénominateur "transparent" veut dire que la carte fonctionne de façon à ce quelle soit totalement transparente pour la carte IICAM. On a un fonctionnement identique à celui que l'on aurait si la carte IICAM était reliée directement à une caméra.

Ce mode est en fait, un mode de réglage de luminosité. On peut choisir une caméra parmi les quatre et la carte réalise alors l'interface entre la carte IICAM et la caméra Micronix.

Dans ce mode, l'EPLD EP1810 libère un chemin direct pour les 6 bits de données de la caméra, entre les drivers de la carte IICAM et la caméra.

Fonctionnement.

Lorsqu'un des modules est sélectionné, le passage à l'état zéro des signaux \overline{GO} et \overline{LEC} de la carte IICAM provoque un front descendant du signal EEC envoyé à la caméra ce qui commande une lecture de celle-ci. La caméra envoie alors ses données ainsi qu'un signal d'horloge (SPC) et d'autres signaux notamment VBC signalant la présence du signal de l'horloge.

Dans ce mode, il faut inverser l'horloge (SPC) provenant de la caméra avant de l'injecter vers la carte IICAM, car la carte IICAM considère que les données sont valides à partir du front montant de l'horloge. De plus, il faut fabriquer un signal start indiquant à la carte IICAM que la caméra est occupée (busy), ce signal est fabriqué par le signal VBC de la caméra Micronix.

Signaux régissant le fonctionnement en mode transparent:

clock IICAM = Mode1. \overline{SPC} START = Mode1. VBC
EEC = Mode1.selectEPLD. \overline{GO} . \overline{LEC} S1 = Mode1.selectEPLD

3.4) Le mode lecture par le bus VME

Ce mode permet de lire l'image stockée en mémoire RAM par l'intermédiaire du bus VME en vue d'un stockage ou d'un traitement futur. Le passage des données de la mémoire RAM vers le bus VME s'effectue par l'intermédiaire d'un buffer. La mémoire RAM du module est vue sous une seule adresse, chaque lecture à cette adresse (\$FF0109) provoque l'incrémentation du compteur d'adresses et c'est la donnée correspondant à cette adresse qui est lue sur le bus VME.

L'EPLD libère un chemin direct entre la mémoire RAM et le buffer de sortie en passant par le buffer de données (S2 = 0).

La mémoire RAM est commandée par les adresses, tout front montant sur \overline{CS} provoque l'incrémentation du compteur d'adresses (la lecture par le bus VME ayant lieu lorsque \overline{CS} est à l'état 0).

Remarque: Lors de la sélection du Mode2, le compteur d'adresses est initialisée à \$7FF, ce qui fait que lors de la première lecture du bus VME, on lit l'adresse \$7FF à la place de l'adresse \$000. Ce petit problème peut être facilement résolu en ne tenant pas compte du premier octet lu, et en faisant 1729 lectures au lieu de 1728.

Ce problème du premier pixel est du à un aléas de fonctionnement. En effet il est très facile de mettre le compteur à l'adresse \$000 lors du passage en mode2 en utilisant la remise à zéro (RAZ) du compteur.

On a alors: RAZ = mode2.fin (fin = A1.A2....A11).

Les huit premiers bits d'adresses, A1 à A8 passent par un ET huit entrées puis par un "CONF" (registre enterré de l'EPLD) pour générer un signal. Ce signal étant moins rapide que le signal passant par le ET quatre entrées prenant en compte les bits d'adresses A9 à A11 puis du ET huit entrées, provoque, lors du passage de l'adresse \$6FF à \$700, une remise à zéro du compteur imprévisible.

Ce problème de retard dans la propagation des signaux à l'intérieur de l'EP1810 n'a pu être résolu à cause de la structure très lourde du compilateur de schéma et de

la structure même des EP1810 qui devient rapidement très rigide lorsqu'un grand nombre de macrocellules est utilisé.

Signaux régissant le fonctionnement en mode lecture par le bus VME.

Signaux internes à l'EPLD EP1810.

compteur enable = mode2

$CK = \overline{CS}$

Signaux externes.

$\overline{WE} = 1$

$\overline{OE} = 0$

$\overline{CE} = 0$

$S1 = 0$

$S2 = 0$

EEC = TimerOUT

3.5) Le mode restitution vers la carte IICAM.

Dans ce mode, l'image stockée dans la mémoire RAM est transmise vers la carte IICAM à une cadence de 4 MHz. La carte IICAM demande une lecture caméra, et c'est en fait le module sélectionné qui simule la caméra et envoie les données contenues dans la RAM. Une fois la mémoire RAM vide, le module se bloque pour le faire fonctionner de nouveau il faut le désélectionner puis le resélectionner. La désélection remet le module dans les conditions initiales.

L'EPLD libère un chemin direct entre la mémoire RAM et les drivers de la carte IICAM en passant par le buffer de données ($S2 = 0$).

Fonctionnement.

L'horloge utilisée est une horloge 4 MHz obtenue en divisant par deux, à l'aide d'une bascule, la fréquence de 8 MHz présente sur le module d'interface XVME085. Les échanges de données avec la carte IICAM se feront donc à 4 MHz (fréquence maximale possible).

Au départ, les compteurs sont initialisés à \$7FF c'est à dire $A1 = A2 = A3 \dots = A11 = 1$ et $A12 = 0$. Lorsque le module est sélectionné, un niveau bas sur \overline{GO} et \overline{LEC} autorise le comptage. Il est nécessaire que \overline{GO} et \overline{LEC} restent simultanément à zéro pendant un temps au moins égal à 250 ns pour qu'il puisse se produire un front montant sur l'horloge (4 MHz) du compteur permettant ainsi le passage à l'adresse \$800 ($A12 = 1$). Le bit d'adresse $A12$ à 1 autorise le comptage et crée le signal START envoyé vers la carte IICAM. On utilise un monostable pour générer la constante de temps (250 ns) nécessaire du signal \overline{LEC} . Ce monostable est réalisé avec un 74LS123. L'horloge envoyée à la carte IICAM est l'horloge 4 MHz inversée. Sur le front descendant, l'IICAM prend en compte la donnée. Le passage de l'adresse \$FFF à l'adresse \$000 fait passer $A12$ à zéro, ce qui inhibe le compteur en faisant passer START à zéro (caméra non "busy").

Signaux régissant le fonctionnement en mode restitution.

Signaux internes à l'EPLD:

$$\text{compteurenable} = (\text{mode3.selectEPLD} \cdot \overline{GO} \cdot \overline{LEC}) + (\text{mode3.selectEPLD} \cdot A12)$$

signaux externes.

$$\text{START} = \text{Mode3.selectEPLD} \cdot A12$$

$$\overline{WE} = 0$$

$$\overline{CE} = 0$$

$$\overline{OE} = 0$$

$$S1 = S2 = 0$$

$$\text{EEC} = \text{timerOUT}$$

3.6) Le mode acquisition simultanée.

Ce mode permet l'acquisition en mémoire RAM de l'image de une, deux, trois, ou quatre caméras. L'EPLD permet le passage des données des caméras vers la RAM. L'entrée dans ce mode, provoque lors du passage à zéro du signal timerOUT du 8254 un front descendant du signal EEC des modules sélectionnés déclenchant un cycle de lecture de leurs caméras respectives. Les données envoyées par les caméras sont stockées dans la mémoire RAM des modules.

Fonctionnement.

Lors de l'entrée en mode 4, une remise à zéro du compteur d'adresses est effectuée, ce qui fait que l'adresse initiale pointée par le compteur est \$000.

Les horloges, SPC de la caméra et CK du compteur, étant inversées, un front descendant sur SPC incrémente le compteur d'adresses, les données étant valides après le front montant de SPC l'écriture dans la RAM se fait quand SPC est à 1.

On a alors:

$$\overline{WE} = \overline{(\text{Mode4.SPC.VBC})}$$

$$\overline{CE} = (\text{Mode4.} \overline{(\text{Mode4.SPC.VBC})}) + \text{fin}$$

Signaux régissant le fonctionnement en mode acquisition:

Signaux internes à l'EPLD:

$$\text{compteur enable} = \text{Mode4.VBC}$$

$$\text{clear} = \overline{(\text{mode4.VBC.selectEPLD})}$$

Signaux externes:

$$\overline{WE} = \overline{(\text{Mode4.SPC.VBC})} \quad \overline{CE} = (\text{Mode4.} \overline{(\text{Mode4.SPC.VBC})}) + \text{fin}$$

$$\overline{OE} = 1 \quad S1 = S2 = 1$$

$$\text{EEC} = \text{Mode4.timerOUT}$$

3.7) Registre de commande

Le registre de commande est un registre 16 bits dont 14 sont disponibles. Il est constitué du registre de commande 8 bits présent sur le module XVME085 et d'un registre 8 bits d'extension, réalisé avec un composant 74LS273.

Le registre de commande est accessible en écriture sous le format d'un mot de 8 bits à l'adresse \$FF0080.

- bit 0 et 1: réservés à la carte XVME085 doivent être laissés à 0.
- bit 2 : M1 sélection du mode de fonctionnement.
- bit 3 : M0 sélection du mode de fonctionnement.
- bit 4: sélection du module 1.
- bit 5: sélection du module 2.
- bit 6: sélection du module 3.
- bit 7: sélection du module 4.
- bit 8: autorisation des interruptions.

3.8) Registre de status.

Le registre de status est un registre 8 bits dont 6 sont disponibles. Il est présent sur le module XVME085 et est accessible en lecture à l'adresse \$FF0081.

- bit 0 et 1: réservés au système.
 - bit 2 : caméra 1.
 - bit 3 : caméra 2.
 - bit 4: caméra 3.
 - bit 5: caméra 4.
 - bit 6: fin d'acquisition.
 - bit 7: état d'autorisation des interruptions.
-

4) Schéma interne des EPLD 1810.

