

N° d'ordre : 21

50376  
1985  
229

50376  
1985  
229

# THÈSE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour l'obtention du titre de ..

DOCTEUR DE L'UNIVERSITE

par

Patrick STELMASZYK



## **ANALYSE DE SCENES DYNAMIQUES PAR RECHERCHE DES CONTOURS DES OBJETS EN MOUVEMENT. APPLICATION A L'AUTOMATISATION D'UN TRAMWAY**

Soutenue le 22 Novembre 1985 devant la Commission d'Examen

MM.	P.	VIDAL	Président
	Y.	DAVID	Rapporteur
	B.	DUBUISSON	Rapporteur
	J.G.	POSTAIRE	Directeur de Recherche

- P L A N -

Dans le chapitre I, nous allons présenter l'intérêt et les difficultés relatifs à la conduite automatique d'un tramway par un système de traitement d'images. L'étude des propriétés des caméras infrarouges et de celles opérant dans le visible, ainsi que des techniques de perception de l'environnement tridimensionnel justifieront l'approche choisie, à savoir une caméra à semi-conducteurs, fixe par rapport à la scène.

L'étude des différentes méthodes permettant d'analyser le mouvement fait l'objet du second chapitre. Celui-ci est divisé en deux parties. La première regroupe les techniques d'analyse structurale, la seconde traite de la correspondance au niveau des pixels. Les limites propres à chaque méthode sont mises en oeuvre.

Après avoir montré que les techniques de différence d'images apportent une réduction importante des données à traiter, nous avons développé, dans le chapitre III, deux méthodes permettant d'isoler les points de contour des objets en mouvement. Une discussion sur les limites de ces méthodes, dont l'une est essentiellement basée sur des techniques monodimensionnelles et l'autre sur une approche bidimensionnelle, permet d'en apprécier les performances.

Dans le chapitre IV, on recherche les modèles rectangulaires des objets dont on a détecté les contours en mouvement. Les positions et dimensions de ces rectangles donnent une description suffisante de la scène. On isole alors le tramway et on calcule les distances le séparant des objets en mouvement. Cette mesure associée à une prédiction du mouvement, permet de mettre en oeuvre un processus de conduite automatique.

Les algorithmes décrits dans le chapitre précédent sont appliqués sur une maquette de simulation représentant des tramways, piétons, voitures évoluant dans un carrefour urbain. Une séquence de 13 images montre que les procédures développées tiennent compte de situations complexes et peuvent générer la commande automatique du tramway. Pour chaque acquisition, les ordres de commande, les distances minimales et les positions prédites sont affichées sur les images.

Ces résultats montrent les limites de l'approche proposée et indiquent les perspectives liées à une application en site réel.

- S O M M A I R E -

---

. CHAPITRE I : METHODOLOGIE DE CONCEPTION D'UN SYSTEME DE CONDUITE AUTOMATIQUE D'UN TRAMWAY.

- I.1. Conduite assistée par ordinateur
- I.2. Choix du type de caméra
  - I.2.1. Caméras infrarouges
  - I.2.2. Caméras opérant dans le visible
  - I.2.3. Propriétés des caméras
- I.3. Analyse de scène à partir du poste de pilotage
  - I.3.1. Vision stéréoscopique
  - I.3.2. Vision monoculaire
  - I.3.3. Vision monoculaire et télémétrie
  - I.3.4. Cas d'une scène dynamique
  - I.3.5. Conclusion sur les caméras embarquées.
- I.4. Conclusion

. CHAPITRE II : METHODES DE DETECTION D'OBJETS EN MOUVEMENT PAR ANALYSE DE SEQUENCES D'IMAGES.

- II.1. Détection d'objets en mouvement par analyse structurelle
  - II.1.1. Correspondance point à point
  - II.1.2. Correspondance des contours
  - II.1.3. Correspondance entre structures
- II.2. Détection d'objets en mouvement par correspondance au niveau des pixels
  - II.2.1. Analyse du gradient spatio-temporel
  - II.2.2. Méthodes de Fourier
  - II.2.3. Méthodes de corrélation
  - II.2.4. Différence d'images
    - II.2.4.a. Sans image de référence
    - II.2.4.b. Avec image de référence

**. CHAPITRE III : DETECTIONS DES CONTOURS DES OBJETS EN MOUVEMENT  
PAR TRAITEMENT MONO et BIDIMENSIONNEL.**

**III.1. Segmentation de scènes dynamiques par analyse de  
séquences d'images**

III.1.1. Choix d'une méthode de détection d'objets  
en mouvement

III.1.2. Approche par analyse de régions

III.1.3. Approche par détection de contours

III.1.4. Principe des méthodes proposées

**III.2. Détection monodimensionnelle**

III.2.1. Possibilité et limites d'un traitement mono-  
dimensionnel

III.2.2. Segmentation de l'image des différences

III.2.3. Extraction des points de contour dans  
l'image courante

III.2.4. Résultats

**III.3. Détection bidimensionnelle**

III.3.1. Multiplication binaire

III.3.2. Multiplication des niveaux de gris

III.3.3. Binarisation de l'image multiplication

**III.4. Robustesse des méthodes proposées.**

**. CHAPITRE IV : MODELISATION DES OBJETS EN MOUVEMENT ET PROCESSUS  
DE CONDUITE AUTOMATIQUE.**

**IV.1. Modélisation des objets en mouvement**

IV.1.1. Détermination du rectangle exinscrit aux  
objets en mouvement

IV.1.2. Modèles rectangulaires des objets en mouvement

IV.1.2.1. Reconstitution d'un objet en mouvement

IV.1.2.2. Elimination du bruit

IV.1.3. Performances des procédures développées

## IV.2. Processus de conduite automatique

IV.2.1. Simulation et initialisation

IV.2.2. Différentes étapes du processus de conduite automatique

IV.2.2.1. Analyse de scène

IV.2.2.1.a. Poursuite du tramway

IV.2.2.1.b. Recherche des objets sur la voie

IV.2.2.2. Stratégie de conduite

IV.2.2.2.a. Risque de collision mesuré

IV.2.2.2.b. Risque de collision estimé

IV.2.2.2.c. Ordres de commande

IV.2.2.3. Structure de commande de la conduite automatique

## . CHAPITRE V : SIMULATION DE LA TRAVERSEE D'UN CARREFOUR EN CONDUITE AUTOMATIQUE.

V.1. Conditions expérimentales

V.2. Analyse de la stratégie de conduite

V.3. Conclusion.

. CONCLUSION

. BIBLIOGRAPHIE.

## CHAPITRE I

METHODOLOGIE DE CONCEPTION D'UN SYSTEME DE  
CONDUITE AUTOMATIQUE D'UN TRAMWAY.

## I.1. CONDUITE ASSISTEE PAR ORDINATEUR

Les techniques de traitement d'images ont commencé à se développer autour des années 1960-1970 avec l'apparition des premiers systèmes permettant de digitaliser les images.

La motivation la plus importante de ces travaux a été la robotique où la reconnaissance des formes en vue de la manutention, puis de l'assemblage automatique des pièces, a fait l'objet de recherches très importantes.

Bien que la majorité des travaux traite encore d'images statiques, il s'est formé un nouvel axe de recherche orienté vers l'analyse du mouvement dans une séquence d'images. Les travaux sont devenus rapidement importants du fait de la diversité des applications, Parmi celles-ci on trouve (HUA 81) :

- Les problèmes militaires : poursuite multicibles, détection et reconnaissance de cibles à partir d'images infrarouges,
- Les problèmes industriels : surveillance dynamique d'un processus industriel, vision dynamique appliquée à la robotique,
- Les problèmes commerciaux : compression du signal T.V.,
- Les problèmes médicaux : étude du comportement des cellules, ainsi que du mouvement du coeur en imagerie aux rayons X,
- La météorologie : poursuite des nuages,
- Les transports : surveillance du trafic routier.

Le principal objectif de l'analyse de scènes dynamiques est de déterminer les informations qui ne peuvent être extraites à partir d'une image unique. C'est pourquoi, les différentes techniques d'analyse de scènes dynamiques reposent sur l'utilisation de séquences d'images, ces images correspondant à des acquisitions prises à intervalles de temps régulièrement espacés.

Le problème consiste à étudier les transformations qui apparaissent entre les images successives en tenant compte du lieu temporel entre ces images. Cette notion de séquence d'images est différente de l'accumulation d'images dans laquelle aucun lien temporel ne relie les images entre elles.

Le domaine d'application dans lequel se situe notre étude est celui des transports. Le relevé des caractéristiques du trafic routier (débit, vitesse, taux d'occupation.....) a fait l'objet de recherches importantes qui ont nécessité l'emploi de capteurs spécialisés (boucles d'induction, câble piezo-électriques....). Cependant, le développement des opérations de contrôle et de régulation du trafic routier entraîne des besoins croissants au niveau de la saisie des données qui est souvent confié à des caméras vidéo. Mais le traitement manuel de ces observations est fastidieux, sinon impossible (DEP 84). Les informations recherchées (origines, destinations, temps de trajet, comptage directionnels aux carrefours, détection de situations de trafic anormales....) peuvent être obtenues par des processus automatiques basés sur l'utilisation d'un (ou plusieurs) système(s) de traitement d'images.

Les possibilités offertes par une telle approche résultent de la richesse des procédures de traitement d'images et il a semblé intéressant d'appliquer ces techniques d'analyse d'images à l'automatisation de la conduite d'un système de transport en commun.

L'automatisation de tels systèmes s'est surtout développé en site propre (métro par exemple). Le VAL, métro de Lille, en est l'exemple le plus convaincant. Mais la construction d'un tel site nécessite des travaux de génie civil impressionnants dont le coût est élevé.

De ce fait, on découvre aujourd'hui l'intérêt économique des transports en surface, tels les tramways, bien que ceux-ci n'aient pas encore bénéficié de l'essor technologique qui a bouleversé, ces dernières années, les systèmes de transport en site propre. On peut néanmoins envisager de protéger la majeure partie du parcours (auquel cas le tramway se comporte comme un métro) et de limiter au maximum les zones dans lesquelles le tramway se trouve mêlé à la circulation routière. Le Mongy, tramway reliant Lille à Roubaix et Tourcoing est un exemple typique de moyen de transport en site semi-propre où seule la traversée des carrefours présente des difficultés d'automatisation. Le problème fondamental

que soulève une conduite automatique d'un tel moyen de transport est celui de la perception de l'environnement permettant d'éviter les collisions avec des objets mobiles au niveau des carrefours. Il va de soi que cette technique de contrôle et de conduite ne constitue qu'un aspect des possibilités de conduite. Elle pourra être associée à des techniques déjà approuvées, telles que l'utilisation de radars, capteurs tactiles, capteurs de proximité, boucles d'induction, câbles piezo-électriques, etc....

Dans notre étude nous nous sommes intéressés au traitement de l'information visuelle saisie par une ou plusieurs caméra(s) reliée (s) à un système de vision par ordinateur. L'objectif de ce travail vise à proposer une méthodologie concrète en vue de détecter, au niveau des carrefours, les objets mobiles présentant un risque de collision avec le tramway et de générer un processus de commande en vue de piloter ce tramway.

Le cahier des charges ne nous imposant pas une méthodologie particulière, il a été nécessaire de concevoir des approches raisonnables du fait que la bibliographie relative aux transports ne nous a pas orienté vers des travaux similaires.

Avant d'étudier les procédures logicielles, il s'est avéré indispensable de déterminer le type de caméra à employer (opérant dans le visible ou dans l'infrarouge) et d'étudier les problèmes relatifs à l'analyse de la scène. Les choix adoptés sont justifiés dans les paragraphes suivants.

## I.2. CHOIX DU TYPE DE CAMERA

Les différentes caméras présentes sur le marché peuvent être différenciées selon qu'elles opèrent dans le domaine du visible ou dans l'infrarouge. Ces différents types de caméra peuvent être utilisées avec n'importe quel système de vision du fait de la compatibilité de leur signal de sortie.

### I.2.1. Caméras infrarouges

Les caméras infrarouges opèrent, de manière générale, dans les longueurs d'ondes comprises entre 8 et 14  $\mu$ , la principale source d'émission étant la self-émission thermique (WAD 78, WAT 84). Ainsi, l'intensité des radiations sur une surface, qui est interprétée comme une valeur de niveaux de gris par le capteur, dépend du pouvoir émissif de la surface et de sa température. Les différences de température sont causées par le soleil et par l'énergie dissipée par les objets eux-mêmes (moteurs, pots d'échappement...). Ainsi l'image infra-rouge se présente comme une carte de température dans laquelle chaque pixel de l'image se voit attribuer un niveau de gris correspondant à la température de l'objet.

Il existe deux sortes de caméras infrarouges :

- caméras à refroidissement (cryogenically-cooled detectors),
- caméras pyro-électrique (pyro-électric vidicons).

Dans les premières, on expose un détecteur à chaque pixel de l'image par un procédé optique. Comme le détecteur est exposé successivement aux différents pixels de l'image, on obtient une information série analogue au signal video standart. Ces caméras ont besoin d'être refroidies par un liquide approprié et demandent une maintenance régulière.

Les caméras pyroélectriques utilisent une surface dont la charge électrique est fonction des variations de température. Selon ce principe, seules les variations de température sont détectées. De ce fait, dans une scène dynamique, on ne détecte sur l'image que les objets en mouvement. Les objets stationnaires, quant à eux, ne sont pas visibles sur cette image du fait de l'absence de variations de température.

### I.2.2. Caméras opérant dans le visible

L'image obtenue au moyen de ces caméras est constituée de pixels dont la valeur est proportionnelle non plus à la température, mais à la luminance. Il existe deux types de caméras (BON 83, MAI 79, THO 81a,) :

- caméras à tube
- caméras à semi-conducteurs.

Dans les premières, une cible photoconductrice reçoit, en chacun de ses points, une charge électrique qui dépend de l'intensité lumineuse reçue. Un faisceau électronique explore successivement ces points et crée un signal vidéo correspondant à la luminance du pixel exploré.

Les caméras à semi-conducteurs peuvent se présenter sous forme de matrice ou de ligne, formée d'éléments sensibles, ces derniers pouvant être des photo-Mos ou des photo-diodes.

### I.2.3. Propriétés des caméras

Les récents progrès technologiques ont permis d'améliorer la structure des caméras infrarouges à refroidissement, en multipliant le nombre de détecteurs par caméra, formant ainsi une ligne ou une matrice. Malgré ces perfectionnements, la qualité de l'image n'est pas satisfaisante. De plus, comme la dynamique de différents détecteurs ne sont pas identiques, il est nécessaire de filtrer l'image. La taille de la fenêtre devant être importante (57 pixels), cette opération est coûteuse en temps (LOU 83).

Les caméras pyroélectriques offrent une image plus acceptable. Certaines techniques (obturateur mécanique par exemple), permettent de disposer des éléments stationnaires présents dans la scène ; mais cette solution a pour effet de bruyé considérablement l'image.

Les caméras infrarouges présentent la particularité de fournir une image indépendante des conditions d'éclairage. Ainsi, de jour comme de nuit, le contenu des images reste similaire. Les images de la même scène en éclairage diurne et nocturne ne sont pas parfaitement identiques, car l'environnement thermique de la scène n'est pas exactement le même ; mais en ce qui concerne un véhicule en mouvement, l'image de cet objet ne présentera pas de variations importantes entre ces deux conditions extrêmes de prises de vue.

Une autre propriété, certainement la plus importante dans notre application, concerne leur faculté d'ignorer les ombres. En effet, dans le domaine du visible, la principale difficulté est due à la présence d'ombres qu'il est souvent difficile de différencier des objets qui les créent car elles les accompagnent dans leurs déplacements. Mais bien que l'utilisation des caméras infrarouges permette de s'affranchir de ces perturbations, on observe des effets similaires dus aux réflexions infrarouges des sources chaudes ou froides.

L'idée selon laquelle les caméras infrarouges peuvent travailler dans de mauvaises conditions de visibilité n'est que partiellement exacte. En fait, dans ces conditions, la supériorité des caméras infrarouges est compromise. Dans le brouillard les avantages diminuent et deviennent nuls si le brouillard est dense. En cas de pluie ou de neige, on ne trouve aucun intérêt de disposer d'une caméra infrarouge. On peut donc conclure qu'au niveau de l'image, caméra infrarouge et caméra opérant dans le visible se trouvent sensiblement sur le même pied d'égalité pour la vision diurne. La principale raison qui freine l'utilisation des caméras infrarouges est liée à leur production qui reste faible. Leur coût est par conséquent très élevé. De plus, afin d'obtenir une transmission des radiations infrarouges, il est nécessaire d'utiliser des optiques spéciales, les verres devenant opaques pour ces valeurs spectrales. Ces objectifs utilisent des lentilles au germanium et leur ouverture doit être grande. Leur prix est par conséquent beaucoup plus élevé que celui des objectifs classiques.

Actuellement, il semble difficile de justifier l'emploi d'une caméra infrarouge. Cependant, un intérêt croissant pour ces caméras est en train de se développer. Les nouvelles applications sortent du domaine militaire où leur emploi était jusqu'alors cantonné (surveillance de feux de forêts, détection de fissures dans les tunnels, médecine...). On peut alors espérer une diminution considérable du prix de ces capteurs. Mais dans l'état actuel de la technologie, notre choix s'est orienté vers les caméras travaillant dans le visible.

Les caméras à tubes bien que d'un prix modique et d'une bonne définition, ne peuvent être utilisées dans notre application. Les distorsions géométriques dues aux défauts d'alignement du canon à électrons ainsi que les dérives thermiques et le vieillissement des circuits électroniques altèrent la qualité de l'image. De plus, ces caméras présentent une instabilité du balayage liée aux systèmes électroniques de dérivation. Ce phénomène introduit des perturbations importantes dans les problèmes de comparaison d'objets et de différence d'images.

Les caméras à semi-conducteur, malgré un coût supérieur, offrent l'avantage de présenter une distorsion très faible. La réduction de leur taille et de leur poids apporte un confort d'utilisation non négligeable. Ces caméras, à la différence de celles à tubes, peuvent être pointées directement sur une source lumineuse intense (soleil ou réflexion de celui-ci sur un pare-brise de voiture) sans risque de détérioration. En outre, l'effet de rémanance, propre au tube à électrons, qui a pour conséquence d'imprégner la surface photosensible des objets stationnaires, n'existe pas dans les caméras à semi-conducteurs. De plus, l'emploi de ces caméras dans le domaines des transports a montré que la durée de vie des caméras à tube était de 1 000 heures alors que celle des caméras à semi-conducteurs était de 10 000 heures (EQU 84).

Dans ce paragraphe, nous avons souligné l'importance du choix du capteur. Grâce à leurs qualités intrinsèques, les caméras à semi-conducteur supplantent celles à tubes qui se voient désormais reléguées, dans la plupart des cas, à un rôle de surveillance manuelle.

Nous utiliserons, par conséquent, une caméra à semi-conducteur (type CCD) en attendant que des progrès technologiques, tant en conception qu'en fabrication, rendent attrayant l'emploi d'une caméra infrarouge.

### I.3. ANALYSE DE SCENE A PARTIR DU POSTE DE PILOTAGE

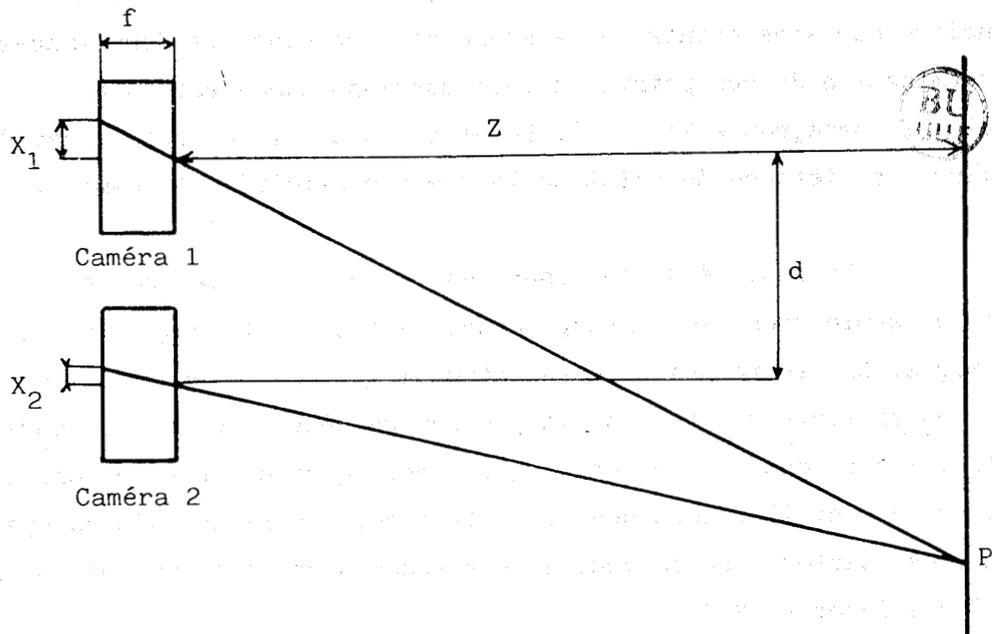
Il semblerait "naturel", afin de conduire le tramway, de placer la caméra à l'intérieur du poste de pilotage. Mais le problème relatif à ce type d'approche concerne la détermination de la profondeur. Cette troisième dimension correspond en fait à la distance séparant le tramway des obstacles potentiels présents dans la scène.

Nous allons présenter les différentes techniques permettant d'extraire l'information profondeur et les illustrer par des applications concrètes. La littérature spécifique aux transports étant très pauvre dans ce type d'application, les exemples choisis seront issus du domaine des robots mobiles qui présente de grandes similarités avec le problème d'un tramway évoluant dans un univers inconnu.

Il n'est pas possible d'exposer, en un chapitre, tous les problèmes relatifs aux robots mobiles. Le lecteur pourra se référer aux articles indiqués s'il désire plus de détails sur les méthodes de localisation des objets tridimensionnels (COR 83) ou sur les stratégies de commande (BRI 83, CHA 81, FER 82).

#### I.3.1. Vision stéréoscopique

Le principe de la stéréovision est basé sur l'utilisation de deux caméras géométriquement décalées, leurs axes optiques étant parallèles, afin d'évaluer la troisième dimension. A partir des deux images ainsi obtenues, il est possible de localiser un point par triangulation (cf. figure I.1.)



- Figure I.1. -

La profondeur  $Z$ , calculée entre le point objet et le plan image, est donnée par la relation suivante :

$$Z = \frac{f \cdot d}{x_1 - x_2}$$

$f$  : distance entre le plan image et le centre des lentilles,

$d$  : distance entre le centre des lentilles,

$x_1$  et  $x_2$  : coordonnées des projections du point objet.

Cette mesure ne peut, évidemment, s'effectuer sur tous les points de l'image. Il est nécessaire, dans un premier temps, de reconnaître certains points de la scène dans chacune des deux images. La détermination de ces points et leur correspondance est un problème délicat (cf. paragraphe II.1.1.). Ensuite, un simple calcul permet de déterminer la distance du point de la scène considéré à la caméra.

La qualité de la mesure est fonction de la précision dans la détermination des correspondances entre points significatifs des deux images. En dehors des erreurs introduites dans la phase d'acquisition, il faut aussi tenir compte des distorsions géométriques des capteurs. En effet, les rayons de projection qui, théoriquement, passent par le centre de la lentille, sont déviés du fait des non-linéarités du système optique. Ces distorsions peuvent être atténuées en utilisant un modèle de la caméra adéquat (COR 83).

Le pilotage du robot "J.P.L. Rover", véhicule autonome d'exploration de la planète Mars, est basé selon ce principe. Un effort important a été engagé au niveau de la précision de la mesure ( $\pm 5$  mm à une distance de 2 m) en modélisant et en calibrant les caméras (YAK 78).

Le guidage du "CART", développé à l'Université de Standford, fait appel à des techniques de stéréovision (MOR 77, MOR 79, MOR 83). Du fait de l'incertitude sur les modèles des caméras et des erreurs de correspondance, il n'était pas possible d'évaluer les distances. Aussi, une variante, appelée "SLIDER STEREO", a été développée. On anime la caméra d'un mouvement latéral afin de prendre 9 prises de vue d'une même scène. Une corrélation des 36 (8!) paires d'images ainsi obtenues permet d'apparier les points choisis sur l'image centrale avec leurs correspondants des 8 autres images. Les 36 distributions sont alors sommées, le pic résultant indiquant la bonne mesure de la profondeur.

Les techniques de stéréovision ne présentent pas la souplesse d'utilisation nécessaire à l'application envisagée du fait de la complexité des traitements (calcul des points caractéristiques, recherche des correspondances) et de l'importance des structures matérielles mises en oeuvre (nécessité de stocker deux ou plusieurs images en mémoire). Celles-ci sont très coûteuses en temps. A titre d'exemple, on relève que le robot "CART", se déplace d'un mètre toutes les 10 à 15 minutes.

Du fait de la simplicité de l'univers dans lequel évoluent les robots décrits et des difficultés rencontrées par leurs auteurs, il ne semble pas réaliste d'utiliser une telle approche pour guider un véhicule dans une scène complexe et réelle.

### I.3.2. Vision monoculaire

Dans le cas où une seule caméra est utilisée, la troisième dimension peut être évaluée par plusieurs techniques. Parmi celles-ci, on peut citer :

#### a) Technique pseudo 3.D. (BOL 82)

On cherche à retrouver des informations sur le relief à partir d'une image 2 D. L'information profondeur est donnée par une interprétation des formes perçues au moyen des outils de l'intelligence artificielle. La scène est alors décrite de façon simple (objet A devant objet B). Cette solution apporte un indice sur la profondeur et non une valeur mesurée.

#### b) Technique des ombres portées (WAL 75)

Connaissant la position de la source lumineuse qui éclaire la scène, on recherche les frontières et les directions des ombres portées par les objets. Cette méthode nécessite la connaissance de la position du plan de référence, mais aussi la description des objets constituant la scène. Ces restrictions limitent cette technique aux scènes de laboratoire.

c) Techniques de la lumière structurée

Ces techniques sont basées sur l'étude de la déformation d'une grille projetée sur la scène.

Plusieurs types de grilles peuvent être utilisées.

- Grilles à fente verticale (OSH 79). On projette sur l'objet une raie de lumière verticale puis on déplace la fente en changeant la direction du faisceau. On arrive ainsi à retrouver les coordonnées tridimensionnelles des points objets par triangulation et on agrège les points appartenant à la même surface de l'objet. Cette approche est très coûteuse en temps.

- Grilles à n fentes. On utilise n fentes parallèles en vue de réduire le temps d'expérimentation. Une telle approche pose des problèmes d'identification de lignes qui limitent le type de scène à analyser.

- Grilles à lignes orthogonales (FRE 79). On fait plusieurs acquisitions sous des angles différents pour avoir une vue totale de l'objet. On calcule ensuite les matrices de transformation pour chaque image afin d'obtenir une représentation tridimensionnelle de chaque élément de surface.

d) Technique de stéréophotométrie (IKE 83)

La position de la caméra et l'éclairage étant connus, il est possible d'établir des relations entre l'orientation des surfaces, leurs niveaux de gris et les caractéristiques de réflexion. Afin de déterminer la profondeur, il suffit de disposer d'une seconde prise de vue de la même scène mais avec un éclairage différent. En pratique, cette méthode nécessite plusieurs acquisitions avec, à chaque fois, un éclairage différent en vue de lever les ambiguïtés relatives aux calculs.

e) Technique K2D

On recherche l'information profondeur à partir de plusieurs (K) images 2D, différentes les unes des autres. L'obtention de ces différentes prises de vue s'adapte aux trois situations suivantes :

- . Caméra fixe et objet mobile (CHE 80)
- . Caméra mobile et scène fixe (WIL 80)
- . Plusieurs caméras fixes.

Les méthodes permettant de retrouver la troisième dimension sont relatives au type de situation rencontré. On remarque que la stéréovision est un cas particulier de la technique K2D.

Pour des robots mobiles utilisant comme moyen de perception de l'environnement la vision monoculaire, l'évaluation du relief de la scène est basée sur une approche différente de celles présentées ci-dessus.

Le système de perception du robot SHAKEY fait appel aux contours de la scène. Les points de contour, obtenus par convolution de l'image avec un masque de type gradient (cf. paragraphe III), sont réunis en petits segments de droite. Ces derniers sont lissés et connectés en vue de réunir toutes les lignes de l'image. Le robot repère alors, sur le sol, les zones exemptes d'obstacles.

Une approche similaire est utilisée pour le robot MEL-DEIC, du Mechanical Engineering Laboratory (Japon). Les contours verticaux possèdent une faible variance de l'intensité lumineuse dans la direction verticale. La connaissance de ces bords verticaux permet, sans faire d'analyse de scènes mais par une méthode basée sur le calcul d'invariants, d'identifier les objets dans la scène et, par conséquent, de piloter le robot.

Du fait des contraintes liées aux différentes techniques permettant d'évaluer les coordonnées tridimensionnelles, il semble difficile d'étendre leur application à l'étude d'une caméra embarquée se déplaçant dans un univers complexe et inconnu.

### I.3.3. Vision monoculaire et télémétrie

Cette technique nécessite deux capteurs. Une caméra, afin de segmenter l'image ou d'isoler les points significatifs et un télémètre (laser ou à ultrason) utilisé pour mesurer la profondeur. Ainsi, la troisième dimension n'est pas directement mesurée en exploitant les informations visuelles. En combinant les données de ces deux capteurs, il est possible de détecter les obstacles présents dans la scène.

Le système de perception du robot ANU (Australian National University) utilise cette technique. Le télémètre (à ultrason) monté sur une plate-forme rotative, permet de balayer le champ de vision de la caméra. Quant au robot développé au L.A.A.S. de Toulouse, il exploite, dans un premier temps, l'information visuelle perçue par la caméra en vue d'isoler des points significatifs. La profondeur est ensuite mesurée en ces points par un télémètre laser. Ce dernier, couplé à un système de miroirs défecteurs, permet d'explorer le même champ que la caméra.

La mauvaise résolution des télémètres à ultrason, du fait d'un cône d'émission important, ne permet pas d'évaluer ponctuellement une distance. Les télémètres lasers s'avèrent nettement plus performants et ce, indépendamment des conditions d'éclairage. Il est ainsi possible de mesurer une distance dans l'obscurité.

Ce type de perception s'avère intéressant mais son emploi est, pour l'instant, limité aux scènes simples composées d'objets polyédriques. Le compromis entre le temps d'acquisition des images et la résolution recherchée semble limiter la complexité de l'univers dans lequel évolue le robot.

On peut cependant remarquer que la mesure ponctuelle de profondeur peut être utilisée pour associer une région de l'image soit à un objet, soit à son ombre et lever ainsi certains ambiguïtés dans une procédure de segmentation délicate.

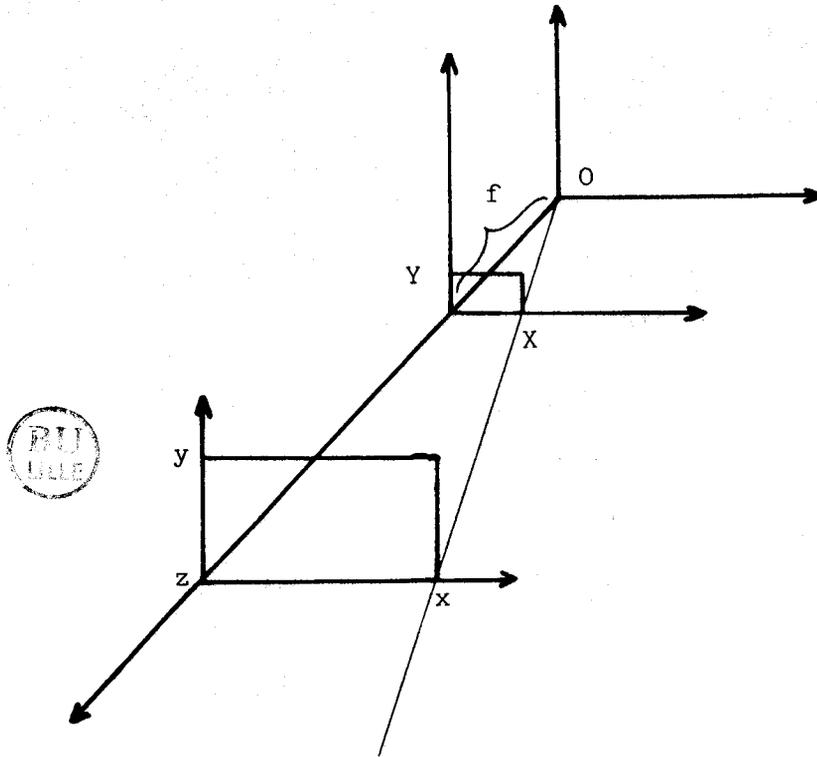
#### I.3.4. Cas d'une scène dynamique

Les diverses approches basées sur l'utilisation de caméras embarquées ont montré que l'univers dans lequel elles évoluent est simple (objet polyédriques et rigides). A ces contraintes, vient s'ajouter l'hypothèse selon laquelle les objets constituant la scène sont stationnaires. On voit que ces restrictions empêchent la généralisation de ces méthodes dans un environnement complexe (et composé d'objets en mouvement) que représente une scène réelle.

Une étude bibliographique sur l'analyse de scènes dynamiques par une caméra en mouvement a souligné le peu d'intérêt que portent les chercheurs à ce domaine, à l'exception cependant de JAIN (JAI 84) qui, par des considérations d'ordre géométrique, arrive à simplifier le problème.

L'approche qu'il a utilisée est basée sur la notion de FOE (Focus of Expansion), qui résulte de l'intersection d'un vecteur 3D représentant la direction instantanée du mouvement de l'observateur avec le plan image. La détermination de ce point dans une scène fixe perçue par une caméra mobile (WIL 80, WIL 81), ou une scène dynamique avec une caméra fixe (PRA 80), permet d'extraire des informations sur les objets telles que la profondeur ou le mouvement des surfaces. JAIN a montré que la connaissance de ce point autorise la détection des objets mobiles.

La projection d'un point objet sur le plan image permet de déterminer les coordonnées de ce point dans l'image.



- Figure I.2. -

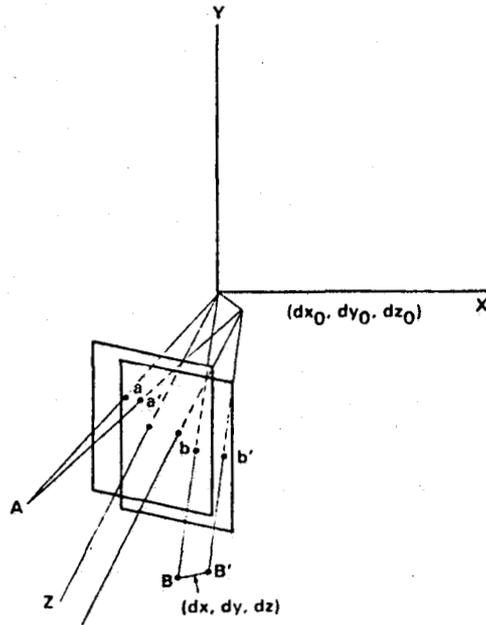
Nous voyons sur la figure I.2. que :

$$\frac{X}{x} = \frac{f}{z}$$

En considérant une distance focale  $f$  unitaire, nous pouvons écrire les relations suivantes :

$$X = \frac{x}{z}$$

$$Y = \frac{y}{z}$$



- Figure I.3. -

Le déplacement  $(dx_0, dy_0, dz_0)$  de la caméra associé à un déplacement éventuel  $(dx, dy, dz)$  du point objet dans l'espace (fig. I.3.), vérifie les équations ci-dessous :

$$X = ((x+dx) - (x_0+dx_0)) / ((z+dz) - (z_0+dz_0))$$

$$Y = ((y+dy) - (y_0+dy_0)) / ((z+dz) - (z_0+dz_0))$$

On en déduit aisément que les coordonnées polaires  $(r, \theta)$  d'un point par rapport au FOE sont données par :

$$\operatorname{tg} \theta = \frac{dzo(y+dy-y_0) - dy_0(z=dz-z_0)}{dzo(x+dx-x_0) - dx_0(z+dz-z_0)}$$

$$r = \sqrt{\left(X - \frac{dx_0}{dzo}\right)^2 + \left(Y - \frac{dy_0}{dzo}\right)^2}$$

Dans le cas où l'objet est stationnaire ( $dx = dy = dz = 0$ ), l'expression  $\operatorname{tg} \theta$  peut s'écrire de la manière suivante :

$$\operatorname{tg} \theta = (y-y_0) - (z-z_0) \frac{dy_0}{dzo} / (x-x_0) - (z-z_0) \frac{dx_0}{dzo}$$

dans laquelle  $\frac{dx_0}{dzo}$  et  $\frac{dy_0}{dzo}$  sont les coordonnées du FOE.

Dans cette équation, la position du FOE étant inchangée pour un même déplacement de la caméra, on voit que pour un point objet stationnaire, l'angle  $\theta$  entre ce point et le FOE est constant.

Ces considérations ont amené JAIN à transformer l'image originale en un tableau bidimensionnel ayant  $r$  pour abscisse et  $\theta$  pour ordonnée. Ainsi, dans cette configuration, les points stationnaires ont une variation de l'angle  $\theta$  nulle et sont représentés par des horizontales. Si ces points sont mobiles, les variations de l'angle  $\theta$  apparaissent le long de l'axe vertical.

L'auteur, par une procédure de différence d'images successives (méthode pourtant réservée au cas d'une caméra fixe et développée en détail au paragraphe II.2.4.a.) effectue une segmentation de la scène. La transformation dans le plan  $(r, \theta)$  permet alors de distinguer les objets mobiles des stationnaires. Cependant, de telles méthodes nécessitent un faible déplacement de la caméra et des objets entre deux acquisitions successives ainsi que la connaissance des paramètres  $dx_0, dy_0$  et  $dzo$  relatifs au déplacement de la caméra.

### I.3.5. Conclusion sur les caméras embarquées

L'étude sur les caméras embarquées a montré la difficulté d'exploiter l'information perçue par la (ou les) caméra(s) en vue de déterminer la troisième dimension. Celle-ci correspond, en fait, à la distance séparant le tramway d'un obstacle potentiel et doit être connue avec précision.

Les techniques de stéréovision semblent difficiles à exploiter en raison du nombre et de la diversité des objets présents dans la scène. Les temps de calcul sont trop importants pour envisager une application temps réel. Les techniques monoculaires ne sont pas envisageables dans notre application. Leur emploi reste limité à l'étude de scènes très simples tout comme celles associées à la télémétrie qui nécessitent un compromis temps d'acquisition - résolution.

De plus, ces différentes méthodes ne sont pas adaptées à notre situation du fait qu'elles supposent les objets constituant la scène stationnaires.

Dans le cas d'une caméra mobile pour l'analyse d'une scène dynamique, les récents travaux de JAIN (JAI 84) ouvrent la voie à de nouvelles espérances mais ne semblent pas utilisables immédiatement à notre application. En effet, les objets en mouvement étant détectés, il faudrait ensuite mesurer la distance les séparant de la caméra et de prédire leur trajectoire en vue de déterminer les risques de collisions éventuels avec le tramway. Ces problèmes semblent freiner une telle approche.

#### I.4. CONCLUSION

Dans ce chapitre, nous avons dégagé les principales caractéristiques des différentes caméras et justifié notre choix, à savoir une caméra opérant dans le visible de type CCD.

Nous avons aussi montré les nombreuses difficultés liées à l'analyse de la scène, pour une caméra placée à l'intérieur du poste de pilotage. Malgré l'intérêt indéniable qu'inspire une telle approche du point de vue scientifique, il s'est avéré nécessaire, afin de proposer une solution efficace et concrète dans les délais impartis, de simplifier le problème en plaçant la caméra au-dessus du carrefour. Le calcul de la profondeur n'étant pas nécessaire, on traite la scène comme étant bidimensionnelle.

Avant d'opter pour une méthode d'analyse de scène dynamique, à partir d'images saisies par une caméra fixe, il convient d'étudier la littérature déjà abondante sur ce sujet. Celle-ci est présentée de façon plus détaillée dans le second chapitre de ce mémoire.

## **CHAPITRE II**

**METHODES DE DETECTION D'OBJETS EN MOUVEMENT  
PAR ANALYSE DE SEQUENCES D'IMAGES.**

La détection d'objets en mouvement dans une séquence d'images a fait l'objet d'une littérature abondante du fait de la diversité des approches proposées. Ce chapitre n'a pas la prétention de donner une liste exhaustive de ces différentes techniques ainsi que de leurs applications. Il vise plutôt à regrouper certaines méthodes qui présentent des analogies. Deux types d'approches sont proposées ; celles basées sur l'analyse de la structure des objets et celles qui analysent l'image au niveau des pixels.

### II.1. DETECTION D'OBJETS EN MOUVEMENT PAR ANALYSE STRUCTURELLE

Dans ces méthodes, on identifie un ensemble d'éléments structurants dans une image et on cherche à les retrouver dans les images suivantes, d'où leur nom de "Feature based method" (AGG 83, ROS 82).

Une grande variété d'éléments structurants a été utilisée dans la littérature parmi lesquels on peut citer les angles, les arcs de cercle, les segments de droite, les barycentres, les surfaces, les axes d'inertie, etc.... Le choix des éléments structurants dépend du domaine d'application et des hypothèses que l'on peut faire sur les parties en mouvement.

#### II.1.1. Correspondance point à point

L'approche générale vise à établir des correspondances entre deux images successives (ou géométriquement décalées dans le cas de la stéréovision), de points significatifs dont les voisinages présentent des propriétés particulières.

On comprend que tous les points de l'image ne présentent pas le même intérêt, certains étant situés dans des régions où l'absence de détails n'autorise aucune distinction. Pour cette raison, la détermination des points significatifs s'avère être une étape importante.

Une approche classique consiste à segmenter l'image (par seuillage par exemple) puis à décrire les contours comme des chaînes de caractères. Les points anguleux sont alors déterminés par l'analyse de cette chaîne en utilisant entre autre la méthode de Rosenfeld et de Weszka (ROS 75) ou de Freeman (FRE 78). Mais la phase de segmentation introduit des erreurs qui affectent la détermination de ces points particuliers.

Aussi, il est plus intéressant de développer des techniques permettant d'appliquer un opérateur spécifique directement sur une image à niveaux de gris, sans effectuer de segmentation préalable.

Maître, Lopez et Krahej (MAI 84) définissent cet opérateur comme devant être isotrope, local, précis et insensible aux bruits. Après avoir appliqué sur toute l'image des opérateurs de type médian et Sobel, ils recherchent les maxima locaux de l'image filtrée qu'ils utilisent comme points significatifs. Lawton (LAW 83) utilise la méthode du passage à zéro du Laplacien pour déterminer les points à haute courbure.

Moravec (MOR 77, MOR 79), quant à lui, a décrit dès 1977, l'opérateur qui porte désormais son nom. Il calcule, dans de petites fenêtres, la somme des carrés des différences des pixels adjacents au pixel considéré et ce dans chacune des quatre directions (horizontale, verticale et diagonales) et en choisit le minimum. La fenêtre contient un point significatif si cette mesure de variation correspond à un maximum local.

Dans le cas où cet opérateur est appliqué sur un contour qui n'est pas collinéaire avec les directions horizontales, verticales ou diagonales, les petites variations dans la distribution des niveaux de gris affectent considérablement la réponse de l'opérateur. Pour cette raison, Drescher et Nagel (DRE 82) lui préfèrent l'opérateur de Baudet. Ces opérateurs et bien d'autres, on fait l'objet de comparaison et d'améliorations qui alimentent ponctuellement la littérature (KIT 82, DRE 83, WIL 83).

Moravec (MOR 77, MOR 79, MOR 83), utilise son opérateur afin d'établir des correspondances en stéréovision. Les points sélectionnés dans l'image 1 sont appariés avec ceux de la seconde en utilisant une méthode qui affine la solution tout au long de l'analyse (fine and coarse). La recherche débute sur des images réduites avec un rapport élevé (1/16). Une fenêtre ( $n \times n$ ) centrée sur le point traité dans la première image est corrélée avec l'image 2. Cette fenêtre, déterminée dans l'image 2, est alors repérée sur cette image pour un nouveau coefficient de réduction

de  $1/8$ . Cette zone, de taille  $(2n \times 2n)$ , devient la nouvelle surface de recherche pour la fenêtre  $(n \times n)$  de l'image 1 réduite également à  $1/8$ . Le processus se poursuit pour des images de moins en moins réduites jusqu'à ce que la fenêtre  $(n \times n)$  de l'image 1 non réduite soit décelée par corrélation dans l'image 2 non réduite.

Les points étant sélectionnés par le même opérateur que précédemment, on peut utiliser une méthode de correspondance par relaxation similaire à celle développée par Rosenfeld, Hummel et Zucker (ROS 76). La probabilité initiale d'une correspondance possible est estimée par une mesure de corrélation. Ces estimations sont réinitialisées de manière itérative en considérant la distribution des probabilités de correspondance dans le voisinage de chaque point. Après un certain nombre d'itérations (de l'ordre de 10), les probabilités deviennent faibles, sauf pour les points ayant trouvé leur correspondant et le processus peut-être arrêté. Cette étude a été aussi menée dans le cadre de la stéréovision (BAR 80).

Cette approche a été améliorée (DRE 82) en développant une procédure de fin d'itération et en remplaçant certains paramètres introduits arbitrairement par des calculs basés sur la connaissance de la structure de l'objet en mouvement. Une approximation polyédrique ensuite effectuée afin de modéliser l'objet qui se déplace dans une scène réelle.

En météorologie, dans l'analyse du déplacement des nuages, l'utilisation de l'algorithme ISODATA (END 71) permet de détecter les différentes zones que constituent les nuages. Les centres de ces régions, appelés centres d'intensité, sont utilisés comme points significatifs. La détermination du mouvement repose sur un processus itératif qui essaye d'apparier ces points significatifs en utilisant une fonction de coût. Cette procédure, qui converge rapidement (de l'ordre de 3 itérations), semble particulièrement adaptée à cette étude, les nuages ne possédant pas de structure interne.

Pour étudier l'impact thermique, dû aux rejets d'eau chaude des centrales nucléaires en bord de mer, on analyse le déplacement de petits flotteurs (SQU 83). Les points significatifs sont représentés, dans cette application, par le barycentre de chaque flotteur. La correspondance est alors basée sur une méthode relationnelle ayant la structure d'arbre de décision qui utilise les critères suivants : maximum de vitesse, continuité temporelle, cohérence spatiale, correspondance unique et caractéristique du flux avoisinant ce point.

Les deux méthodes précédentes (END 71), (SQU 83), donnent des résultats satisfaisants du fait de la nature des déplacements. Ceux-ci sont réguliers, sans accélérations, ni changements de direction brutaux. Ces conditions ne sont pas vérifiées dans les scènes de trafic routier.

Les méthodes de correspondance de points caractéristiques réduisent énormément le nombre de points à traiter et par conséquent diminuent les temps de calcul. En outre, de telles techniques, à l'exception des deux dernières méthodes présentent l'avantage d'apporter des indices sur la structure de l'objet qui pourraient servir de guide pour une segmentation ultérieure.

Ces méthodes s'avèrent robustes dans les scènes à deux dimensions et en stéréovision. Mais dans une scène de trafic routier, dans laquelle un univers tridimensionnel est considéré comme une scène à deux dimensions, l'image des objets en mouvement peut varier selon leur position par rapport à la caméra et à la source d'éclairage. La détection des points significatifs étant affectée par ces variations, le processus de correspondance amène des ambiguïtés rendant difficile un appariement correct.

### II.1.2. Correspondance des contours

Les méthodes de correspondance des contours décrites dans ce paragraphe utilisent les contours pour représenter les objets constituant la scène. Cependant, dans certaines applications, d'autres paramètres tels que surface, périmètre, angles et axes d'inertie peuvent être introduits. Ces techniques analysent le mouvement avec un haut niveau de compréhension, notamment dans le cas des occlusions.

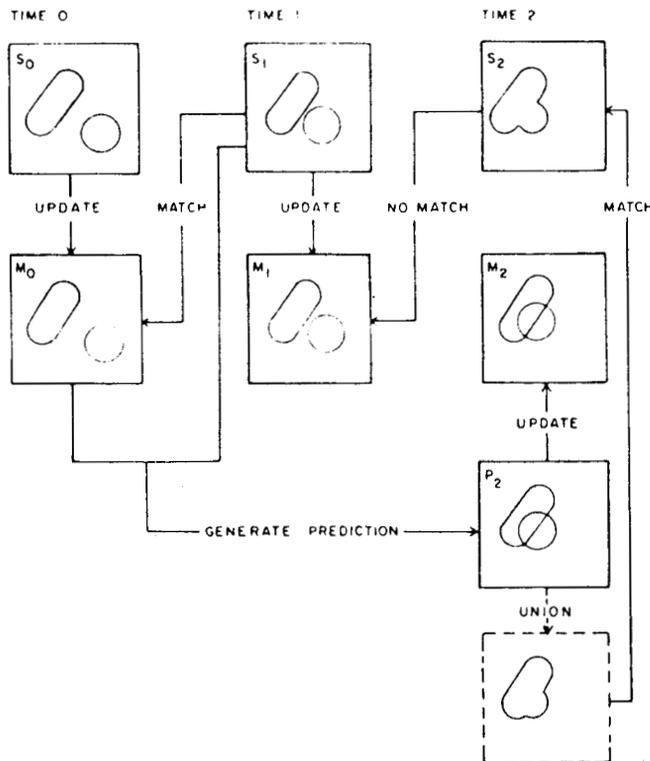
En météorologie, la direction et la vitesse des vents sont calculées en utilisant le déplacement des nuages. Les images, acquises par satellite géostationnaire, sont binarisées et divisées en secteurs (LEE 70). Chaque secteur, contenant les contours des nuages constitue un modèle que l'on recherche dans l'image suivante. L'origine du vecteur déplacement est le centre du secteur de la première image. L'extrémité correspondant au centre du secteur de la deuxième image. Cette méthode, simple et rapide, donne des résultats satisfaisants. Toutefois, dans le cas où les nuages sont composés de plusieurs couches possédant chacune leur propre déplacement, aucune mesure ne peut être effectuée.

Les travaux de AGGARWAL et DUDA (AGG 75) peuvent s'appliquer, de moins en théorie, à ce type de situation. Afin d'éviter des complications introduites par l'utilisation d'images réelles et pour permettre une plus grande généralité, un modèle idéal constitué de polygones a été utilisé. Chaque polygone rigide, opaque mais pouvant présenter des cavités, représente une couche de nuages. Afin d'analyser leur mouvements, tout en tenant compte des recouvrements des couches, les auteurs ont développé le concept de sommets "vrais" ou "faux". Les premiers étant des sommets appartenant réellement au polygone, les seconds quant à eux sont formés à partir des recouvrements intervenus. En l'absence de bruit, les sommets "vrais" peuvent être facilement identifiés, ces angles restant constants quand les polygones se déplacent. De plus, la valeur de ces angles est généralement inférieure à 180 degrés. La correspondance est établie à partir de la nature des sommets, des longueurs des côtés du polygone et de la valeur des angles.

On définit un modèle, réactualisé à chaque itération, en tenant compte des estimations sur les vitesses angulaires et linéaires. Cette méthode repose sur deux hypothèses qui interdisent une application plus universelle :

- le modèle doit être de forme polyédrique,
- les changements topologiques (c'est-à-dire l'apparition ou disparition des sommets) sont limités à deux à chaque prise de vue.

Ces restrictions ont été levées dans l'algorithme développé par CHOW et AGGARWAL (CHO 77), qui utilisent des descripteurs pour chaque image. Ces derniers, invariant en rotation et translation, sont la surface et les axes d'inertie. Tant qu'aucun recouvrement n'intervient dans la scène, on utilise un simple processus de correspondance basé sur ces descripteurs et sur un modèle. Dans le cas contraire, l'algorithme requiert l'emploi d'un modèle prédictif généré à partir du calcul des vitesses et de la description du modèle. L'image considérée est alors mise en correspondance avec la figure du modèle prédictif. Si cette opération s'avère correcte, le système suppose que les estimations de vitesse sont exactes et réinitialise le modèle en conséquence. Cette approche est illustrée sur la figure II.1. où l'utilisation du modèle prédictif s'effectue à l'instant  $t_2$ .



- Figure II.1. -

MARTIN et AGGARWAL (MAR 79), afin d'analyser le mouvement dans une séquence d'images de laboratoire composée de figures curvilignes, se sont intéressés à la fonction de courbure. Cette fonction, représentée sur un graphe avec pour abscisse la longueur du contour et pour ordonnée la pente de la tangente, caractérise la forme de l'objet indépendamment de sa position et de son orientation. Dans cette représentation, les segments de droite horizontaux du graphe correspondent aux lignes droites de l'image et les segments de droites obliques correspondent à des arcs de cercle dont la courbure est proportionnelle à la pente. Les deux images étant ainsi décrites, elles peuvent être comparées en étudiant les similarités dans la forme de leurs courbes, ce qui permet de déduire les composantes horizontales et verticales du déplacement.

Toutes ces méthodes se trouvent dans l'impossibilité d'établir une correspondance correcte dans le cas de changement d'échelle. Pour pallier cette difficulté, certains auteurs (TSU 80) ont introduit des homotéties en vue de constituer un modèle flexible (flexible template matching). Ces travaux analysent le mouvement d'une séquence d'images simples (un dessin animé) représentant un animal se déplaçant dans un environnement tridimensionnel. Dans un premier temps, on détermine une région  $R_1$  dans la première image et on recherche, dans la seconde image, les régions pouvant correspondre à  $R_1$ . Cette analyse est basée sur l'analyse de la position du centre de la région et sur son périmètre. La meilleure correspondance est donnée pour le déplacement le plus faible. Cette méthode est successivement employée pour chaque région de la première image. Lors d'un changement d'échelle, une procédure de modèle flexible est utilisée. Cette opération consiste à appairer de courts segments d'un modèle avec une figure présentant des distorsions partielles. Cette approche est montrée sur la figure II.2., sur laquelle les parties A, B, C et D de la figure II.2.a. peuvent être respectivement appariées avec A', B', C' et D' de la figure II.2.b.

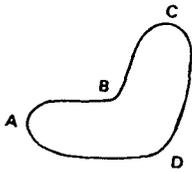


Figure II.2.a.

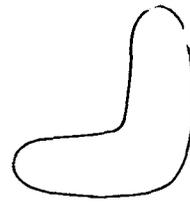
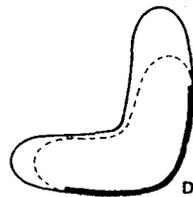
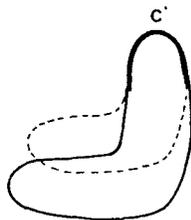
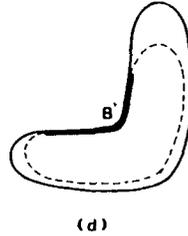
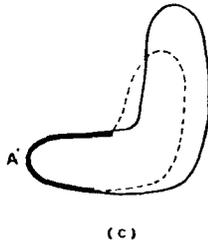


Figure II.2.b.



- Figure II.2. -

Une autre méthode, classique en Reconnaissance des Formes, voit son application s'élargir à l'étude d'images dynamiques. Il s'agit de la transformée de HOUGH qui est utilisée pour détecter les lignes et les courbes (DUD 72) et les symétries circulaires (MAI 84). La procédure de correspondance apparaît comme une généralisation de cette méthode. En considérant  $B = \{(X_i, Y_i)\}_{i=1}^n$  la liste des points de contour de l'objet et  $p = (X, Y)$  un point arbitraire, on définit cette transformée par  $H(B, p) = \{d_i\}_{i=1}^n$  avec  $dx_i = X - X_i$  et  $dy_i = Y - Y_i$

L'utilisation d'un accumulateur A, initialement fixé à 0, permet l'identification de l'objet, représenté par  $H(B, p)$ , dans une liste  $e_j = \{X_j, Y_j\}$ . Cet algorithme (AGG 81) est donné ci-dessous :

$V_{ej} = (X_j, Y_j) \quad e_j$ , points de contours de l'image candidate

$V_{di} = (d_{Xi}, d_{Yi}) \quad H(B,p)$ , points de l'objet

alors  $A(X_j + d_{Xi}, Y_j + d_{Yi}) = A(X_j, Y_j) + 1$

La translation recherchée est donnée par le maximum de l'accumulateur A. Cette recherche est toujours calculée par rapport à un point p, choisi généralement comme étant le barycentre de l'objet considéré. Les résultats sont les mêmes que ceux obtenus par corrélation binaire pour un temps de calcul de 30 à 50 fois plus rapide. Les rotations et changements d'échelle peuvent être, moyennant une augmentation du temps de calcul et de la mémoire nécessaire, facilement traités.

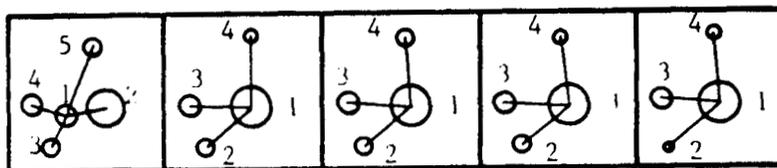
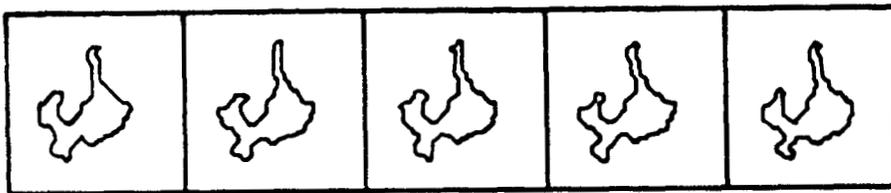
Toutes les méthodes développées présentent, à l'exception de celle utilisant la transformée de Hough, un niveau d'abstraction trop élevé pour pouvoir être généralisées au cas d'images réelles. Le bruit, les textures des objets et du fond, les variations d'illumination ainsi que les distorsions dues au mouvement seraient autant de facteurs qui affecteraient le processus de segmentation. La procédure de correspondance n'offrirait plus la fiabilité désirée.

La transformée de Hough, quant à elle, reste peu sensible aux bruits et s'utilise sur des images réelles en vue de la poursuite de cibles en temps réel (SAM 84). Les objets doivent cependant être suffisamment contrastés par rapport au fond. Cette condition n'est pas suffisamment vérifiée pour les scènes de trafic routier pour que cette méthode puisse être utilisée dans le cadre de cette étude.

### II.1.3. Correspondance entre structures

Les méthodes de correspondance entre structures ont pour but de décrire l'information visuelle au niveau de la structure de l'objet par un formalisme basé sur l'utilisation d'un graphe.

LEVINE, NOBLE et YOUSSEF (LEV 83), ont développé un algorithme original afin de détecter et d'analyser les changements de taille et de structure de cellules. Ces dernières sont d'abord modélisées par des polygones qui sont utilisés pour former un graphe (cf. Figure II.3.).



- Figure II.3. -

Un processus de correspondance est alors établi au niveau des graphes sur lesquels on calcule un indice de similarité qui intègre des paramètres tels que la surface, le périmètre et la position relative des polygones. La valeur de cet indice est élevée lorsqu'il n'y a pas de changement de structure important. Dans le cas où une mauvaise segmentation modifie cette structure, l'indice de similarité devient inférieur à un seuil. Une procédure tente alors de modifier la forme du graphe en rassemblant ou en séparant les polygones. Dans le cas où aucune amélioration ne peut être apportée, l'image est simplement ignorée. Si, pour une nouvelle acquisition la nouvelle valeur de l'indice reste toujours inférieure à ce seuil, un changement morphologique est alors détecté.

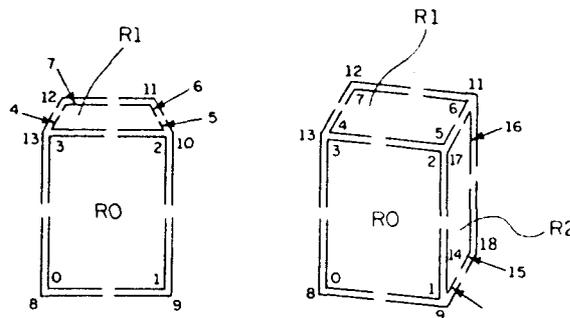
Cette approche permet d'obtenir des résultats satisfaisants. Le contraste élevé entre les cellules et le fond, facilite l'étape de segmentation et permet ainsi une description fiable de la structure de la cellule.

ROACH et AGGARWAL (ROA 79) considèrent des images composées d'objets polyédriques se déplaçant dans un environnement tridimensionnel en vue de déterminer les mouvements de ces objets par rapport à la caméra. Un pré-traitement permet d'extraire les contours et de localiser les sommets. Dans cette application la correspondance est établie par un système hiérarchique. Au niveau le plus haut, chaque objet est représenté par son barycentre dont la prédiction sur sa position dans l'image suivante permet de vérifier si ce mode de représentation est suffisant pour alimenter une procédure d'appariement. Dans le cas contraire, on fait une description grossière de l'objet (par exemple A est à gauche et en dessous de B). Les figures présentant les mêmes relations de proximité peuvent ainsi être associées. Lorsque cette description devient insuffisante, la correspondance s'effectue sur les positions relatives des faces du polyèdre.

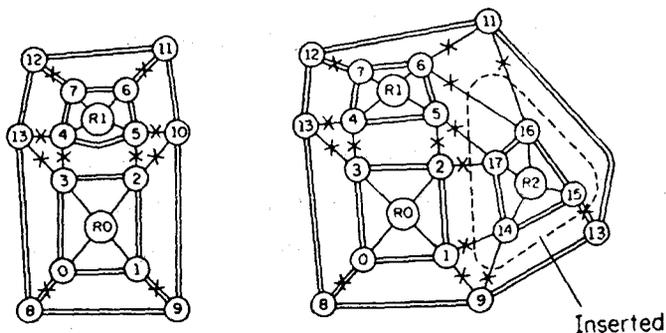
La complexité des mouvements autorisés est importante et les occlusions sont parfaitement détectées. Mais une généralisation de cette méthode au cas d'objets complexes et non uniformes ne semble pas envisageable. De plus, il apparaît que l'idée, intuitive, de représenter le mouvement en profondeur comme étant fonction du rapport de la surface de l'objet dans la nouvelle image sur sa surface dans l'image précédente, ne peut être utilisée pour toutes les classes d'objets.

Le déplacement d'objets tridimensionnels sur un fond uniforme a aussi été traité par JACOBUS, CHIEN et SELANDER (JAC 80). Un premier traitement permet de sélectionner les points significatifs. La description de l'objet est complétée par l'adjonction des contours, sommets et régions.

Afin d'établir une relation structurelle entre ces paramètres, les auteurs introduisent la notion de "demi-contour" (half chunk). Chaque demi-contour représente une partie de contour associé à une face de l'objet. Cette définition est illustrée sur la figure II.4.a., où l'on a associé à la face  $R_0$  les demi-contours (0,8), (1,9), (2,10) et (3,13). Cette représentation permet de décrire cinq types de connections à savoir : deux relations entre les points extrêmes des demi-contours, deux autres sur l'adjacence des régions, la dernière étant le point qui sépare ces deux régions. Ces connections permettent de représenter la structure topologique de l'objet par un graphe (figure II.4.b.) qui est utilisé dans un algorithme de correspondance.



- Figure II.4.a. -



- Figure II.4.b.

Toutes les structures n'autorisant qu'une seule association sont d'abord traitées. Ce processus individuel trouve les compatibilités entre voisins et permet de fusionner les structures. Cette étape utilise des mesures de similarité basées sur l'association ou la destruction des noeuds et des arcs.

Cette méthode permet d'analyser des situations complexes mais nécessite l'emploi d'images peu bruitées et bien contrastées.

Les méthodes basées sur la correspondance des structures analysent, à un niveau de compréhension élevé, l'évolution du mouvement. L'apparition (ou disparition) de nouvelles faces, les déformations et les occlusions peuvent être traitées. Mais ces algorithmes ne sont appliqués que sur des scènes simples et de bonne qualité composés d'objets de niveaux de gris homogènes et fortement contrastés. L'application de ces méthodes à un univers complexe ne semble pas envisageable. La principale raison résidant dans la difficulté d'effectuer une modélisation simple et fiable permettant la construction d'un graphe exploitable.

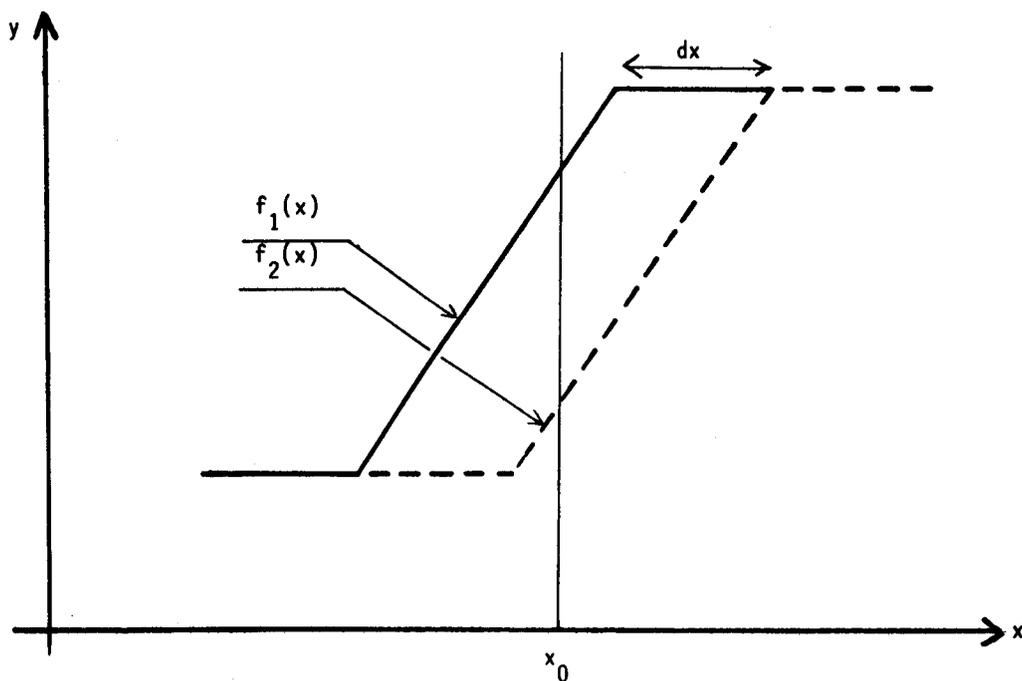
## II.2. DETECTION D'OBJETS EN MOUVEMENT PAR CORRESPONDANCE AU NIVEAU DES PIXELS

Dans ces méthodes, on considère l'image non plus comme une association d'éléments structurants, mais comme un ensemble de pixels. La comparaison s'effectue alors au niveau du pixel afin d'estimer un vecteur de déplacement. D'où leur nom de "Pixel based methods".

Comme aucune technique de reconnaissance de forme n'est nécessaire, l'application de ces méthodes s'adapte à des situations plus générales.

### II.2.1. Analyse du gradient spatio-temporel

L'analyse du gradient spatio-temporel utilise les changements d'intensité en un point de l'image dans le temps et dans l'espace, afin de déterminer le déplacement des surfaces constituant l'objet. Ce processus est illustré, dans le cas d'un déplacement mono-dimensionnel par la figure II.5., dans laquelle une surface caractérisée pour une courbe d'intensité  $f$ , se déplace vers la droite. La fonction  $f_1(x)$  représente cette courbe à l'instant  $t_0$ . La fonction  $f_2(x)$  représente cette même courbe à l'instant  $t_1 = t_0 + dt$  pour un déplacement  $dx$ .



- Figure II.5. -

Considérant un point  $x_0$ , choisi arbitrairement, on peut écrire la relation suivante :

$$f_2(x_0) - f_1(x_0) = p \, dx \quad \text{où } p \text{ représente la pente de la courbe.}$$

Dans le cas où le déplacement  $dx$  est faible, le développement au premier ordre de la série de Taylor peut être utilisé :

$$f_2(x_0) - f_1(x_0) = - f'(x_0) \, dx$$

$$\text{soit } \Delta f(x) = - f'(x_0) \, dx$$

L'extension au cas bidimensionnel nous donne :

$$\Delta f(x,y) = - \text{Grad } f(x,y) \cdot D$$

où  $D$  représente le vecteur déplacement.

En considérant deux points voisins, on obtient deux équations indépendantes qui permettent de trouver les composantes de ce vecteur déplacement. En pratique, cette opération est effectuée sur un nombre de points beaucoup plus important.

Dans ce cas, trois approches sont utilisées :

- Les méthodes d'optimisation locale, qui minimisent un terme d'erreur pour résoudre un ensemble d'équations linéaires (moindres carrés, pseudo-inverse) (THO 81b, LIM 75, LUC 81, DAV 83, YAC 83).
- Les méthodes d'optimisation globale, qui utilisent une procédures d'optimisation itérative (HOR 81, WOH 83).
- Les méthodes d'agrégation qui visent à regrouper les points de l'image possédant la même vitesse (CAF 76, FEN 79, THO 80)

Ces méthodes offrent une précision de l'ordre du pixel. Cependant, quatre sources d'erreurs viennent altérer les résultats à savoir :

- le bruit de mesure qui affecte la fonction intensité,
- les variations du gradient le long du contour,
- les variations des vecteurs déplacements,
- la variation de la luminance de l'objet.

Afin de minimiser leur importance KEARNES et THOMPSON (KEA 82) proposent de filtrer l'image (par lissage (YAC 83) ou défocalisation (FEN 79)), puis de seuiller le gradient. On remarque cependant que dans le cas d'un contour parallèle au mouvement, aucun calcul ne peut être effectué.

Ces techniques ont été appliquées sur des points significatifs (DAV 83, YAC 83).

LUCAS et KANADE (LUC 81) ont étendu cette méthode au cas de la stéréovision et ont développé un algorithme permettant d'extraire l'information profondeur.

Certains auteurs, quant à eux, se sont intéressés au développement au second ordre de la série de Taylor (PRA 81, NAG 83, SNY 80).

Les méthodes d'analyse du gradient spatio-temporel, appelées encore méthodes différentielles, n'autorisent qu'un déplacement très faible (quelques pixels) d'où la nécessité d'effectuer un échantillonnage temporel rapide et un échantillonnage spatial serré. Des techniques récursives, utilisables pour des déplacements plus importants ont été développées (BER 83). Cependant, ces algorithmes ne convergent pas rapidement, le nombre d'itérations nécessaires croissant rapidement avec l'ampleur du déplacement.

Les principales applications des méthodes différentielles concernent la transmission d'images, domaine dans lequel le but recherché est le codage de l'image qui, en supprimant les informations redondantes, diminue le débit global d'informations à transmettre (LAB 82, LIM 75, CAF 76, BER 83, CAF 83).

### II.2.2. Méthodes de Fourier

L'utilisation des méthodes de Fourier présentées par HUANG (HUA 81) s'appuie sur l'analyse de la phase de la transformée de Fourier.

Considérons  $g(x,y)$  l'image translatée de  $f(x,y)$  telle que :

$$g(x,y) = f(x-dx, y-dy)$$

Les transformées de Fourier  $F(u,v)$  et  $G(u,v)$  des fonctions  $f(x,y)$  et  $g(x,y)$  vérifient l'équation :

$$G(u,v) = F(u,v) \exp (- 2\pi j (udx + vdy))$$

La différence des phases entre ces deux fonctions se calcule aisément :

$$\Delta\phi(u,v) = \angle G(u,v) - \angle F(u,v) = - 2\pi (udx + vdy)$$

Deux couples de fréquence  $(u,v)$  permettent alors une détermination de  $dx$  et  $dy$ .

Le temps de calcul peut être considérablement réduit en étudiant la projection de l'image sur les axes, ce qui donne :

$$\begin{aligned} f_1(x) &= \int f(x,y) dy & f_2(y) &= \int f(x,y) dx \\ g_1(x) &= \int g(x,y) dy & g_2(y) &= \int g(x,y) dx \end{aligned}$$

Dans ce cas, la différence des phases est donnée par  $\Delta\phi_1(u)$  et  $\Delta\phi_2(v)$  :

$$\Delta\phi_1(u) = G_1(u) - F_1(u) = - 2\pi u dx$$

$$\Delta\phi_2(v) = G_2(v) - F_2(v) = - 2\pi v dy$$

où  $G_1(u)$ ,  $F_1(u)$ ,  $G_2(v)$  et  $F_2(v)$  sont les transformées de Fourier respectives de  $g_1(x)$ ,  $f_1(x)$ ,  $g_2(y)$  et  $f_2(y)$ .

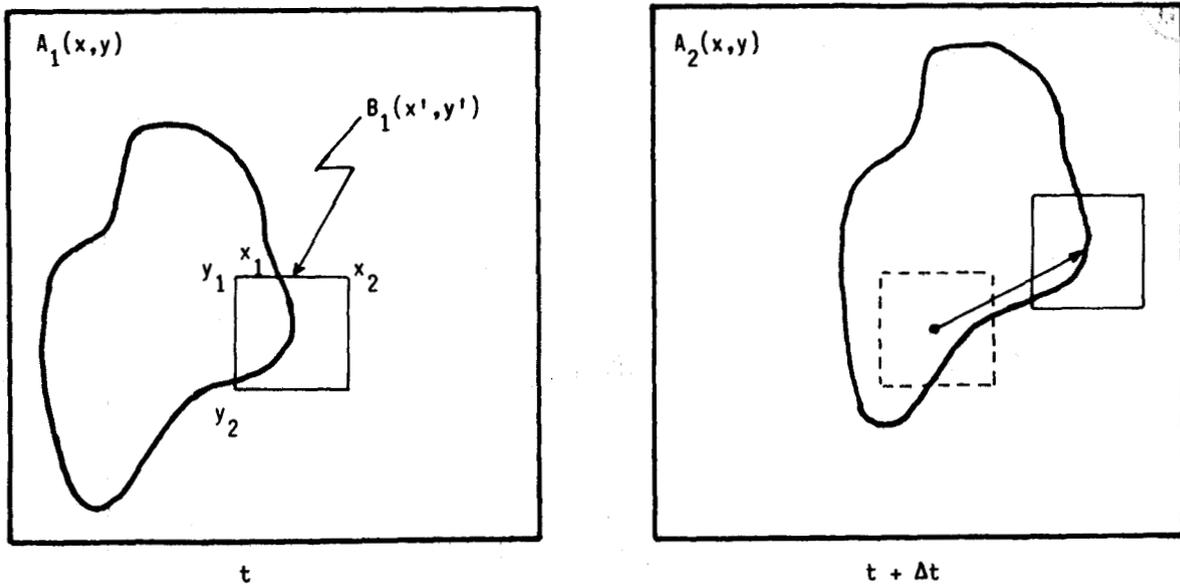
Notons que la phase est déterminée à  $2k\pi$  près et que cela impose une restriction sur la dimension du déplacement.

Malgré des temps de calcul rapides notamment grâce à l'utilisation de circuits câblés de transformée de Fourier rapide (FFT), ces méthodes restent peu employées. On ne peut que citer les travaux de ARKIN (MAR 78) qui traitent de l'estimation du mouvement des nuages. Les résultats obtenus ont montré la difficulté d'analyser le mouvement de façon cohérente. Les méthodes de corrélation semblent mieux adaptées à ce genre de situation.

Les méthodes de transformée de Fourier ne donnent des résultats satisfaisants que dans le cas où on dispose d'un fond uniforme. D'autre part, une extension de ces algorithmes aux rotations et aux changements d'échelles impliquent des transformations complexes qui ne rendent plus cette approche attractive du point de vue temps de calcul.

### II.2.3. Méthodes de corrélation

Les méthodes de corrélation consistent à trouver la correspondance d'une fenêtre de référence dans une autre image. Cette approche est illustrée sur la figure II.6.



- Figure II.6. -

Soient  $A_1(x,y)$  et  $A_2(x,y)$  les images prises aux instants  $t_1$  et  $t_2$  et  $B_1(x',y')$  la sous image incluse dans  $A_1(x,y)$ .

La taille de la fenêtre  $B_1$  est  $\Delta x, \Delta y$  avec :

$$\Delta x = |x_2 - x_1|$$

$$\Delta y = |y_2 - y_1|$$

Apparier  $B_1$  à une partie de  $A_2$  revient à rechercher le maximum d'une fonction de corrélation qui peut être formulée par la relation suivante :

$$C(x,y) = \frac{\sum_{i=1}^{\Delta x} \sum_{j=1}^{\Delta y} b_1(x'+i, y'+j) a_2(x+i, y+j)}{\sqrt{\sum_{i=1}^{\Delta x} \sum_{j=1}^{\Delta y} b_1(x'+i, y'+j)^2 \sum_{i=1}^{\Delta x} \sum_{j=1}^{\Delta y} a_2(x+i, y+j)^2}}$$

où  $b_1(x,y)$  représente le pixel de coordonnée  $x,y$  de la fenêtre  $B_1(x,y)$  et  $a_2(x,y)$  représente le pixel de coordonnée  $x,y$  de l'image  $A_2(x,y)$

Ces méthodes ont été utilisées dès 1970 afin de déterminer la vitesse et la direction des nuages à partir d'un satellite géostationnaire (LEE 70). Les résultats s'avéraient satisfaisants dans le cas où les nuages n'étaient composés que d'une seule couche. Cependant, il s'est avéré qu'une simple méthode de correspondance de contours binaires donnait des résultats similaires tout en diminuant les temps de calcul.

SPOER (SPO 83) emploie les mesures de corrélation sur des images de laboratoire simulant le déplacement d'un objet conjointement à celui du fond. Cette situation correspond au cas d'une caméra embarquée à bord d'un appareil volant et analysant une scène dynamique. La procédure développée permet de séparer l'objet du fond.

SQUALLI a utilisé ces mesures pour calculer le déplacement de petits flotteurs permettant la mesure de courantologie (SQU 83). Mais il lui a préféré la méthode relationnelle déjà mentionnée qui présente l'avantage d'être 4 à 8 fois plus rapide que la corrélation.

Un algorithme performant a été développé afin d'effectuer le suivi d'un véhicule dans une scène réelle (BER 80). Cet algorithme permet de traiter les recouvrements et prend en compte les changements de taille quand l'objet se rapproche ou s'éloigne de la caméra. Le modèle est alors éclaté en plusieurs parties et chaque sous fenêtre corrélée indépendamment. L'initialisation du processus nécessite cependant l'emploi d'un opérateur qui "pointe" l'objet que l'on désire poursuivre.

Ces méthodes de corrélation, bien que coûteuses en temps, sont très largement utilisées. Elles restent cependant très sensibles aux changements d'éclairement et de réflectance qui accompagnent généralement le mouvement. C'est pourquoi d'autres mesures de corrélation ont été développées parmi lesquelles nous citerons :

- La corrélation centrée, dans laquelle on soustrait l'intensité moyenne de la fenêtre de référence. Cette opération qui augmente la durée des calculs, tend à éviter de faire correspondre une région d'intensité moyenne faible avec une autre d'intensité moyenne élevée.

- La corrélation filtrée où l'on convolue les images avec un filtre spatial afin d'obtenir un pic de corrélation plus prononcé. La réponse impulsionnelle du filtre utilisé doit être choisie afin de maximiser la dynamique de ce pic.

La taille et la forme de la fenêtre jouent un rôle important (BUR 82). En effet, une fenêtre trop petite ne comporte pas d'éléments significatifs. Avec une fenêtre trop grande, les distorsions apparues empêchent un calcul fiable. Bien que la plupart des travaux utilisent une fenêtre rectangulaire, de meilleurs résultats seraient donnés en travaillant avec des fenêtres de type Gaussien dans lesquelles le poids le plus important est donné au centre de la fenêtre et décroît progressivement au fur et à mesure que l'on s'en éloigne. Les calculs trop coûteux en temps, empêchent généralement d'utiliser ce type d'approche.

La principale limite des méthodes de corrélation concerne la difficulté d'établir une correspondance entre un objet qui se déplace d'une zone claire à une zone sombre (zone d'ombre par exemple) ou inversement. Les reflets, dûs à la réflexion de la source d'éclairage sur une partie de l'objet (soleil sur le pare-brise d'une voiture) amènent des variations importantes sur les niveaux de gris de ces surfaces dans les différentes images constituant la scène. Une mesure de corrélation sur des fenêtres entourant ces zones apporte des résultats erronés et induit des erreurs dans la procédure de correspondance.

De plus, si l'on désire poursuivre un objet, il faut initialiser la procédure en le pointant manuellement à moins d'effectuer un découpage de l'image entière et de corréler indépendamment chaque sous-image. Cette approche ne semble pas attractive du point de vue temps de calcul.

## II.2.4. Différence d'images

### II.2.4.a. Sans image de référence

Une comparaison entre deux images successives montre, dans le cas d'une caméra immobile, que les régions qui diffèrent sont celles affectées par le mouvement. Cette constatation est la base de la technique d'utilisation de la différence d'images pour détecter le mouvement.

Pour une question de commodité, on appelle image courante (notée C) la dernière image acquise, et image précédente (notée P) celle qui la précède immédiatement dans la séquence d'images analysée. Soit  $D = C - P$  la différence pixel à pixel entre les images C et P.

La différence D entre deux images successives comportant un objet de niveau de gris homogène en mouvement fait apparaître des zones à important niveau de gris et des zones à faible niveau de gris. En travaillant en différence relative, on voit apparaître quatre zones distinctes dans le cas, le plus fréquent, où il y a chevauchement des deux positions de l'objet (cf. figure II.7.).

#### . ZONE 1

Une première zone, située "à l'avant" de l'objet dans le sens du déplacement est constituée de pixels qui appartenaient au fond dans l'image précédente et qui sont recouverts par l'objet dans l'image courante.

L'amplitude des différences de niveaux de gris est importante dans cette zone.

#### . ZONE 2

Une seconde zone, située "à l'arrière" de l'objet, est constituée de pixels qui appartenaient à l'objet dans l'image précédente et qui ont été découverts suite au mouvement, et appartiennent donc au fond dans l'image courante.

L'amplitude des différences de niveaux de gris est importante dans cette zone.

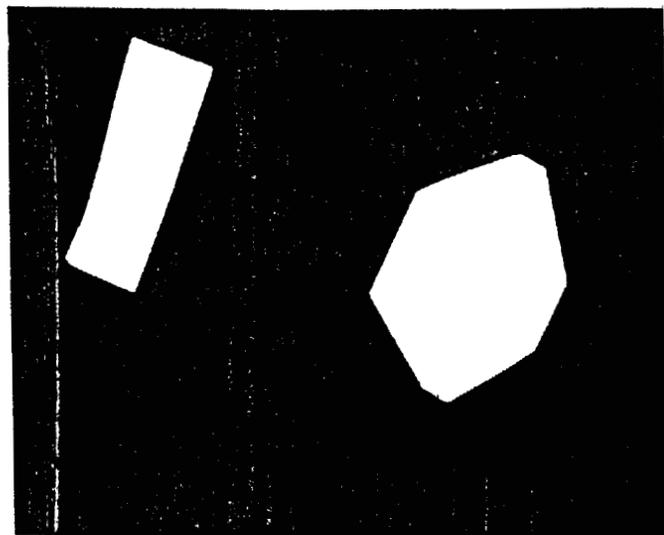
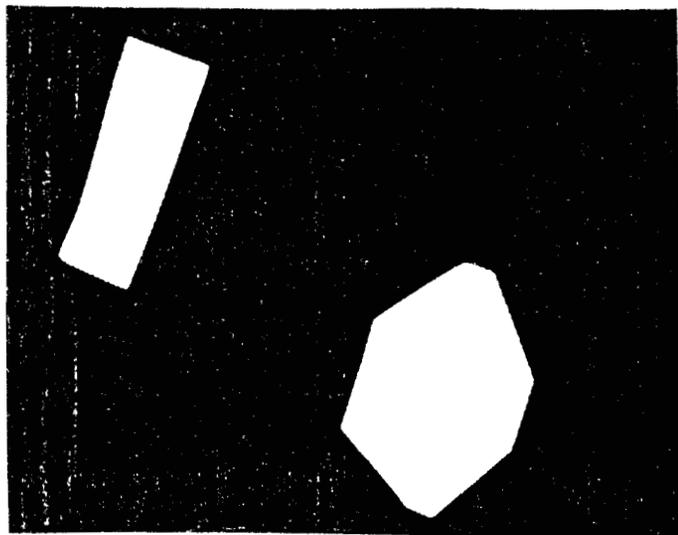
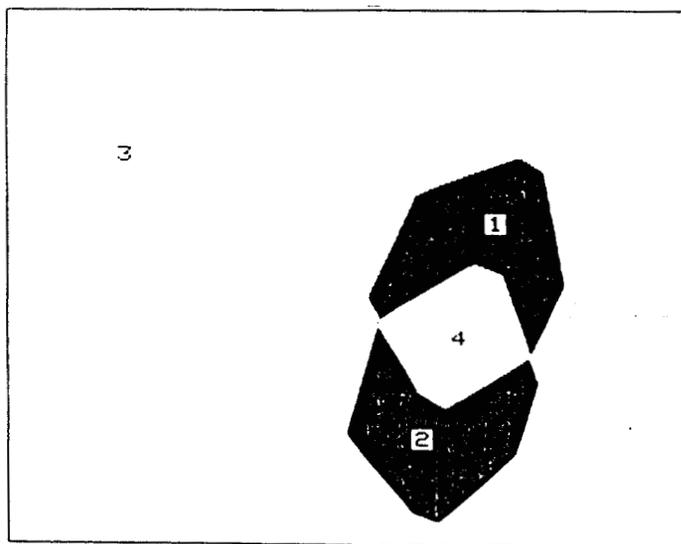


Image précédente P

Image courante C



Différence  $D = C - P$

FIGURE II.7.

. ZONE 3

Une troisième zone est constituée de pixels qui appartiennent au fond et aux objets stationnaires dans les deux images.

L'amplitude des différences de niveaux de gris y est faible.

. ZONE 4

Une quatrième zone, qui n'apparaît que lorsqu'il y a chevauchement des images successives de l'objet, est constituée de pixels qui appartiennent à l'objet en mouvement dans les deux images.

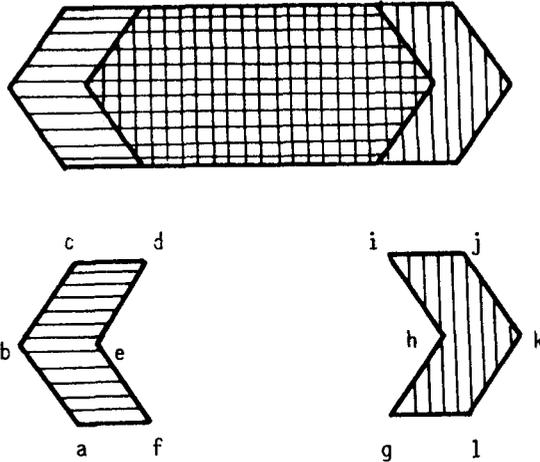
L'amplitude des différences de niveaux de gris y est faible.

Diverses techniques ont utilisé cette propriété pour caractériser les objets et leurs mouvements.

Une étude théorique a été menée sur des images comportant des objets polygonaux (YAL 80, YAL 82). Nous avons vu figure II.7. que l'image D fait apparaître l'existence d'un trou qui correspond à la partie commune aux images D et C (ZONE 4). La localisation de cette région permet d'utiliser un processus d'accroissement de régions (région growing) afin de réunir les zones 4 et 1.

La localisation de cette zone 4 peut être difficile dans le cas où la direction du mouvement coïncide avec l'orientation d'un ou plusieurs côtés de l'objet. Cette zone étant isolée, car non adjacente aux zones 1 et 3, des considérations d'ordre géométrique s'avèrent nécessaires. On détecte les paires de segments de droites collinéaires et d'égale longueur correspondant aux côtés des régions présentes dans l'image C. La longueur de ces côtés indique le déplacement du polygone. La direction du mouvement est donnée par leurs orientations.

Cette situation est illustrée sur la figure II.8. qui représente le polygone dans les images successives ainsi que l'image D, dans le cas d'un mouvement parallèle aux côtés horizontaux de l'objet.



- Figure II.8. -

Les paires de segments de droite collinéaires sont donnés par  $|(c,d),(i,j)|$  et  $|(a,f),(g,p)|$ . Les points  $(d,i)$  et  $(f,g)$ , qui bordent l'espace vide entre chaque paire de segments de droite sont appelés points de collinéarité. Le simple fait de joindre ces points sépare la partie commune aux deux images du fond. Cette zone déterminée, le processus d'accroissement de surface peut être utilisé.

Il consiste alors à examiner la luminance des points adjacents à cette région dans l'image réelle. Si leurs valeurs correspondent à la moyenne observée à l'intérieur de cette zone les points sont agrégés.

Ce procédé permet d'extraire les objets en mouvement avec une bonne précision. Mais les hypothèses sur ces derniers sont sévères. Ils doivent être polyédriques, homogènes et exempts de cavité.

Ces contraintes empêchent une généralisation de cette méthode dans le cas d'objets plus complexes.

Des techniques de classification de régions ont été traitées par certains auteurs. Celles-ci s'intéressent aux contours présents dans l'image courante C et l'image précédente P. Dans ces images binarisées, 1 représente un point de contour et une valeur nulle représente les autres points. Dans l'image D, binarisée elle-aussi, on associe à chaque point non nul une valeur correspondant :

- à la présence d'un point de contour dans C
- à la présence d'un point de contour dans P
- à la présence d'un point de contour dans C et P
- à la présence d'un point de contour qui n'appartient ni à C ni à P
- à la présence d'un point qui n'est pas un point de contour

En considérant le rapport donné par le nombre de points de contour appartenant à C sur le nombre de points de contour appartenant à P, on peut aisément déduire l'appartenance des régions considérées soit à la zone 1 (recouvrement du fond par l'objet), à la zone 2 (apparition du fond) ou à la zone 4 (recouvrement et apparition à la fois). On peut alors, par un simple processus d'aggrégation isoler l'objet en mouvement (JAI 79).

On remarque que dans le cas où l'on connaît les niveaux de gris des objets en mouvement et du fond (exemple objets noirs se déplaçant sur une zone blanche ou inversement) on peut utiliser le signe de la différence des images (YAC 78). Les notions de différence positive (DP) et différence négative (DN) sont introduites et vérifient les règles suivantes :

$$\begin{aligned} \cdot \text{ Si } C(x,y) - P(x,y) > T_1 & \text{ alors } DP = 1 \\ & \text{ sinon } DP = 0 \end{aligned}$$

$$\begin{aligned} \cdot \text{ Si } C(x,y) - P(x,y) < T_1 & \text{ alors } DN = 1 \\ & \text{ sinon } DN = 0 \end{aligned}$$

On voit que les régions présentes dans DP appartiennent aux objets en mouvement dans l'image C alors que celles détectées dans DN se rattachent à leurs positions dans P. Connaissant les zones communes, déterminées par la méthode précédente par exemple, on peut les réunir avec celles détectées dans DP afin d'avoir une représentation totale et immédiate des objets qui se déplacent dans la scène.

Cette méthode nécessite une connaissance a priori sur la nature des objets et du fond. Elle ne peut donc être utilisée dans le cas d'une scène de circulation routière.

Certains travaux, basés sur l'analyse des points de contours de D et de leur appartenance à l'image C ou P, permettent de décrire l'évolution de la scène par utilisation d'un arbre de classification (JAI 80, JAI 81.a., JAI 81.b.).

Plusieurs situations sont alors détectées à savoir :

- Recouvrement du fond par un objet
- Apparition du fond
- Occlusion
- Translation d'un objet
- Approche d'un objet
- Eloignement d'un objet
- Déplacement total d'un objet de son ancienne position
- Nouvelle position de l'objet totalement déplacé
- Indéterminé.

Cet algorithme ne permet pas une analyse complète de la scène mais apporte des informations utiles dans l'interprétation grossière des changements intervenus. Une analyse plus précise, dans les zones présentant un intérêt spécifique à l'application, peut alors être effectuée. De plus, les méthodes basées sur la classification des régions de l'image des différences demandent une binarisation de l'image courante C, de l'image précédent P et de l'image D.

La robustesse de ces méthodes est fonction du choix de trois seuils dont dépend le résultat final. Cette approche ne peut évidemment pas s'inscrire dans un cadre plus général visant à détecter et analyser les objets en mouvement dans une scène réelle où les variations d'éclairement sont importantes et où le seuillage des images brutes ne donne aucun résultat utilisable (WIL 81).

les méthodes présentées dans ce paragraphe reposent sur l'utilisation de deux images consécutives. Il peut cependant être intéressant d'utiliser plusieurs acquisitions. YACHIDA, ASADA et TSUJI (YAC 81) ont développé une méthode permettant de suivre le déplacement de cellules ou de poissons dans un bassin. Comme il s'avérait difficile de segmenter directement ces images du fait de la présence d'ombres (de niveau de gris intermédiaire entre le fond et les objets) et de bruit, il s'est avéré plus judicieux d'utiliser les changements temporels intervenus. L'image des différences entre les deux premières acquisitions est seuillée puis utilisée comme modèle auquel on réunit le résultat des différences des images successives. Ce processus permet de donner rapidement une représentation totale des objets en mouvement.

Dans un but de réduction de donnée en vue de la transmission d'images, une méthode similaire a été utilisée (LEN 83). Les différences  $D$  entre les images  $n-1$  et  $n$ , puis entre les images  $n$  et  $n+1$ , sont tout d'abord seuillées. Un simple ET logique entre ces deux résultats détermine, sans aucune ambiguïté, la zone commune. Un processus d'agrégation des régions donne la forme de l'objet en mouvement dans l'image  $n$ . Cette approche nécessite un échantillonnage temporel serré afin qu'existe un recouvrement entre les deux images. De plus, l'objet est détecté à l'instant  $n$ , soit avec un retard correspondant à une acquisition. Mais il faut souligner que les deux dernières méthodes présentées ne nécessitent que le seuillage d'une seule image, l'image des différences.

II.2.4.b. Avec image de référence

Nous avons vu que la principale difficulté des techniques de différences d'images concerne l'agrégation des régions 1 et 4, permettant ainsi la représentation totale de l'objet en mouvement.

Cependant, dans certaines applications, il peut être possible de disposer d'une image dans laquelle seuls les éléments stationnaires sont présents. Cette image est alors utilisée comme image de référence (notée R).

Sur l'image D', avec  $D' = C - R$ , on détecte immédiatement chaque objet introduit dans la scène. En effet, l'image D' ne contient que deux régions. La première (zone 1) est constituée de points résultant de la différence entre l'objet de l'image C et la partie recouverte par cet objet de l'image R. La seconde (zone 2), quant à elle, n'est formée que de la différence des points stationnaires présents dans les images C et R. Cette deuxième région constitue le fond de la différence des images et son niveau de gris est très faible. Ainsi, un simple seuillage permet de détecter immédiatement la position et la forme des objets en mouvement dans la scène.

Cette approche utilisée notamment dans le cadre de l'analyse de trafic routier (DIC 84, WAT 84), permet une compression de données importante et peut s'inscrire, de ce fait, dans une application temps réel. AOLI (AOK 84) applique cette même technique, mais se limite à l'analyse d'une ligne d'image unique. La référence est, dans ce cas, constituée de la valeur moyenne de la ligne exempte d'objets.

L'obtention de valeurs proches de zéro pour les points du fond de la différence des images, implique un éclairage de la scène constant. Cette condition n'est pas vérifiée dans le cas d'images extérieures où l'on observe des variations très rapides de luminosité (soleil caché par un nuage par exemple).

Aussi, l'étape de binarisation de l'image D' est basée sur l'utilisation de deux seuils distincts. Le premier, permet de sélectionner les points correspondants à un mouvement ; le second, quant à lui, détecte les petits changements de luminance du fond. Ces petites variations sont au fur et à mesure additionnées à l'image de référence R qui est ainsi réactualisée à chaque nouvelle acquisition.

Mais l'impossibilité, dans la plupart des cas, de disposer d'une image représentant le fond dépourvu d'objets non stationnaires, rend cette approche inexploitable. DRESCHER et NAGEL (DRE 78) ont levé cette contrainte en considérant deux images dans lesquelles le déplacement de l'objet en mouvement est important de manière à ce que leurs positions dans ces deux images ne se chevauchent pas. La première acquisition est utilisée comme référence.

La différence D', entre chaque nouvelle acquisition et l'image de référence R, fait apparaître des régions de type 1. On mesure les niveaux de gris des contours pour chaque région en se référant à l'image courante C. Lorsque la somme de ces niveaux de gris le long du contour atteint une valeur élevée, on détecte une zone appartenant à l'objet en mouvement. Cette approche impose un contraste important entre le fond et les objets en mouvement.

Chaque groupe de points, de taille suffisante, est considéré comme candidat et représenté par un ensemble de paramètres (valeur moyenne de l'intensité, surface, périmètre, centre de gravité....). Les objets de chaque image de la séquence peuvent ainsi être représentés par un vecteur multidimensionnel.

On suppose qu'entre deux images de la séquence, les paramètres mesurés sur le même objet ne présentent pas de variations importantes. Les deux points représentant le même objet dans deux images successives, seront alors très proches, alors que des points représentant des objets différents seront éloignés. Un simple critère basé sur la mesure d'une distance Euclidienne permet d'établir les correspondances entre les objets de deux images.

Les différentes limitations inhérentes à cette méthode (contraste important entre le fond et les objets, faibles variations d'éclairage pendant la séquence, décalage important entre les deux premières acquisitions) semblent interdire son application dans le cas d'une scène complexe.

TANG, SNYDER et RAJALA (TAN 82) ont développé une méthode permettant d'extraire les objets en mouvement dans une séquence d'images réelles, en recherchant une sous-image ou fenêtre de référence. Dans cette application, les deux premières acquisitions n'ont pas besoin d'être suffisamment espacées comme dans (DRE 78) et l'utilisation d'une image ne comportant que le fond et les objets stationnaires de la scène n'est pas requise comme dans (DIC 84, WAT 84). La différence entre deux images successives est simplement seuillée. Les zones ainsi détectées sont fusionnées selon un critère géométrique de proximité. Le rectangle encadrant ces régions ainsi mises en évidence forme la sous-image de référence R. Tant que l'objet apparaît dans cette fenêtre, la différence, des images successives de la séquence, présente des régions non nulles comprises à l'intérieur de R.

Le cas contraire indique que l'objet est entièrement sorti de la fenêtre. On connaît alors la nature du fond et en se référant aux images précédentes on peut extraire l'objet en mouvement.

D'autres travaux utilisent aussi cette notion de fenêtre de référence (JAI 84) en introduisant les notions de différences cumulées positives (DCP) ou négatives (DCN), dans lesquelles la première acquisition est considérée comme référence. Les différences s'effectuent toujours entre l'image courante et cette référence. Les points où la différence est positive (respectivement négative), sont ajoutés sous forme binaire à la DCP (respectivement DCN). Quand l'image binaire ainsi formée dans la DCP cesse de s'étendre en surface alors que l'image dans la DCN continue de croître, on peut affirmer que l'objet a quitté sa position initiale.

Comme dans la méthode précédente, l'objet en mouvement peut être facilement isolé et détecté dans les images courantes.

Ces deux dernières méthodes rejoignent, par une approche différente, celles développées par THOMPSON((THO 80), cf. paragraphe II.2.1.) et YACHIDA ((YAC 81), cf. paragraphe II.2.4.a.) dans lesquelles la segmentation ne s'effectue pas uniquement sur l'analyse des niveaux de gris. Dans ces méthodes, c'est la combinaison des changements intervenus dans la séquence d'images avec l'analyse des niveaux de gris qui permet d'isoler l'image des objets en mouvement. Cependant, ces méthodes nécessitent une phase d'apprentissage pendant laquelle les objets doivent quitter totalement leur emplacement d'origine. Cette approche semble difficilement exploitable dans le cas d'une scène complexe composée d'une quantité importante d'objets en mouvement.

Dans le chapitre suivant, nous allons proposer deux méthodes basées sur la différence d'images successives, permettant d'isoler les contours des objets en mouvement.

## CHAPITRE III

DETECTIONS DES CONTOURS DES OBJETS EN MOUVEMENT  
PAR TRAITEMENT MONO ET BIDIMENSIONNEL.

### III.1.SEGMENTATION DE SCENES DYNAMIQUES PAR ANALYSE DE SEQUENCES D'IMAGES

#### III.1.1. Choix d'une méthode de détection d'objets en mouvement

La diversité des approches permettant de détecter les objets en mouvement souligne l'absence d'une méthodologie universelle adaptée à tous les problèmes d'analyse de scènes dynamiques.

Dans notre application, la principale difficulté résulte de la complexité de la scène à analyser. Nous ne possédons aucune connaissance a priori sur le nombre d'objets présents dans la scène, sur leurs formes et sur leurs tailles. Les objets peuvent, entre deux prises de vue, apparaître dans le champ de vision de la caméra, comme ils peuvent en disparaître. Ils peuvent aussi présenter un arrêt momentané pendant plusieurs acquisitions et redémarrer de façon inattendue. Les déplacements observés entre deux prises de vue peuvent être importants (de l'ordre de 10 - 20 pixels), tout comme ils peuvent être faible (de l'ordre du pixel). Quant à leur trajectoire elle peut être soumise à des changements brutaux et imprévisibles.

A ces constatations d'ordre général, on remarque que l'image des objets en mouvement se déforme entre deux acquisitions, en raison de plusieurs facteurs :

- Les réflexions de la source d'éclairage sur une partie de l'objet (soleil sur pare-brise par exemple) qui présente, temporairement, des zones à niveau de gris élevé,
- Les effets de perspective relatifs à la position de la caméra par rapport à la scène,
- Les distorsions dues aux capteurs.

Le passage entre deux zones différemment éclairées (zone ensoleillée - zone d'ombre), le manque de contraste entre le fond et les objets au cours de leur déplacement et les variations importantes d'éclairage durant la séquence sont autant de facteurs supplémentaires qui conditionnent le choix d'une méthode permettant de détecter les objets en mouvement.

Il semble que, dans ces conditions, les méthodes utilisant l'analyse structurelle soient vouées à l'échec tout comme celles basées sur l'analyse du gradient spatio-temporel pour lesquelles seuls les déplacements faibles sont autorisés. Les méthodes de Fourier ne paraissent pas adaptées à ce type d'approche du fait de la difficulté d'analyser le mouvement de façon cohérente. Quant aux méthodes de corrélation, elles seraient trop coûteuses en temps s'il fallait analyser indépendamment toutes les sous-fenêtres constituant l'image, la corrélation de tous les objets de la scène n'étant pas envisageable dans nos hypothèses de travail.

Notre choix s'est, par conséquent, porté sur la différence d'images successives. Une telle approche permet une réduction importante des données à traiter en ce sens que seuls les objets présentant un mouvement sont détectés, le fond et les objets stationnaires étant éliminés. L'utilisation d'une image de référence n'est pas envisageable, tout comme la détermination d'une fenêtre de référence qui nécessite une phase d'apprentissage pendant laquelle l'objet doit quitter totalement son emplacement d'origine.

Dans ce chapitre, nous allons montrer comment la différence d'images successives permet, en se référant à l'image courante, d'isoler les objets en mouvement.

### III.1.2. Approche par analyse de régions

Nous avons vu au paragraphe II.2.4.a. que la différence d'images  $D$  comporte des indices sur les déplacements intervenus. Sur cette image  $D$ , nous avons observés 4 types de régions distincts, à savoir :

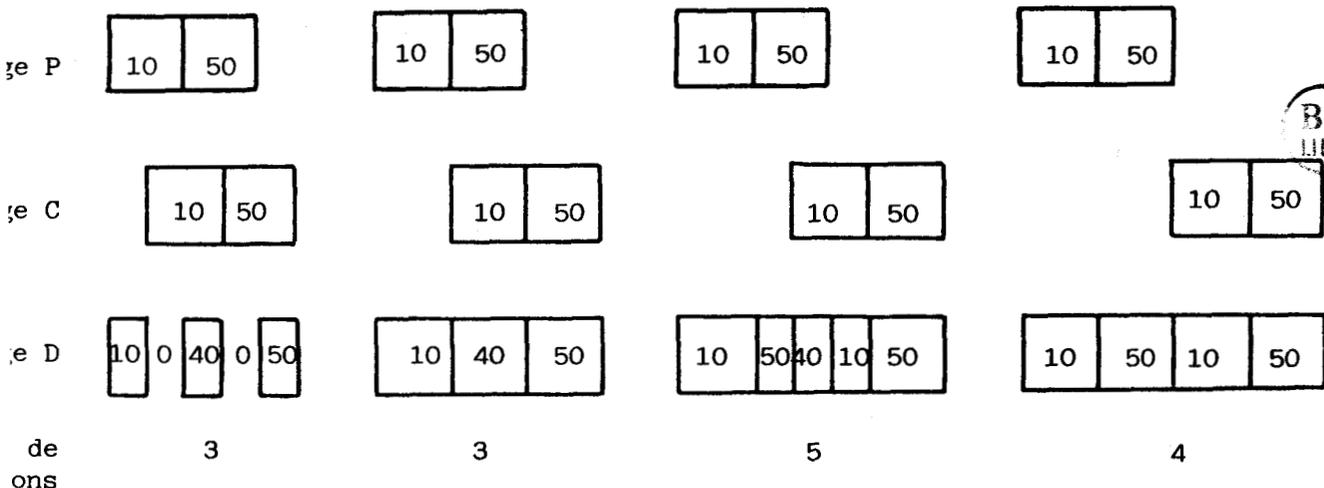
- zone qui appartenait au fond et qui est recouverte par l'objet,

- III.3. -

- zone qui appartenait à l'objet et qui fait apparaître le fond,
- zone constituant le fond et les objets stationnaires,
- zone constituée de pixels appartenant à l'objet en mouvement dans les deux images.

On pourrait envisager d'utiliser un processus de segmentation en vue d'isoler les régions constituant l'image et de vérifier leur appartenance à l'un des types de zones présentés ci-dessus. Mais l'un des problèmes relatif à la plupart des méthodes de segmentation concerne la détermination du nombre de classes présent dans l'image. Mis à part le cas des scènes simples où ce paramètre peut être connu (en cytologie par exemple), la détermination du nombre de classe est, comme pour toutes les méthodes d'agrégation, un problème pour lequel il n'existe pas de solutions entièrement satisfaisantes.

D'autre part, la notion de nombre de classes présent dans l'image des différences D, n'a pas de sens. Celle-ci est non seulement fonction des niveaux de gris des objets en mouvement, mais aussi de leurs déplacements relatifs. Cette propriété est illustrée sur la figure III.1. où 4 déplacements différents d'un objet constitué de 2 zones de niveaux de gris homogène (10 et 50) font apparaître des régions différentes en nombre et en niveaux de gris. La segmentation d'une telle image par une procédure de binarisation masquerait des caractères importants de l'image D.



- Figure III.1. -

### III.1.3. Approche par détection de contours

Plutôt que de traiter les zones de niveau de gris comme des indices sur les déplacements intervenus entre deux images successives, on peut s'intéresser aux transitions importantes présentes entre les régions. Ces transitions apparaissent comme des points de contours des objets en mouvement et appartiennent soit à l'image P, soit à l'image C, soit aux deux à la fois.

Plusieurs méthodes permettent de détecter les contours d'une image et ont fait l'objet d'une littérature abondante (DAV 75, ABD 79, BAS 79, GEU 83).

Sans vouloir entrer dans une étude approfondie, on peut cependant relever que les contours, dans une image numérisée, apparaissent comme un ensemble de points possédant une forte variation du niveau de gris moyen par rapport à leur environnement.

Certains travaux ont montré la possibilité d'extraire les contours lignes après lignes, un test de connexité permettant une segmentation totale de l'image (LEL 79, CAM 84). D'autres ont caractérisé un contour par un ensemble de paramètres (orientation, épaisseur, largeur, hauteur (ABD 79). Un ensemble de filtres, ou masque, correspondant aux diverses valeurs des données qui caractérisent un contour idéal, ont pu ainsi être déterminé. La convolution de l'image avec ces masques donne une réponse importante en présence des points de contour. Parmi ces masques, on relève les opérateurs de type gradient (Roberts, Sobel, Prewitt, Kirsch, Compass Gradient....) (BAS 79, ABD 79) et ceux du type Laplacien (Robinson, zéro crossing....) (SCH 78, HIL 83, GEU 83).

Chaque point de contour ainsi mis en évidence peut donner une information exploitable sur la position de l'objet dans l'image courante.

#### III.1.4. Principe des méthodes proposées

Les points de contour présents dans l'image D nous informent qu'un déplacement d'objets constituant la scène, s'est produit entre deux prises de vue.

Les points de contour dans l'image courante C indiquent la position de tous les objets de la scène.

En combinant les indications apportées par ces deux images, il est possible d'isoler les seuls points de contour en mouvement qui appartiennent à l'image courante.

Dans cette optique, nous allons présenter deux méthodes. La première, développée au chapitre suivant, repose sur une procédure monodimensionnelle dans laquelle l'aggrégation des résultats obtenus, ligne après ligne, permet d'isoler les contours des objets en mouvement. L'intérêt d'une telle approche s'inscrit dans une application basée sur l'utilisation d'une caméra linéaire. En analysant le mouvement le long d'une seule ligne, on peut obtenir une information suffisante sur le déplacement.

La seconde méthode, présentée au paragraphe III.3., s'inscrit dans un cadre plus général et utilise des procédures bidimensionnelles.

La robustesse de ces deux méthodes sera analysée au paragraphe III.4.

### III.2. DETECTION MONODIMENSIONNELLE

#### III.2.1. Possibilité et limites d'un traitement monodimensionnel

On peut envisager de traiter l'information contenue dans chaque ligne d'une image indépendamment des lignes voisines en vue d'analyser le mouvement des objets constituant la scène.

Nous avons vu (paragraphe II.2.4.b.) qu'afin d'évaluer le flux du trafic routier, AOKI (AOK 84) considère une ligne perpendiculaire à la chaussée. Cette ligne étant arbitrairement fixée, on peut aisément mémoriser le fond exempt d'objets. Cette mesure sera utilisée comme ligne de référence. La différence pixel à pixel entre la ligne appartenant à l'image courante et cette ligne de référence fait apparaître la présence (ou absence) d'objets sur la voie.

Cette différence est seuillée et la continuité des objets détectés, dans les acquisitions précédentes, vérifiée. L'information vitesse peut être facilement déduite en analysant deux lignes parallèles suffisamment espacées. En réinitialisant la (ou les) ligne(s) de référence, on tient compte des variations d'éclairage.

Le traitement s'effectuant sur une seule ligne, le nombre d'informations à traiter est beaucoup plus faible que dans le cas d'une analyse complète de l'image matricielle. Cette méthode peut, par conséquent, s'inscrire dans un traitement temps réel. Cependant, une telle approche ne permet de détecter que les déplacements perpendiculaires à la ligne traitée et limite, par conséquent son application à des situations simples.

Afin de détecter et poursuivre des véhicules en mouvement à partir d'images infrarouges, LE GUILLOUX (LEG 84) s'est intéressé aux lignes horizontales de même rang entre deux acquisitions. La densité des niveaux de gris de ces lignes est tracée et on remarque que les segments présentant la même structure au même endroit représentent les objets stationnaires. Les zones ne présentant pas de correspondant sont détectées et soulignent un changement dû au mouvement.

La méthode utilisée est basée sur une technique semblable à celle d'un éditeur de texte où la recherche des associations possibles entre deux caractères s'appuient sur l'utilisation d'une fonction de coût. Celle-ci est différente de zéro si les éléments sont distincts et nulle dans le cas contraire. Cette approche permet de ramener le problème de l'association de caractères à la recherche du chemin minimal dans un tableau bidimensionnel.

Dans le cas d'images dynamiques, l'association s'effectue entre les niveaux de gris de deux lignes de même rang appartenant à des images successives. Les disparitions et apparitions d'objets sont représentées respectivement par des segments horizontaux et verticaux dans le tableau bidimensionnel. Les segments parallèles à la diagonale correspondent à la partie dynamique présente dans la ligne. Le déplacement est quantifié en mesurant l'écart par rapport à la diagonale. Cette méthode est illustrée à la figure III.2. où l'image des déplacement indique les mouvements à droite et à gauche des deux objets mobiles constituant la scène.

La méthode semble parfaitement adaptée au type d'application traité par l'auteur en ce sens que seuls des mouvements horizontaux sont observés. Elle ne semble pas envisageable dans notre étude du fait de la diversité des déplacements autorisés (translations de directions quelconques et rotations).

Les travaux cités dans ce paragraphe ont souligné la possibilité de traiter une ligne d'image de façon indépendante. Cependant de telles techniques n'autorisent que la détection de mouvements parallèles (LEG 84) ou perpendiculaires (AOK 84) à la ligne traitée. Il semble intéressant de développer des procédures permettant de détecter, par un type d'approche similaire, des mouvements de direction quelconque par rapport à la ligne considérée. Ces procédures font l'objet des paragraphes suivants.

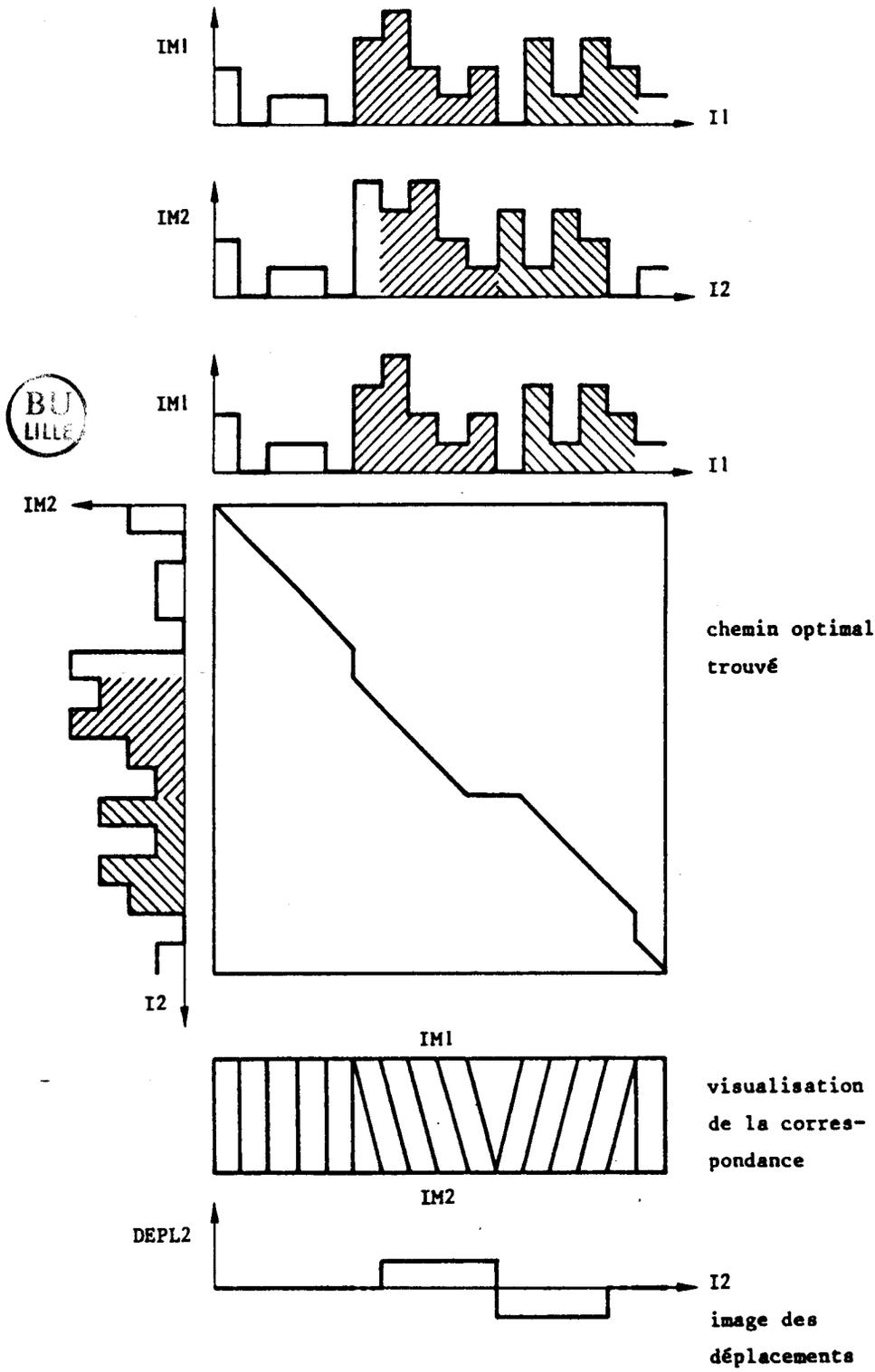


FIGURE III.2.

### III.2.2. Segmentation de l'image des différences

Les points de contour de l'image D sont des indices sur les déplacements intervenus (paragraphe III.1.3.). La détection de ces points peut s'inscrire dans un processus monodimensionnel basé sur l'analyse des niveaux de gris de chaque ligne. Différentes techniques peuvent être exploitées. La segmentation d'une ligne peut se formaliser par un automate d'état fini selon une table de transition basée sur l'évolution des niveaux de gris des pixels successifs (CAM 84). On peut ainsi relever les indices des transitions en particulier les débuts et fins de segments. Une telle approche monodimensionnelle, permet en outre, d'exploiter les techniques de traitement de signal utilisées notamment en génie biomédical (CHE 83, LEC 83).

Mais ne désirant pas une décomposition aussi complète de la ligne, nous avons choisi, dans un même ordre d'idée, d'extraire uniquement les points voisins présentant une discontinuité importante de leurs niveaux de gris. Ces transitions sont détectées selon le schéma suivant. On compare le niveau de gris du point de référence au niveau de gris du point suivant. Si la variation entre ces niveaux reste inférieure à un seuil, on agrège les deux premiers points et on compare le niveau de gris du troisième et de la référence inchangée. Si la variation du niveau de gris entre le premier et le troisième point reste inférieur au seuil, on agrège les trois premiers points et le processus est itéré. Dès que la différence de niveau entre le premier point et un point de la ligne dépasse le seuil, on détecte une transition en ce point et le processus est réinitialisé en prenant comme point de référence le niveau de gris du point de transition détecté.

L'algorithme de cette procédure de segmentation pour une image de dimensions (256 x 256) est donné ci-dessous :

Pour J de 1 à 256

TQ I < 256

Si VAL<sub>D</sub> (I,J) < Seuil

alors K = I

marquer début de segment = K

PIXVAL = VAL<sub>D</sub> (K,J)

TQ VAL<sub>D</sub> (K,J) = PIXVAL ± Seuil et K ≤ 256

K = K + 1

ftq

marquer fin de segment = K

I = K

Sinon I = I + 1

fsi

ftq

fpour

"VAL<sub>D</sub> (I,J)" représente le niveau du gris de l'image des différence D aux points de coordonnées I, J. "Seuil" indique les variations admises autour de la référence sans détection de transition.

Cet algorithme permet de définir les zones de mouvements encadrées par leurs points de transitions. Pour identifier la position des contours de l'objet dans l'image courante C, il convient de vérifier quels points de transition sont effectivement des points de contour de l'image courante.

Nous nous attacherons donc maintenant à ce problème.

### III.2.3. Extraction des points de contour dans l'image courante

La détection des points de contour dans l'image courante C, pourrait être effectuée par une procédure monodimensionnelle en segmentant cette image par un processus semblable à celui présenté au paragraphe précédent. Les points ainsi sélectionnés dans l'image courante C et dans l'image des différences D appartiendraient donc aux contours des objets en mouvement. Cette approche ne semble pas intéressante du point de vue temps de calcul, la scrutation de la ligne complète de l'image courante étant nécessaire.

Il est plus judicieux de baser notre approche sur un test d'appartenance aux points de contour de l'image C. La méthode développée s'inspire des travaux de BLANZ et REINHARD (BLA 81). Les auteurs, afin d'isoler des cellules dans une image matricielle ont placé une fenêtre de taille fixe (11 x 1) pour tous les points de l'image. Pour chacun de ces points, ils calculent la somme des niveaux de gris des pixels situés à gauche du centre, qu'ils soustraient à la somme des niveaux de gris des pixels placés à droite. Cette mesure apporte une information intéressante sur les frontières de l'objet. En effet, une valeur nulle indique que la fenêtre est placée dans une zone homogène en niveaux de gris. La mesure est maximale quand apparaît une transition entre deux zones de niveaux de gris différents. La présence d'un point de contour est alors détectée.

Le principe de scrutation des pixels de gauche et de droite à l'intérieur d'une fenêtre pour détecter un point de contour peut être retenu pour notre application. Dans un premier temps, nous avons mémorisé les points de transition de l'image D. Pour chaque point ainsi retenu, on place une fenêtre dans l'image courante et on vérifie si le point central de cette fenêtre est un point de contour. Plutôt que d'effectuer une mesure globale sur l'ensemble de la fenêtre comme dans (BLA 81), il est préférable, afin d'optimiser les calculs, d'effectuer des tests locaux successifs. Lorsque les conditions relatives à un test ne sont pas vérifiées, on arrête la procédure ; le pixel traité n'appartient pas à un point de contour de l'image courante.

La taille des objets constituant la scène étant faible, on a choisi de travailler sur une fenêtre de taille inférieure à celle utilisée par (BLA 81). Une fenêtre de taille 5 x 1 s'est avérée suffisamment porteuse d'informations. L'étude des deux pixels de gauche et des deux pixels de droite permet de détecter la présence d'un point de contour, ce dernier validant les conditions suivantes :

- Les deux points situés à droite du point analysé possèdent les mêmes niveaux de gris à un seuil près,
- Même condition pour les deux points situés à gauche,
- Les niveaux de gris des points de gauche sont différents de ceux des points de droite d'une quantité supérieure au seuil.

Ces conditions nous ont amené à développer certains tests dont la procédure, appliquée à tous les points sélectionnés dans l'image D de taille 256 x 256 pixels, est décrite ci-dessous :

Pour I = 3 à 254

Si VAL<sub>D</sub>(I,J) = début ou fin de segment

Pour G = 1,2

Pour D = 1,2

Si (VAL<sub>C</sub>(I-G,J) > VAL<sub>C</sub>(I+D,J) + Seuil ou  
< VAL<sub>C</sub>(I+D,J) - Seuil)

Si (VAL<sub>C</sub>(I+2) + Seuil >> VAL<sub>C</sub>(I+1,J) >> VAL<sub>C</sub>(I+2,J) - Seuil)

Si (VAL<sub>C</sub>(I-2,J) + Seuil >> VAL<sub>C</sub>(I-1,J) >> VAL<sub>C</sub>(I-2,J) - Seuil)

Alors VAL<sub>C</sub>(I,J) est un point de contour

fsi

fsi

fsi

fpour

fpour

fsi

fpour

"VAL<sub>C</sub>(I,J)" et "VAL<sub>D</sub>(I,J)" représentent le niveau de gris du pixel de coordonnées I,J respectivement dans l'image des différences D et dans l'image courante C.

#### III.2.4. Résultats

Les procédures exposées dans les paragraphes précédents ont été appliquées sur des images réelles. Les images III.3. et III.4. représentent respectivement l'image précédente P et l'image courante C. Entre ces deux prises de vues, la voiture s'est déplacée vers la droite et le piéton vers la gauche. Tous les autres objets de la scène sont restés stationnaires. Pour atténuer les variations d'éclairage entre deux acquisitions, les images subissent une égalisation d'histogramme avant tout traitement.

En appliquant sur cette scène les différentes procédures développées dans les paragraphes précédents, on obtient l'image III.5. La valeur du seuil dans l'étape de segmentation (paragraphe III.2.2.), ainsi que dans l'étape d'extraction des points de contour (paragraphe III.2.3.) a été arbitrairement fixée à 13 sur une échelle s'étageant du niveau 0 au niveau 256.

Sur l'image III.5. apparaissent clairement les points de contour appartenant au véhicule et au piéton dans leur position dans l'image courante. Les autres points ont été éliminés à l'exception de quelques fausses détections correspondant aux points isolés. Ceux-ci peuvent être aisément éliminés en utilisant les techniques de lissage.

La méthode qui, rappelons-le, est basée sur l'utilisation de procédures monodimensionnelles, permet de lever les contraintes relatives aux mouvements analysés par les méthodes de LE GUILLOUX et AOKI (LEG 84, AOK 84). En effet des translations de direction quelconque ainsi que des rotations et changements d'échelles sont instantanément détectés par les procédures développées dans ce chapitre.

La principale limitation inhérente à cette méthode concerne l'ajustement des seuils. Une faible variation autour de la valeur sélectionnée dans l'exemple, modifie fortement l'image des objets en mouvement. Aussi, il semble préférable d'utiliser une telle approche dans des applications où le choix du seuil est aisé (scènes fortement contrastées et éclairage constant). Dans les paragraphes suivants, nous allons étudier une méthode plus robuste, dans laquelle l'ajustement d'un seuil n'est pas critique.

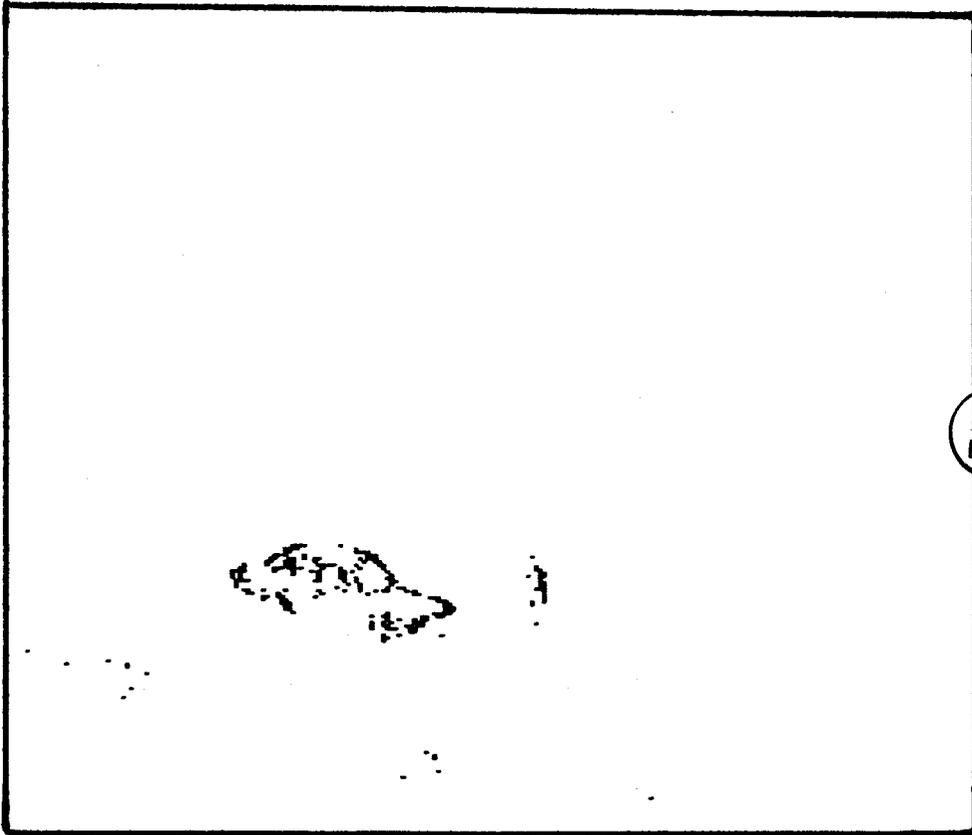


BU  
LILLE

- Figure III.3. -



- Figure III.4. -



BU  
ILLE

FIGURE III.5.

### III.3. DETECTION BIDIMENSIONNELLE

#### III.3.1. Multiplication binaire

Nous avons vu (par III.1.3.) que l'application d'un opérateur de type gradient, noté  $G(.)$ , à la différence  $D$  entre deux images successives permet de détecter, sur l'image  $G(D)$ , les contours d'un objet en mouvement dans ses positions occupées dans l'image courante ( $C$ ) et dans l'image précédente ( $P$ ).

Pour distinguer les points de contour relatifs à l'objet dans sa position courante de ceux relatifs à l'objet dans sa position précédente, les contours détectés dans l'image  $G(D)$  sont comparés à tous ceux détectés dans l'image  $G(C)$  obtenus en appliquant le même opérateur de type gradient  $G(.)$  à l'image courante  $C$ .

Les niveaux de gris des images  $G(D)$  et  $G(C)$  pourraient être comparés à un seuil pour sélectionner les pels représentatifs des contours. Un simple ET logique entre les images  $G(D)$  et  $G(C)$  ainsi binarisées permettrait de ne retenir que les pels relatifs aux contours des objets en mouvement dans leur position courante.

Cette approche est illustrée sur les figures représentant une scène simple comportant des objets contrastés et de niveaux de gris homogène. Les images  $G(P)$ ;  $G(C)$  et  $G(D)$  binarisées, sont représentées sur les figures III.6.a., III.6.b. et III.6.c. La figure III.6.d. montre le résultat du "ET logique" entre ces deux images. Cette approche permet d'isoler de façon satisfaisante les contours de l'objet en mouvement dans l'image  $C$ .

Cependant, une telle méthode n'est applicable que dans le cas de scènes présentant un contraste élevé entre le fond et les objets en mouvement. Dans le cas de scènes réelles, la réponse d'un opérateur de type gradient n'est pas constante tout au long d'un contour qui présente toujours des fluctuations au niveau des variations spatiales des niveaux de gris. De nombreux points de contour peuvent donc être éliminés lors d'une binarisation par comparaison à un seuil fixe. L'ajustement de ces seuils est donc délicat et il peut être nécessaire de procéder à une fermeture de contour.

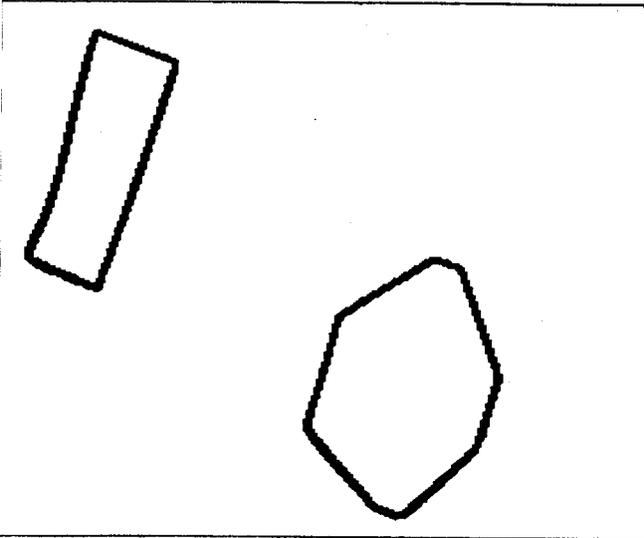


FIGURE III.6.a.

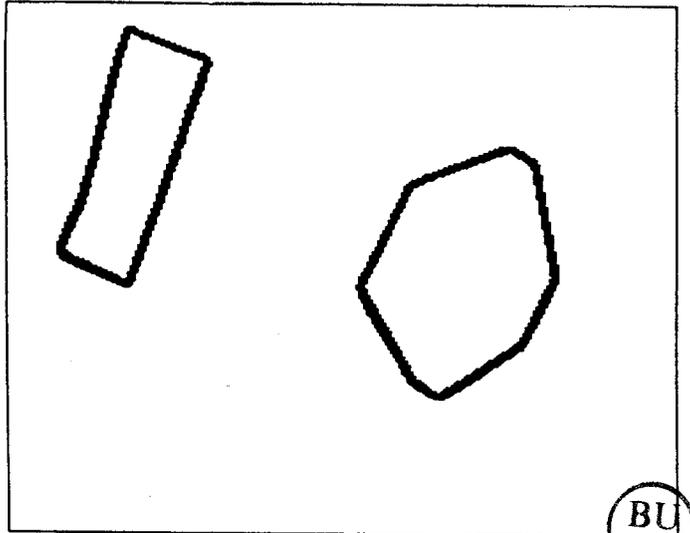


FIGURE III.6.b.

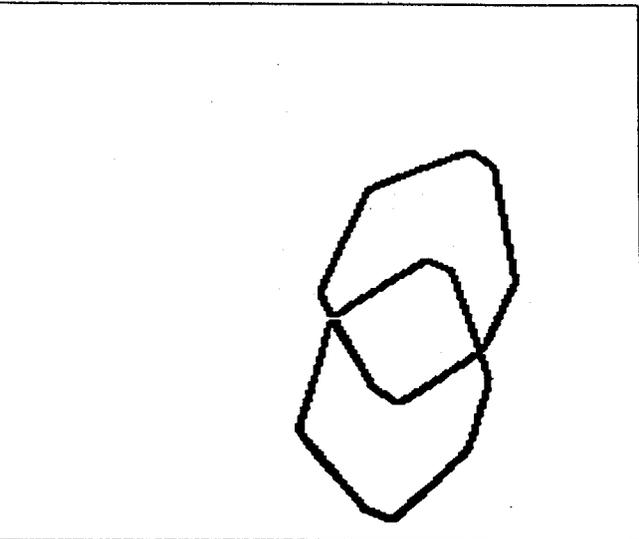


FIGURE III.6.c.

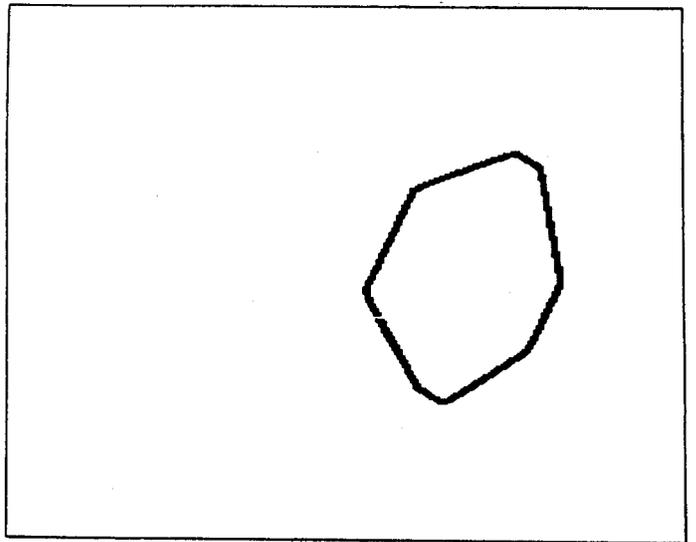


FIGURE III.6.d.

### III.3.2. Multiplication des niveaux de gris

Afin de rendre la méthode plus robuste et plus facile à implanter, on ne cherche pas à binariser les images  $G(D)$  et  $G(C)$ , mais à seuiller uniquement l'image résultant d'une opération numérique entre  $G(C)$  et  $G(D)$ .

On peut considérer que  $G(D)$  fonctionne comme un détecteur de mouvement. L'image  $G(C)$ , quant à elle, peut être assimilée à un détecteur de position. Pour qu'un point soit sélectionné, il suffit que ces deux détecteurs affichent, indépendamment, une valeur suffisante. Dans le cas où l'une au moins des deux valeurs est faible, il est évident que le point traité ne peut être choisi.

L'opération numérique qui répond le mieux à ces constatations est la multiplication pixel à pixel des niveaux de gris de ces images notée  $(G(C) \times G(D))$ .

L'extraction des contours en mouvement dans leur position courante s'effectue alors par seuillage de l'image  $(G(D) \times G(C))$ . Cette approche offre l'avantage majeur de ne nécessiter qu'un seul seuil dont la détermination est analysée sur un exemple au paragraphe suivant.

Différents types de gradient ont fait l'objet d'expérimentations. En désirant conserver une application temps réel, nous avons exclu les opérateurs de type KIRSH, CHEN et FREI, GRADIENT BOUSSOLE, PREWITT etc,... qui nécessitent un nombre de calculs trop important (opérateurs composés de 8 ou 9 masques). Nous nous sommes donc limités aux opérateurs de ROBERTS, PREWITT et SOBEL qui ont été successivement utilisés sur les images communes à la méthode exposée dans le paragraphe III.2.

Les images  $G(C)$  et  $G(D)$  définies précédemment sont représentées dans le cas des opérateurs de ROBERTS, PREWITT et SOBEL, respectivement par les figures III.7.a., III.7.b. - III.8.a., III.8.b. et III.9.a., III.9.b. Sur les images  $G(C)$  apparaissent les contours de tous les objets de la scène. Sur l'image  $G(D)$ , on distingue les contours de chaque objet en mouvement dans ses positions dans l'image P et dans l'image C. La multiplication pixel à pixel  $\{G(C) \times G(D)\}$  des images  $G(C)$  et  $G(D)$ , suivie d'une binarisation pour un seuil optimal, ne laisse apparaître que les contours des objets en mouvement correspondant à leur position dans l'image courante C (figures III.7.c., III.8.c., III.9.c.).

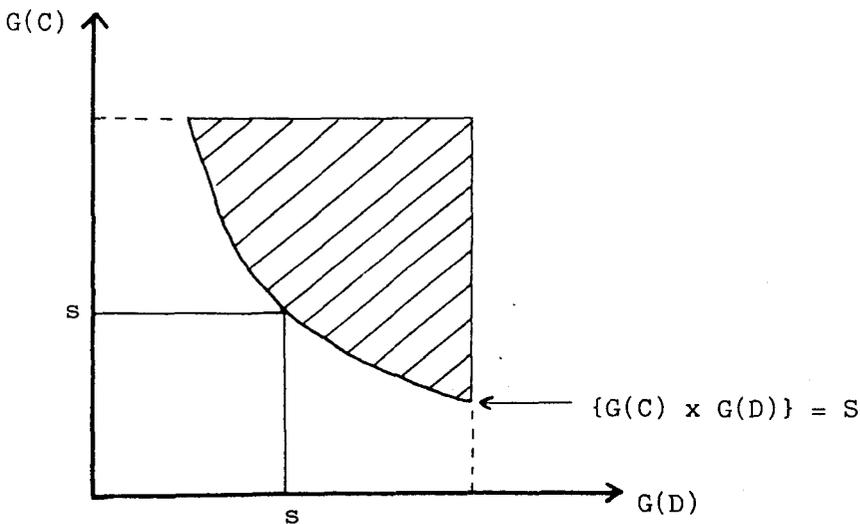
### III.3.3. Binarisation de l'image multiplication

Le seuillage de l'image  $\{G(C) \times G(D)\}$  permet d'extraire les contours des seuls objets en mouvement. Cette opération vérifie la relation suivante :

$$|G(C)| \times |G(D)| \geq S$$

dans laquelle  $S$  représente la valeur du seuil.

Les points répondant à cette condition sont situés dans la partie hachurée du graphe de la figure III.10.



- Figure III.10. -

- III.20. -  
OPERATEUR DE ROBERTS

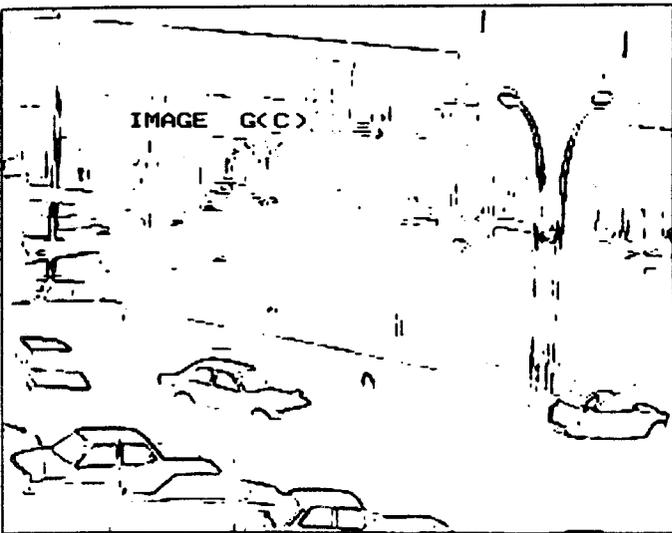


FIGURE III.7.a.

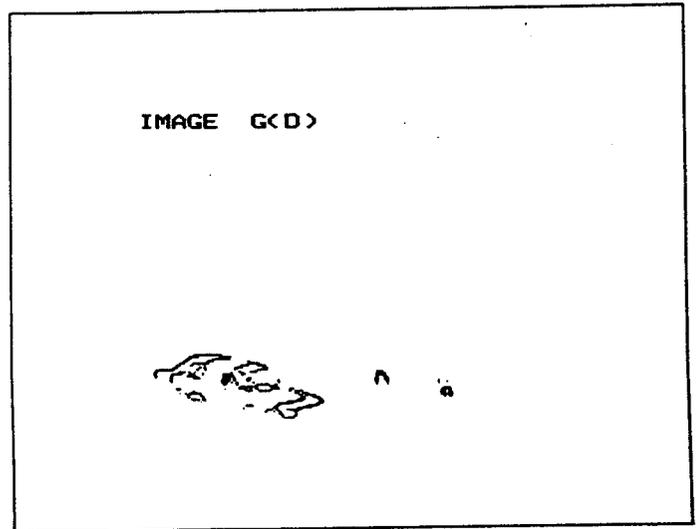


FIGURE III.7.b.

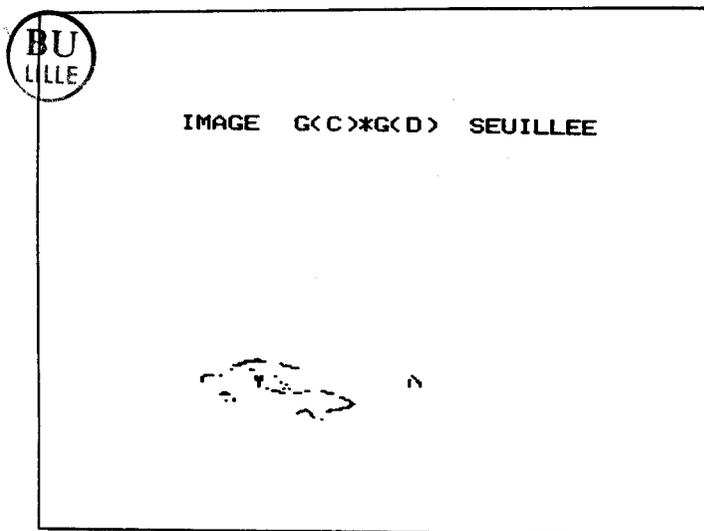


FIGURE III.7.c.

OPERATEUR DE PREWITT

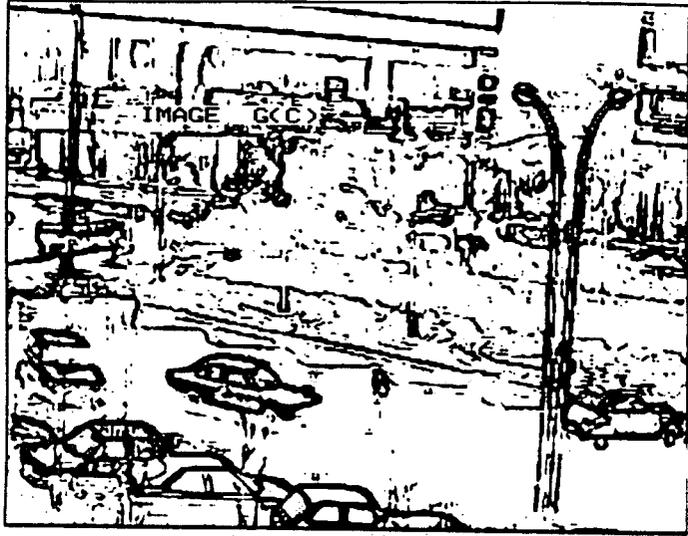


FIGURE III.8.a.

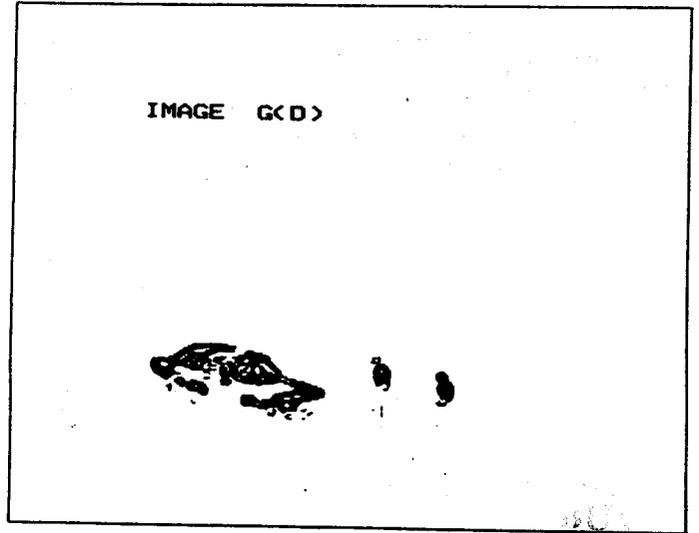


FIGURE III.8.b.

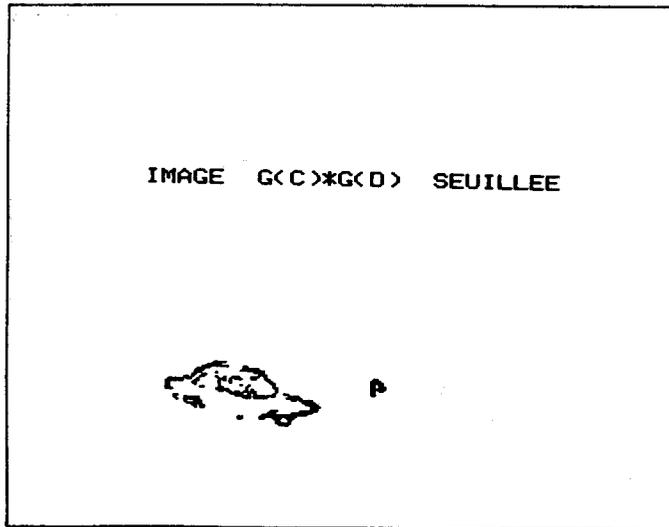


FIGURE III.8.c.

OPERATEUR DE SOBEL



FIGURE III.9.a.

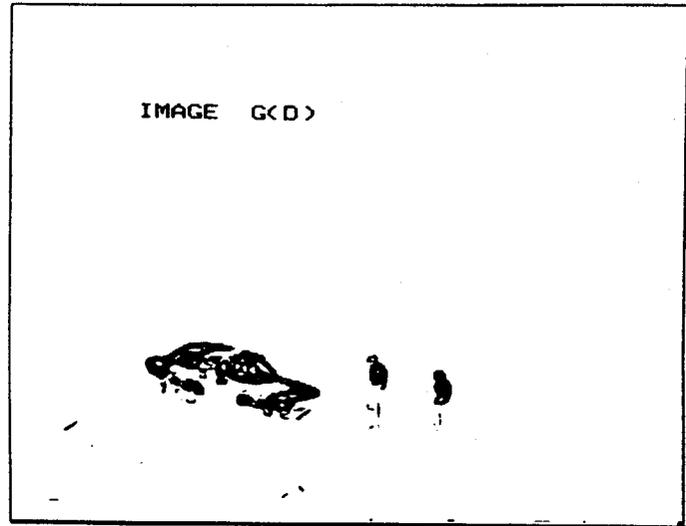


FIGURE III.9.b.

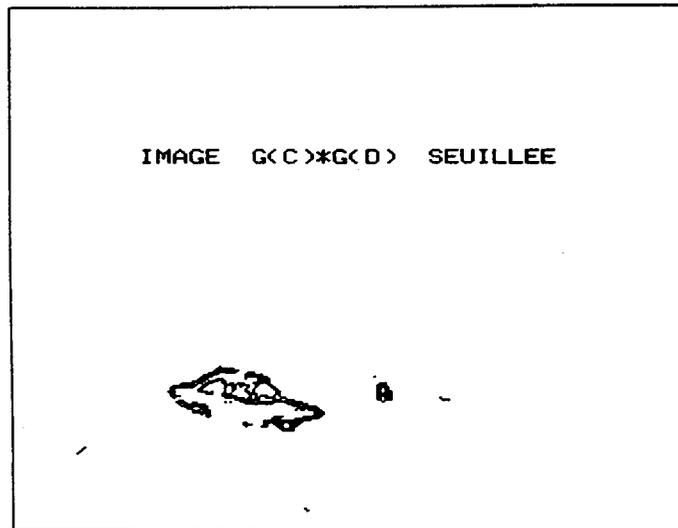


FIGURE III.9.c.

Il faut noter que pour ne pas dépasser la capacité de digitalisation du système (256 niveaux), les niveaux de gris de chaque image sont divisés par un coefficient avant multiplication. La valeur de celui-ci est de 4 pour l'opérateur de ROBERT et de 16 pour ceux de PREWITT et SOBEL, ce qui correspond, du point de vue de l'implantation matérielle, à un décallage à droite de 2 à 4 bits. Cette opération présente en outre l'avantage d'augmenter la dynamique de l'image  $\{G(C) \times G(D)\}$ .

La diversité des scènes rencontrées en traitement d'images a vu s'épanouir une multitude de méthodes de seuillage. L'ajustement du seuil est en général un problème très délicat qui conditionne les résultats de l'analyse ultérieure de l'image. La méthode la plus courante se réfère à l'analyse de l'histogramme des niveaux de gris de l'image dont il s'agit de rechercher les vallées. Mais lorsque les modes correspondant aux régions à détecter se superposent, l'histogramme devient inexploitable. On peut alors faire appel à une connaissance a priori sur la taille des objets à détecter [DOR 62]. On peut aussi diviser l'image en fenêtres sur lesquelles on effectue un seuillage local. La principale difficulté consiste à restituer une image binarisée globale cohérente [CHO 72]. La méthode de relaxation permet, quant à elle, d'améliorer la séparabilité des niveaux de gris mais repose sur l'hypothèse d'une certaine homogénéité des régions à mettre en évidence.

Dans notre approche, le choix d'une méthode s'avère délicat en ce sens qu'aucune forme ou répartition spatiale particulière n'est recherchée mais seulement un ensemble de points plus ou moins connexe. Nous allons montrer qu'il n'est pas nécessaire de développer une procédure de seuillage particulière. L'ajustement d'un seuil n'est pas critique, ce qui facilite son choix. Ces affirmations s'appuient sur une expérimentation sur des scènes variées, autant dans le monde réel que dans celui créé en laboratoire. Afin d'illustrer notre propos sur un exemple, nous utiliserons l'image  $\{G(C) \times G(D)\}$ , avec comme opérateur celui de SOBEL. L'histogramme d'une telle image (figure III.11.) montre que les niveaux de gris associés aux contours ne forment pas un mode différencié. Cependant, le résultat de la binarisation avec plusieurs seuils s'étageant du niveau

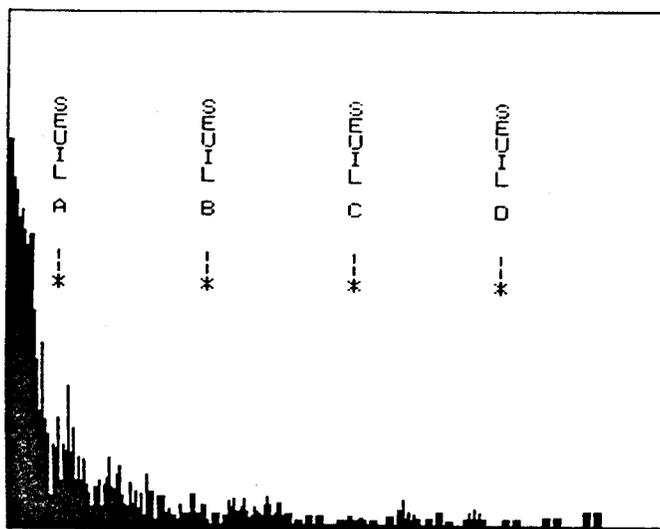


Figure III.11



Figure III.12.a

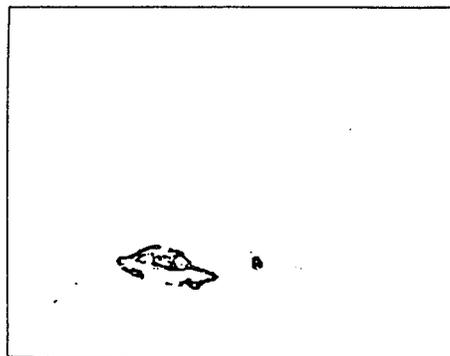


Figure III.12.b

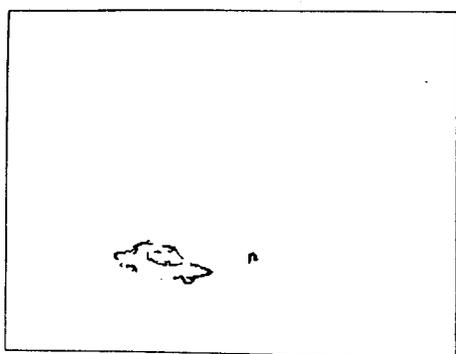


Figure III.12.c

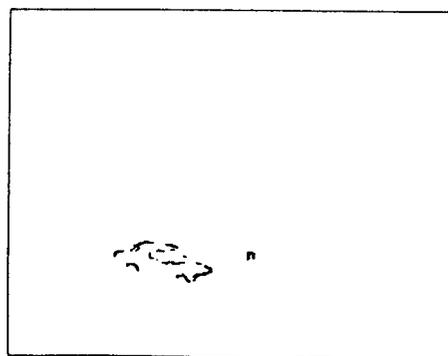


Figure III.12.d

23 (seuil A) au niveau 190 (seuil D) (0 représentant le blanc et 256 le noir) montrent la robustesse de la méthode (figures III.12.a., III.12.b., III.12.c. et III.12.d.).

En effet, il faut descendre jusqu'à un seuil de l'ordre de 30 pour faire apparaître des contours qui n'appartiennent pas à des objets en mouvement, ce qui correspond à un niveau de seuillage très proche du blanc. Ainsi, il apparaît que les niveaux de gris associés aux contours sont nettement plus élevés que les niveaux de gris parasites dûs aux bruits de mesure et aux variations d'éclairage sur le fond et les objets stationnaires.

La méthode proposée peut être comparée à celle développée par HAYNES et JAIN (HAY 83) qui utilisent le produit de l'image courante et de la différence des images sans toutefois introduire la notion de gradient de la différence des images.

La comparaison de l'histogramme de la figure et de l'histogramme du produit  $G(C) \times D$  obtenu à partir des mêmes images selon la méthode de HAYNES et JAIN (figure III.13.) montre que cette dernière a tendance à réduire la dynamique de l'image produit. Les résultats de diverses tentatives de binarisation avec différents seuils confirment cette constatation et montrent que cette méthode est très sensible au choix du seuil (cf. Figures III.14.a., III.14.b., III.14.c. et III.14.d.).

#### III.4. ROBUSTESSE DES METHODES PROPOSEES

Nous avons présenté deux méthodes d'analyse de séquences d'images qui permettent de détecter les contours d'objets en mouvement et de déterminer leur position instantanée dans une scène dynamique.

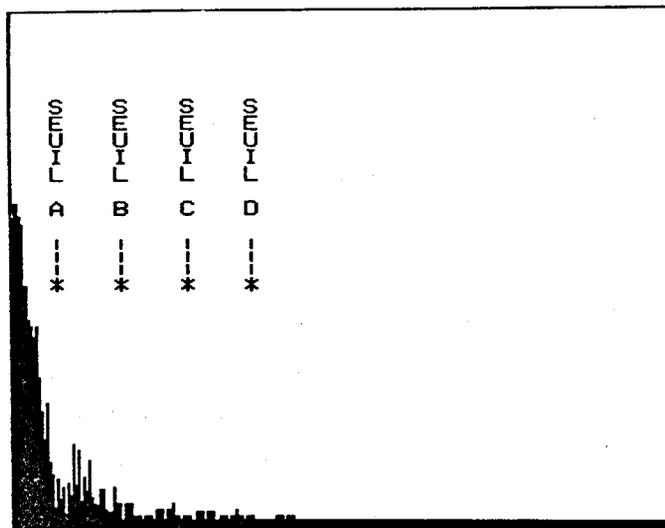


Figure III.13.

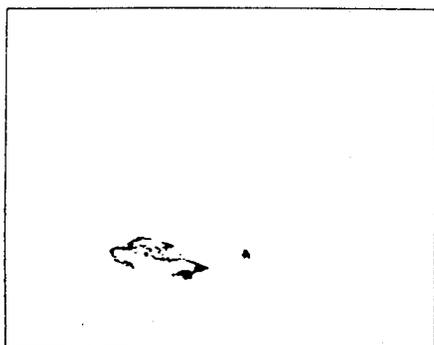


Figure III.14.a

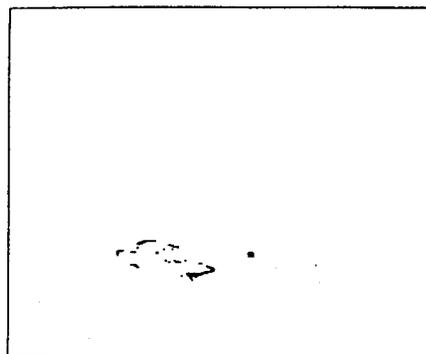


Figure III.14.b

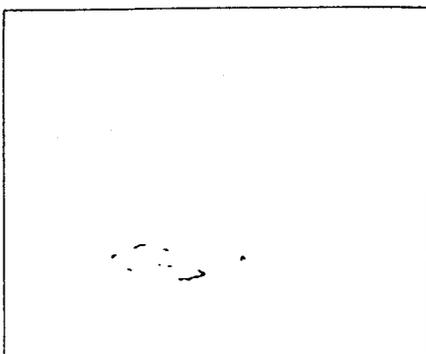


Figure III.14.c

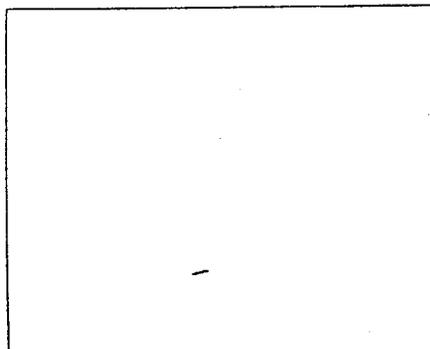


Figure III.14.d

La robustesse de ces méthodes, qui ne nécessitent aucune connaissance a priori sur la scène analysée et tolèrent de mauvaises conditions d'éclairage, a été testé sur des images réelles et de laboratoire.

Dans le cas monodimensionnel la qualité de l'image des objets en mouvement est fonction du choix du paramètre "seuil" utilisé lors de la procédure de segmentation et d'extraction des points de contour. Une légère variation autour de la valeur nominale, choisie dans l'exemple, fait apparaître des modifications considérables. Cependant, dans certaines applications qui nécessitent l'emploi d'une caméra linéaire, il peut être intéressant d'utiliser une telle méthode qui détecte le mouvement indépendamment de sa direction par rapport à la ligne. Si la scène analysée est fortement contrastée et les variations d'éclairage faibles, la valeur du paramètre "seuil" peut être fixée manuellement au début de l'opération et rester inchangée durant l'analyse de la séquence d'image.

La méthode bidimensionnelle, basée sur la multiplication des gradients s'est rapidement avérée plus intéressante et moins sensible aux bruits. Des différents types d'opérateurs utilisés, celui de SOBEL a donné une représentation des contours plus prononcée. Ceux de PREWITT et de ROBERTS apportent cependant suffisamment d'informations pour pouvoir être utilisés dans notre application. La validité des résultats obtenus par ces 3 opérateurs ne nous ont pas incité à utiliser d'autres types de masques.

Dans le cas où le fond n'est pas uniforme, la procédure peut isoler des points de contour en dehors de la position de l'objet dans l'image courante. Lorsque l'objet se déplace, il peut faire apparaître un fond texturé ou présentant des non-homogénéités de niveaux de gris. Le gradient coïncide pour certains points avec ceux de l'objet dans l'image précédente et sont par conséquent isolés dans l'étape de multiplication. Afin d'éviter ces fausses détections, on pourrait envisager de sélectionner les points qui répondent non seulement aux conditions sur les modules du gradient, mais aussi sur leur orientation. Ainsi, pour les points vérifiant

l'équation  $G(C) \times G(D) > S$ , on ne sélectionnerait que ceux pour lesquels la direction de  $G(C)$  serait identique à la direction de  $G(D)$ . Les pénalités que cette modification apporterait au niveau des temps de calcul, nous ont conduit à rejeter cette approche.

Le seuillage de l'image  $\{G(C) \times G(D)\}$  a montré que l'ajustement du seuil n'était pas critique. Un processus de binarisation automatique peut être aisément élaboré. Des expériences, sur différents types de scène, ont montré que le calcul de la valeur maximale du produit  $\{G(C) \times G(D)\}$ , mesuré sur l'ensemble des points des images, suffisait pour déterminer un seuil correct. Ce dernier ayant pour valeur  $\max \{G(C) \times G(D)\}/2$  valide, de manière robuste, le résultat du processus de détection des contours des objets en mouvement.

La modélisation des objets en mouvement fera l'objet du chapitre suivant. Nous verrons comment des modèles rectangulaires permettent d'analyser la scène en vue de piloter le tramway.

## **C H A P I T R E I V**

**MODELISATION DES OBJETS EN MOUVEMENT  
ET PROCESSUS DE CONDUITE AUTOMATIQUE.**

## IV.1. MODELISATION DES OBJETS EN MOUVEMENT

### IV.1.1. Détermination du rectangle exinscrit aux objets en mouvement

Les méthodes développées dans le chapitre précédent, font apparaître les points de contour des objets en mouvement. Bien qu'elle soit très précise, cette description des objets en mouvement est peu utilisable. Une étape importante de l'analyse d'une scène dynamique consiste à comprimer l'information sous une forme plus compacte et donc plus facilement exploitable.

Il existe de nombreuses méthodes qui permettent de modéliser les contours des objets et qui conduisent à des modèles suffisamment simples pour être manipulés aisément et suffisamment élaborés pour minimiser la différence entre le contour et son modèle.

Mais les grandes variations de contraste, d'éclairage et de réflectance des objets au cours des prises de vues ne permettent pas de détecter des contours complètement fermés. Dans les conditions de prises de vues réelles, le contour extérieur et les contours intérieurs aux objets en mouvement présentent des discontinuités parfois importantes. Il serait donc inutile de vouloir modéliser des contours souvent dégradés par des modèles aussi précis que ceux retenus par les méthodes d'approximation de contours (MIC 78, STU 80) ou par l'utilisation de la transformée de HOUGH (DUD 72, MAI 84).

Plutôt que d'essayer de modéliser la forme des contours en mouvement, il est préférable de tenter d'obtenir une représentation plus globale de chacun de ces objets en mouvement. Le principal intérêt d'une représentation globale réside dans la réduction d'informations, qui permet d'envisager une application temps réel. Cette remarque nous conduit à utiliser un modèle rectangulaire, dont les côtés sont parallèles et perpendiculaires au balayage vidéo, pour lequel la compression de données est énorme. En effet, l'image brute, composée de 256 sur 256 pixels,

est digitalisée sur 256 niveaux de gris, soit 8 bits. La représentation d'un objet par un rectangle requiert uniquement 4 paramètres, chaque paramètre étant codé sur 8 bits (0 à 255). Dans le cas de k objets en mouvement dans la scène, on observe un coefficient de réduction de données de :

$$\frac{256 \times 256 \times 8}{k \times 8 \times 4} \approx \frac{16}{k} 10^3$$

Dans notre application, on relève en moyenne un nombre d'objets de l'ordre de 10. La compression ainsi obtenue ( $16 \cdot 10^2$ ) souligne l'intérêt d'une telle approche dans un processus temps réel.

La recherche des rectangles encadrant les objets en mouvement ne peut s'effectuer directement sur l'image résultant du seuillage de la multiplication des images G(C) et G(D) du fait des nombreuses discontinuités présentes dans les contours. En effet, l'affectation des points de contour à un même objet en mouvement ne peut se faire que sur des propriétés de connexité (ROS 70). Celles-ci sont renforcées par un prétraitement du produit des images G(C) et G(D) utilisant un filtrage logique élémentaire. On détermine le nombre de pixels égaux à un dans une fenêtre centrée en chaque point de l'image binaire  $\{G(C) \times G(D)\}$ . On affecte alors, au centre de la fenêtre, la valeur unité si le nombre de ses pixels voisins égaux à 1 est supérieur à un seuil fixe. Le regroupement des points, obtenus à la suite de ce prétraitement, est effectué sur la base d'une analyse de leur connexité. Cette opération conduit à une représentation des objets en mouvement non plus par des segments de contour discontinus, mais des ensembles de points connexes ou amas, dont l'enveloppe extérieure délimite les contours de l'objet. Le regroupement des points constituant chaque amas et la détermination du rectangle exinscrit peuvent être conduits simultanément par la procédure développée par STUCK (STUC 80).

Ce traitement monodimensionnel analyse chaque ligne et détermine par simple détection des transitions opposées (noir sur blanc ou inversement), des segments appartenant aux objets. Au fur et à mesure

du balayage, les segments sont réunis s'ils vérifient un test d'adjacence. Ainsi, pour que le segment  $(m, n)$  de la ligne  $i$  soit réuni au segment  $(m', n')$  de la ligne  $i+1$ , il suffit qu'au moins l'une des 4 conditions ci-dessous soit vérifiée :

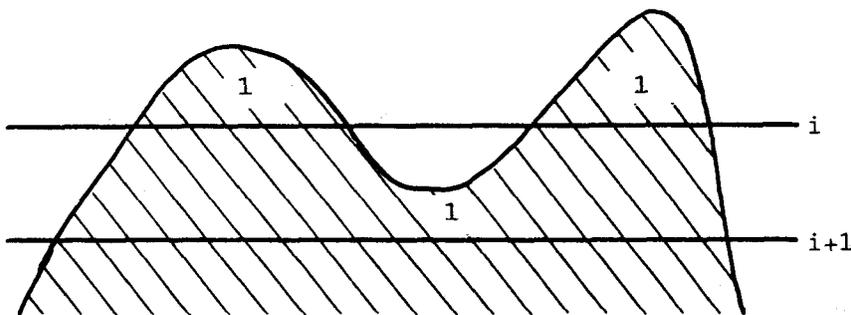
$$m \in [m', n']$$

$$n \in [m', n']$$

$$m' \in [m, n]$$

$$n' \in [m, n]$$

La figure IV.1. illustre un regroupement de deux segments. Les segments 1 et 2 de la ligne  $i$  sont rassemblés à la ligne  $i+1$ .



- Figure IV.1. -

A partir de ces considérations simples, on détermine les minima et maxima des abscisses et ordonnées de chaque amas représentant un objet en mouvement. Cette recherche s'effectue ligne après ligne et ne nécessite, par conséquent, qu'un faible encombrement mémoire.

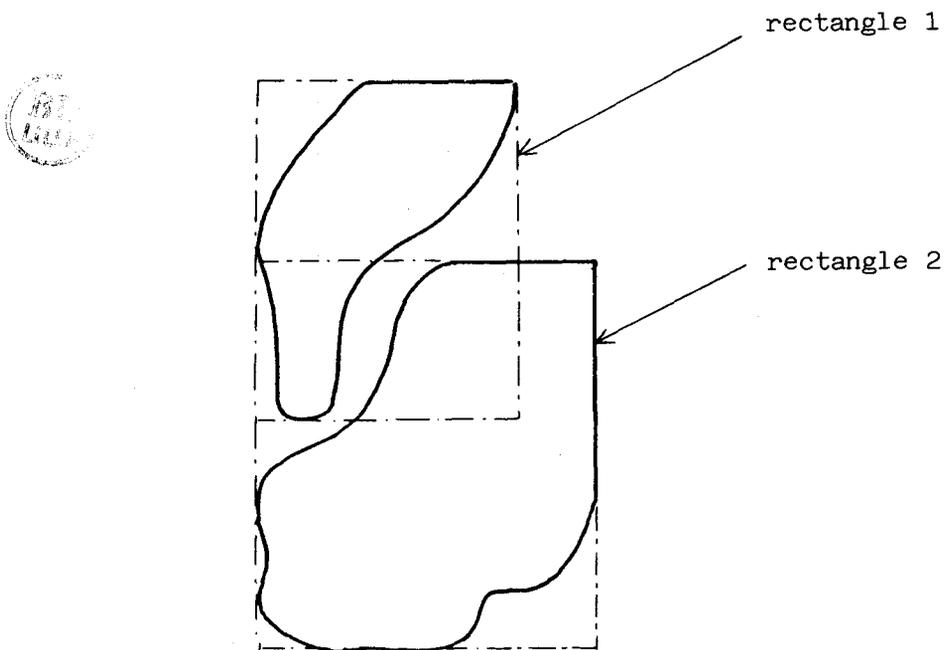
#### IV.1.2. Modèles rectangulaires des objets en mouvement

##### IV.1.2.1. Reconstitution d'un objet en mouvement

La procédure de lissage logique qui permet de regrouper des points en mouvement appartenant au même objet peut laisser subsister des discontinuités importantes de telle sorte qu'un objet en mouvement peut-être mis en évidence sous la forme d'amas séparés, bien que voisins.

En effet, le manque de contraste entre certaines parties de l'objet et le fond, ainsi que les variations d'éclairage et de réflectance entre deux acquisitions successives altèrent la représentation de l'amas correspondant à l'objet en mouvement. De plus, certains objets sont constitués de surfaces de niveaux de gris homogène. Comme il n'existe pas de points de transition représentatifs à l'intérieur de ces surfaces, il est difficile d'isoler les points de contour internes et, par conséquent, des amas significatifs.

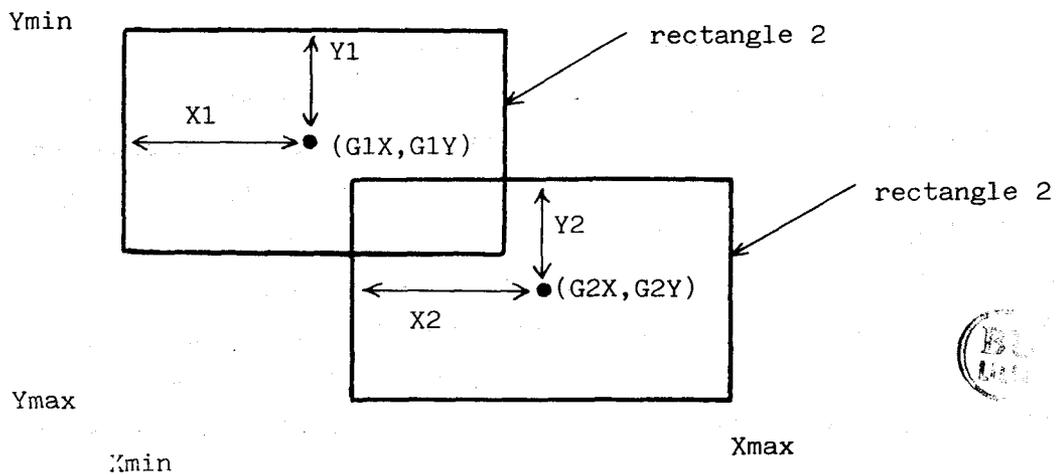
Dans ce cas, les rectangles encadrant les différentes régions relatives à un même objet sont souvent imbriqués comme l'indique la figure IV.2.



- Figure IV.2. -

- IV.5. -

Cette propriété nous amène à fusionner les rectangles dont la réunion comporte un ensemble non vide. Cette réunion peut-être vérifiée en décrivant le rectangle par ses côtés (notés  $2X$  et  $2Y$ ) et les coordonnées de son centre (notées  $G_X$  et  $G_Y$ ) comme le montre la figure IV.3.



- Figure IV.3. -

En posant :

$$\Delta X = |G2X - G1X| - |X2 + X1|$$

$$\Delta Y = |G2Y - G1Y| - |Y2 + Y1|$$

La réunion des rectangles 1 et 2 vérifie la relation suivante :

$$\Delta X < 0 \quad \text{ET} \quad \Delta Y < 0$$

Le rectangle unique ainsi obtenu aura pour sommets  $X_{min}$ ,  $X_{max}$ ,  $Y_{min}$  et  $Y_{max}$  (cf. figure IV.4.) donnés par les relations suivantes :

$$X_{\min} = \min (G1X - X1 , G2X - X2)$$

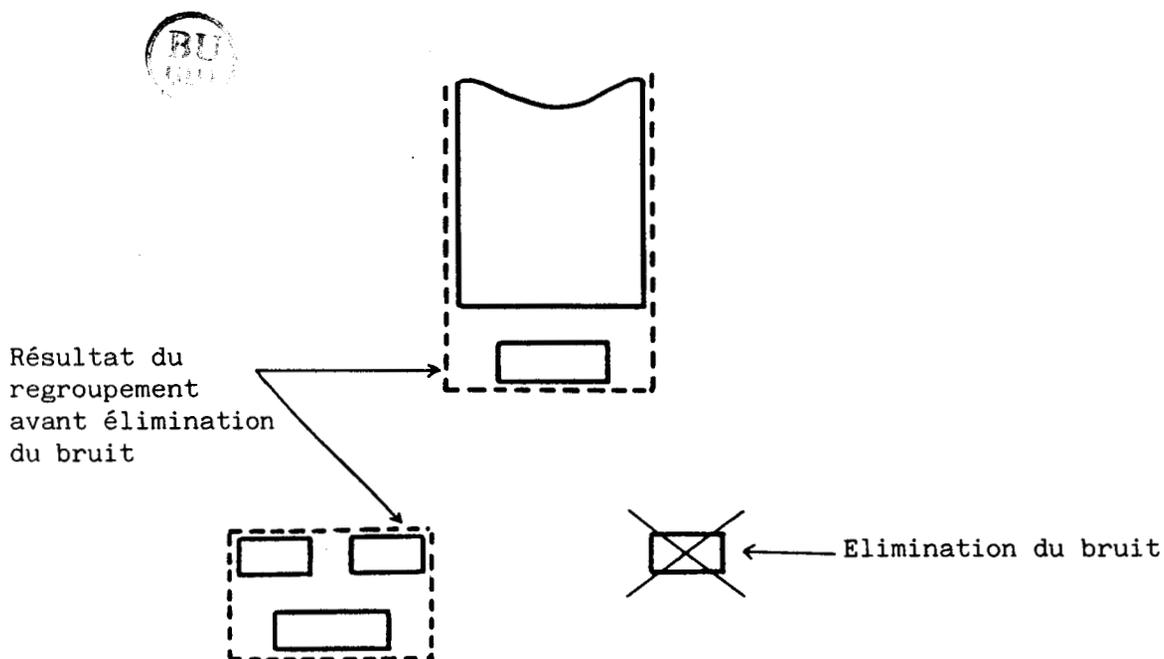
$$X_{\max} = \max (G1X + X1 , G2X - X2)$$

$$Y_{\min} = \min (G1Y - Y1 , G2Y - Y2)$$

$$Y_{\max} = \max (G1Y + Y1 , G2Y + Y2)$$

#### IV.1.2.2. Elimination du bruit

En définissant un critère de surface en dessous duquel tout rectangle doit être soit supprimé, soit regroupé à ses voisins, on peut éliminer les effets du bruit. Supprimer directement tout rectangle de surface inférieure à un seuil pourrait amener à éliminer des objets en mouvement détectés sous la forme de petits rectangles voisins. Mais en commençant par fusionner tous les petits rectangles avec leurs voisins avant de procéder à l'élimination de rectangles isolés, on parvient à préserver les objets en mouvement et à supprimer le bruit (cf. figure IV.4.).



- Figure IV.4. -

#### IV.1.3. Performances des procédures développées

Nous allons illustrer, sur trois exemples utilisant la scène réelle des figures III.3. et III.4., la performance des procédures présentées aux paragraphes précédents.

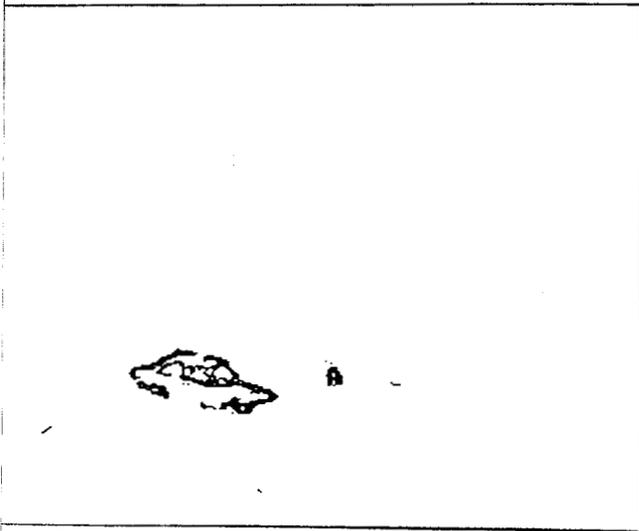
Les figures IV.5.a. et IV.5.b. représentent respectivement l'image des contours des objets en mouvement et les amas constitués. La recherche des rectangles exinscrits fait apparaître 4 rectangles : le premier modélisant le véhicule, le second le piéton, les deux derniers correspondant à du bruit. Ces deux petits amas étant isolés ils peuvent, par conséquent, être éliminés (figure IV.5.d.). La justesse des modèles ainsi créés peut être appréciée en les dessinant sur l'image courante (figure IV.6.). Cette figure montre la précision avec laquelle le véhicule est représenté. Quant au modèle correspondant au piéton il n'encadre que la partie supérieure du fait d'un manque de contraste entre le fond et une partie du piéton qui n'autorise pas la détection de points de contour significatifs.

La perte d'informations résultant d'un seuil de binarisation trop élevé modifie l'image des contours en mouvement (figure IV.7.a.) et fait apparaître plusieurs amas distincts (figure IV.7.b.). Ces différents amas sont mis en évidence sur la figure IV.7.c. qui montre que le véhicule n'est plus composé d'un seul et unique amas mais de quatre amas pour lesquels les rectangles exinscrits sont imbriqués. En fusionnant ces différents amas et en supprimant ceux qui se trouvent isolés, on vérifie que les modèles ainsi déterminés encadrent parfaitement les objets en mouvement. On remarque que la figure IV.7.d. est identique à la figure IV.5.d.

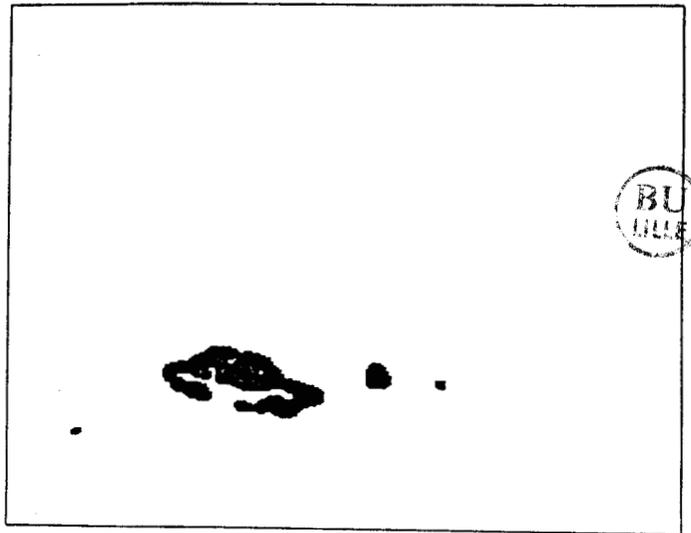
L'image IV.8.a. reprend l'image IV.7.a., sur laquelle nous avons généré artificiellement du bruit. Après avoir recherché les amas ainsi que les rectangles exinscrits (figure IV.8.b.), nous avons réuni certains amas (figure IV.8.c.) et éliminé le bruit (figure IV.8.d.).

La taille du rectangle encadrant le véhicule est légèrement supérieure à celle trouvée sur les figures IV.7.d. et IV.5.d., du fait de l'agrégation de 3 petits rectangles. Mais cette fusion n'altère pas dans de grandes proportions la modélisation finale de cet objet.

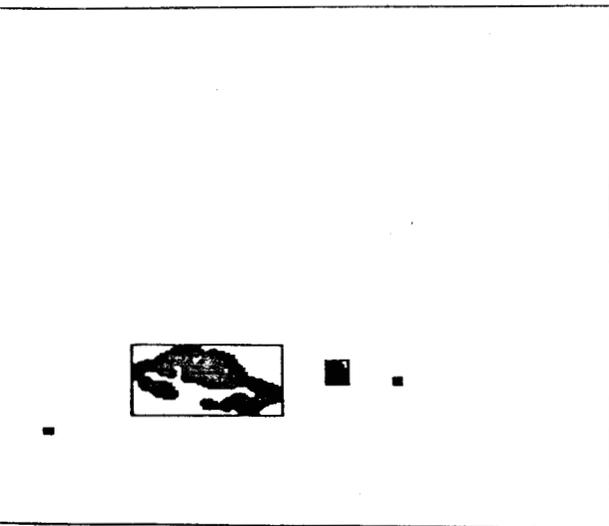
Les résultats obtenus sur les exemples présentés dans ce paragraphe s'avèrent satisfaisants même dans le cas où les contours de l'objet en mouvement constituent des amas disjoints et soulignent la robustesse des procédures par rapport au bruit.



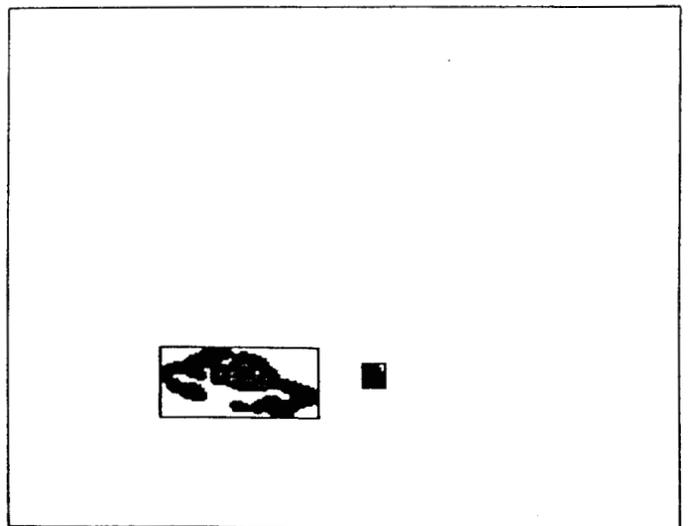
- Figure IV.5.a. -



- Figure IV.5.b. -



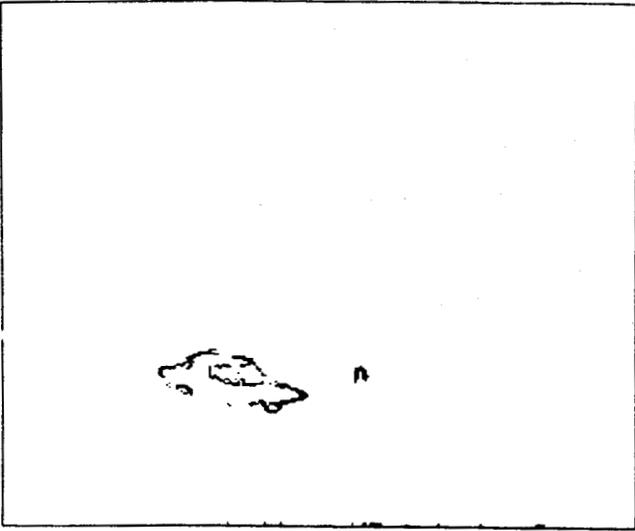
- Figure IV.5.c. -



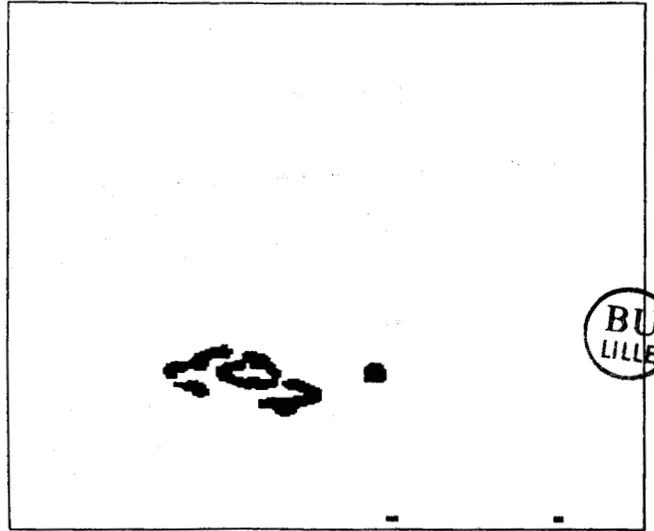
- Figure IV.5.d. -



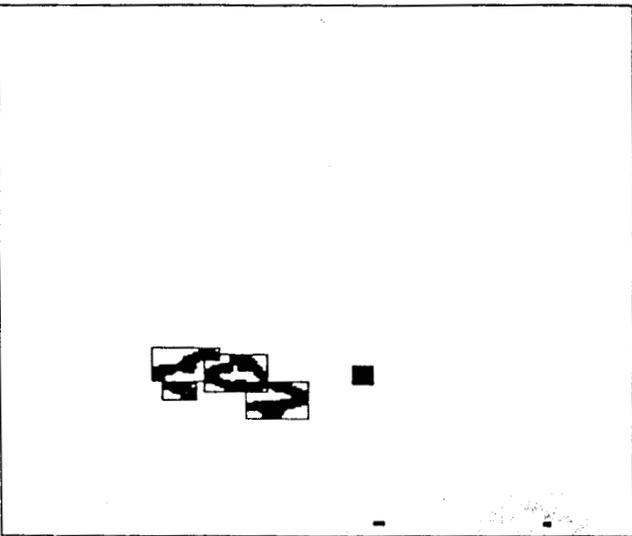
- Figure IV.6. -



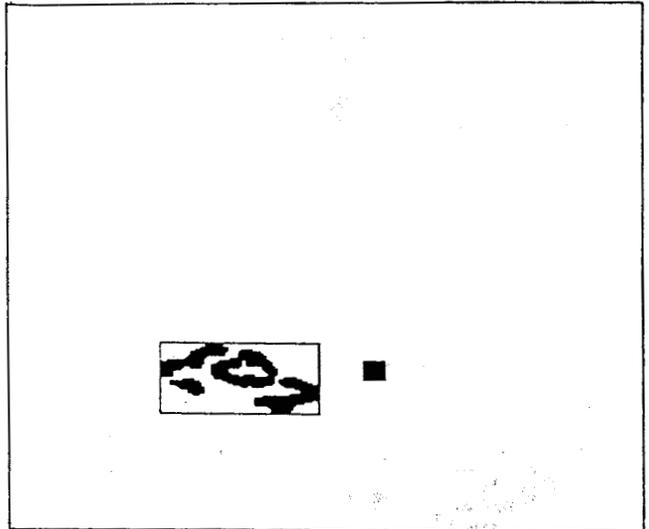
- Figure IV.7.a. -



- Figure IV.7.b. -

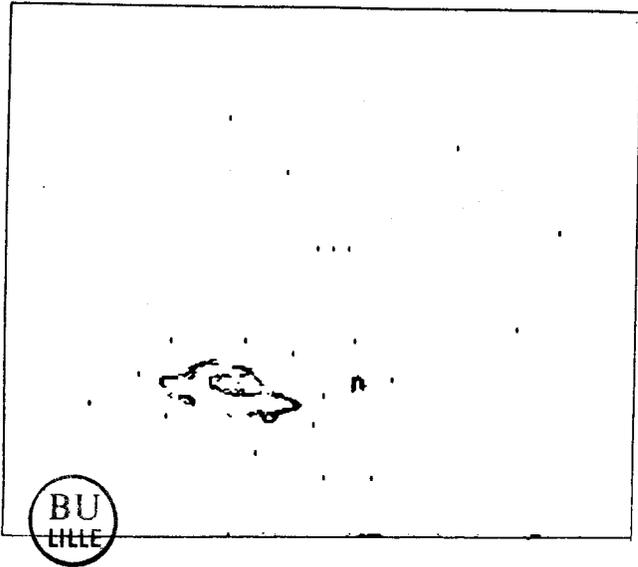


- Figure IV.7.c. -

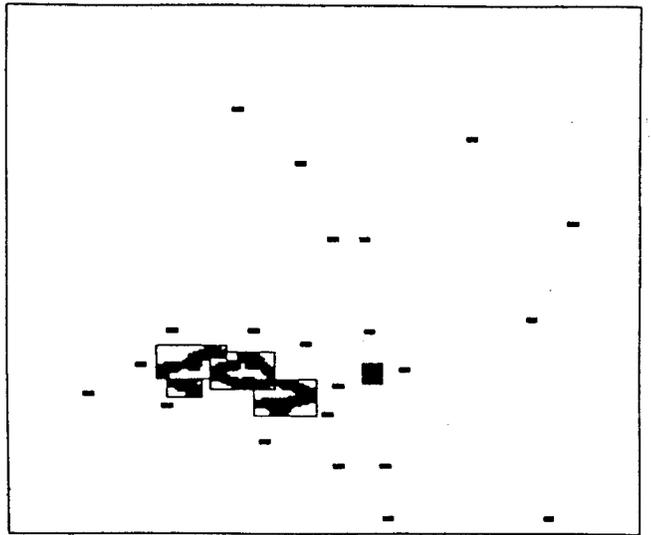


- Figure IV.7.d. -

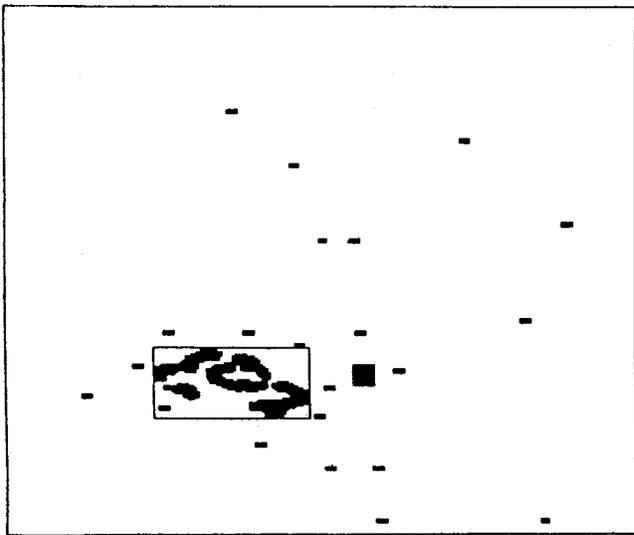
- Figure IV.7. -



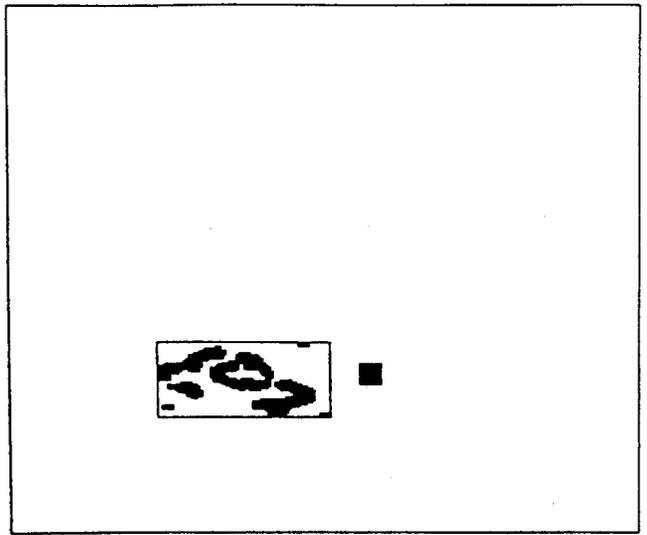
- Figure IV.8.a. -



- Figure IV.8.b. -



- Figure IV.8.c. -



- Figure IV.8.d. -

## IV.2. PROCESSUS DE CONDUITE AUTOMATIQUE

### IV.2.1. Simulation et initialisation

Les résultats obtenus sur une scène réelle montrent comment les points de contour des objets en mouvement peuvent être exploités en vue de modéliser ces objets par des rectangles. Il serait tentant, au vu de ces travaux, d'envisager d'étudier la conduite automatique du tramway sur des scènes réelles. Mais une telle approche ne présenterait pas la souplesse d'utilisation nécessaire pour notre étude.

Aussi, comme les problèmes de traitement d'images sont maîtrisés, il s'est avéré plus intéressant de développer la stratégie de conduite non pas sur site réel, mais à partir d'une maquette de simulation. Ce choix nous permet de créer toutes les situations désirées par simple déplacement des objets constituant la scène.

Cette maquette de simulation occupe une surface de 3m sur 1,5m sur laquelle se déplacent des véhicules tels que voitures, camions, motos et des personnages. Le tramway, motorisé, circule sur des rails. Des arbres, ainsi que des caténaires font fonction d'objets fixes. Le carrefour ainsi simulé a pour échelle 1/22,5. Il faut souligner qu'aucun éclairage artificiel n'est utilisé, la maquette étant placée dans une salle recevant la lumière naturelle.

Etant maître de la position de la caméra, nous l'avons placé perpendiculairement au-dessus de la scène, de telle sorte que les rails du tramway soient sensiblement parallèles au balayage vidéo de la caméra.

La conduite automatique du tramway dans un univers inconnu ne peut s'effectuer sans un minimum de connaissances. Aucune hypothèse ne peut être faite sur la forme, la taille, les déplacements et les niveaux de gris des objets constituant la scène. Par contre, les paramètres relatifs au tramway, à savoir sa longueur, sa largeur ainsi que sa trajectoire sont connus a priori. Dans une phase d'initialisation on déplace, manuellement, un curseur sur l'image représentant la scène afin de repérer les extrémités droite et gauche du milieu de la voie en vue de déterminer son ordonnée moyenne. Le tramway étant présent sur les images utilisées pour cette phase

d'initialisation, on note ses dimensions en plaçant le curseur sur ses extrémités. Nous disposons ainsi des dimensions d'un masque rectangulaire représentant le tramway. La procédure d'initialisation doit également donner le sens de déplacement (gauche-droite ou inversement) du tramway sur la voie. Cette information est donnée par l'intermédiaire du clavier, car les tramways se déplacent toujours dans le même sens sur chaque voie. La connaissance du sens de déplacement permet de détecter l'entrée du tramway dans le champ de vision de la caméra. C'est à partir de ce point que sera initialisée la procédure de poursuite.

#### IV.2.2. Différentes étapes du processus de conduite automatique

Le processus de conduite automatique se décompose, selon un schéma bien structuré, en deux parties distinctes à savoir :

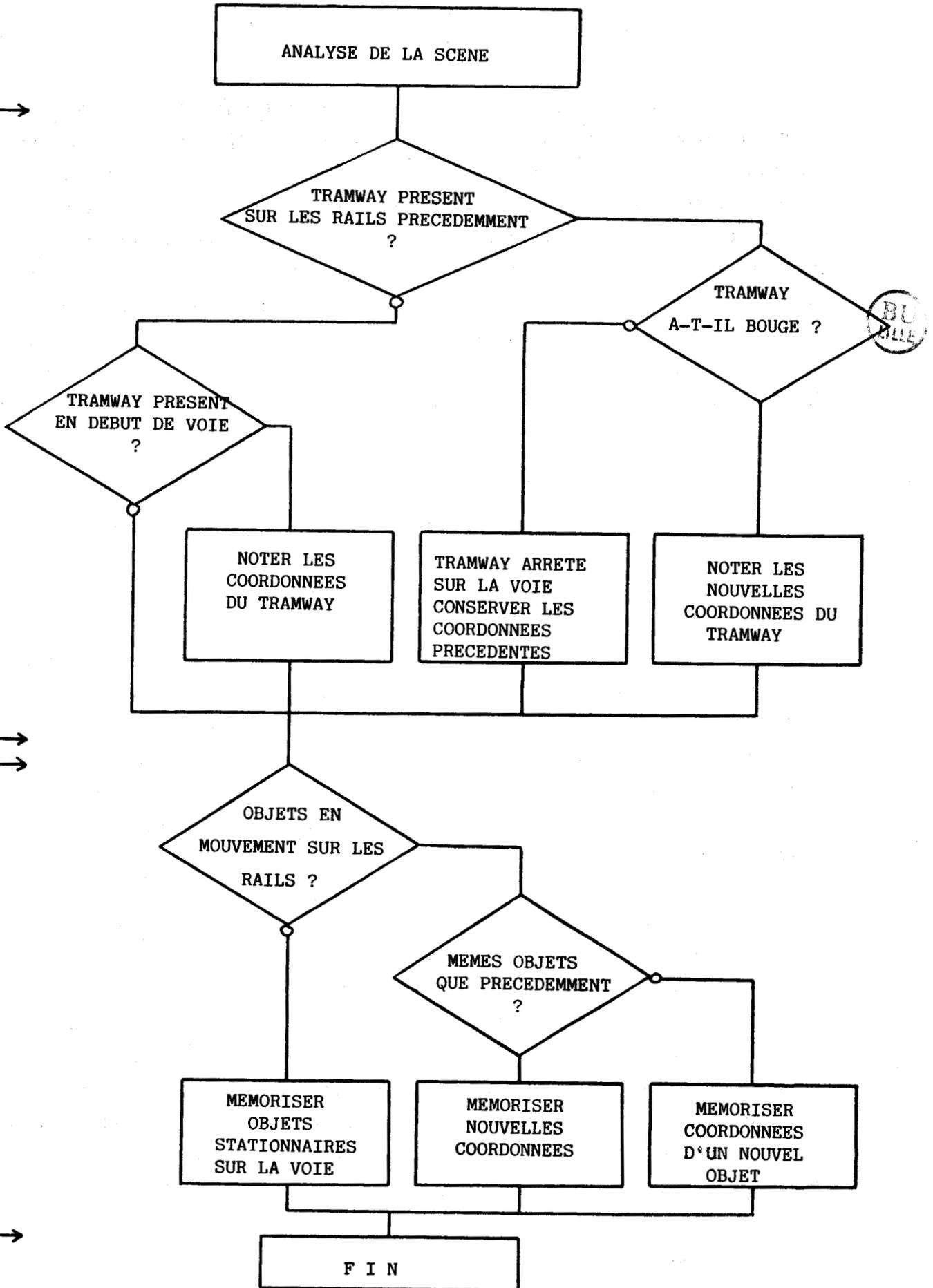
- 1 - Analyse de scène
- 2 - Stratégie de conduite

La première partie consiste à déterminer les modèles rectangulaires qu'on associe au tramway et aux objets qui, de par leurs mouvements peuvent présenter un risque de collision avec le tramway.

La seconde partie traite de la stratégie de conduite en vue de piloter le tramway en évitant les collisions.

##### IV.2.2.1. Analyse de scène

L'organigramme présenté sur la figure IV.9. illustre, de façon concise, les différentes étapes de la procédure d'analyse de scène. La première partie de cet organigramme concerne la poursuite du tramway. La seconde, quant à elle, vise à rechercher les objets se trouvant dans une "zone critique" correspondant à la voie du tramway et son voisinage immédiat. Les objets situés à l'extérieur de cette zone critique seront étudiés ultérieurement.



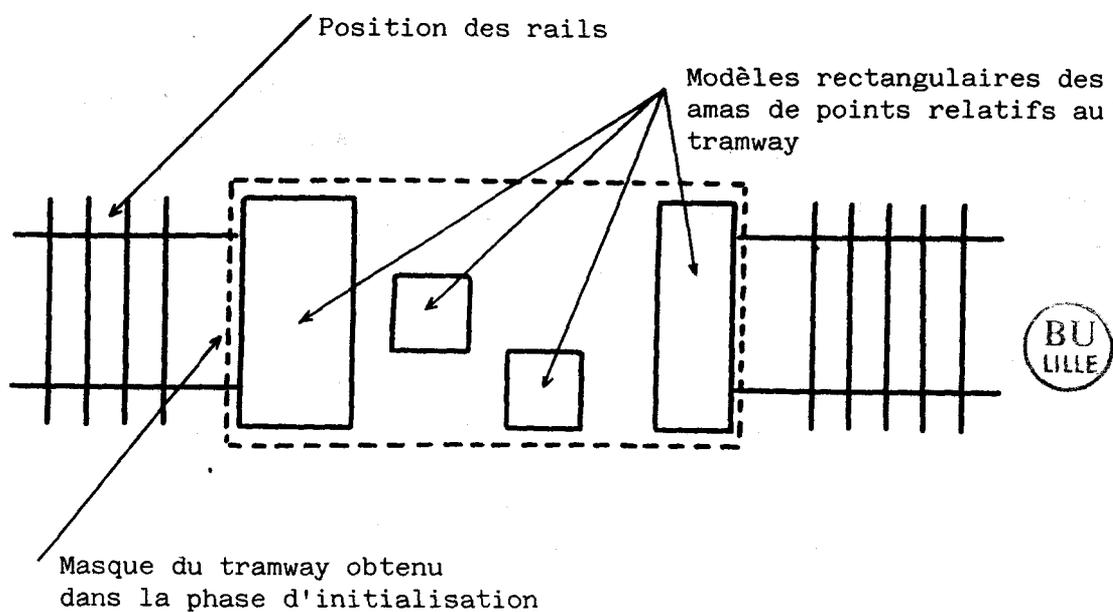
La connaissance des objets présent à l'intérieur de cette "zone critique" est très importante car, du fait de la proximité du tramway, ils peuvent présenter des risques de collision. Une attention particulière devra être apportée à la poursuite des objets dans cette zone, en prenant toutes les précautions nécessaires pour conserver la trace des objets après une éventuelle immobilisation. En effet, la technique utilisée, basée sur les différences d'images successives, isole les seuls objets en mouvement. Ainsi, si un objet s'arrête dans cette zone, il devient invisible pour la procédure de détection et n'est pas pris en compte dans la détermination du risque de collision. La collision avec le tramway devient donc inéluctable.

#### IV.2.2.1.a. Poursuite du tramway

A la première acquisition, le tramway est supposé hors du champ de vision de la caméra. On attend son arrivée en début de voie et on initialise la procédure de poursuite du tramway décrite ci-dessous.

On recherche pour chaque nouvelle acquisition les coordonnées du tramway qui ont ainsi réactualisées. Cette recherche ne s'effectue pas sur l'image entière mais à l'intérieur d'une fenêtre délimitée à partir de la position du tramway dans l'image précédente et du déplacement maximal que celui-ci peut effectuer entre deux acquisitions (la période d'échantillonnage et la vitesse maximale du tramway permettent de calculer ce déplacement).

La connaissance, a priori, des dimensions du masque rectangulaire représentant le tramway, permet de simplifier la détection de ce dernier dans la fenêtre. En effet, il est très simple de détecter la position de l'avant du tramway dans la fenêtre et, connaissant le sens de déplacement, de positionner le masque sur la voie. Cette procédure présente l'énorme avantage d'éliminer tous les rectangles associés aux amas de points en mouvement relatifs au tramway. En effet, compte-tenu de la taille du tramway et de son homogénéité en niveaux de gris, il est peu probable qu'il soit détecté sous la forme d'un amas unique (figure IV.10.).



- Figure IV.10. -

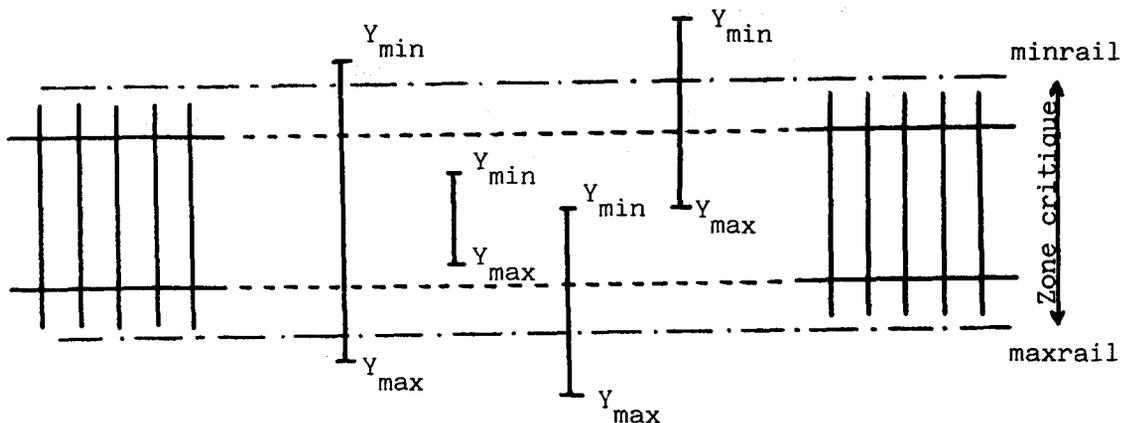
La procédure de poursuite doit permettre de conserver la position du tramway, même lorsque celui-ci est à l'arrêt dans le champ de la caméra. Nous avons déjà indiqué que la procédure utilisée permet de ne détecter que les contours des objets en mouvement. Il est donc nécessaire de mémoriser les coordonnées du masque représentant le tramway à chaque acquisition afin de conserver sa position en cas d'arrêt.

IV.2.2.1.b. Recherche des objets sur la voie

La procédure de conduite automatique doit prendre en compte les objets en mouvement dans la scène et ceux qui s'y seraient arrêtés. S'il est nécessaire de surveiller les déplacement de tous les objets en mouvement dans la scène analysée, il est par contre inutile de mémoriser leur position au cas où ils s'arrêteraient loin des rails.

Pour cette raison nous avons défini une "zone critique", constituée par les rails et leur voisinage immédiat, dans laquelle il est impératif de mémoriser la position des objets arrêtés. En effet, pour les objets à l'arrêt dans la scène, seuls ceux situés à l'intérieur de cette zone critique présenteront des risques de collision avec le tramway.

Dans le cas où les rails sont parallèles au balayage horizontal, cette zone critique peut être définie par la bande horizontale comprise entre les ordonnées minimales et maximales. Plus la bande sera large et plus grande sera la marge de sécurité. Un objet, présent à l'intérieur de la zone critique et dont le modèle rectangulaire a pour abscisses XMIN et XMAX et pour ordonnées YMIN et YMAX, peut s'inscrire dans l'une des 4 configurations représentées sur la figure IV.11.



- Figure IV.11. -

Ainsi, pour tous les objets de la scène situés à l'avant du tramway, on affecte un indice de présence sur voie à l'objet considéré si l'une au moins des conditions ci-dessous est vérifiée :

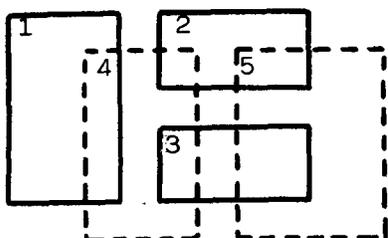
- . minrail  $\in$  [Ymin, Ymax]
- . maxrail  $\in$  [Ymin, Ymax]
- . Ymin  $\in$  [minrail, maxrail]



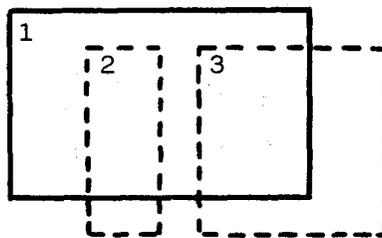
Cet indice de présence sur voie, qu'on notera I.P.V., impose la mémorisation des coordonnées de l'objet. Ces coordonnées sont réactualisées dans le cas où l'objet continue de se déplacer dans la zone critique.

Lors de la poursuite de ces objets, il peut arriver qu'entre deux images successives, leurs modèles rectangulaires s'éclatent sous la forme d'un nombre supérieur de rectangles, ou inversement, que plusieurs rectangles représentant le même objet fusionnent en un nombre inférieur de modèles rectangulaires dans l'image suivante. Cette situation est la conséquence directe d'un manque de contraste entre certaines parties de l'objet et le fond, des variations d'éclairage et de réflectance, ou du passage d'un objet entre deux zones différemment éclairées. Toutefois, pour une période d'échantillonnage temporel suffisamment courte, il y aura chevauchement des modèles dans l'image courante et dans l'image précédente de telle sorte qu'il sera possible d'établir une correspondance entre les modèles rectangulaires représentant le même objet dans deux images successives.

Cette situation est illustrée sur les figures IV.12.a. et IV.12.b. sur lesquelles les rectangles en trait continu décrivent l'objet dans l'image précédente, les rectangles en pointillés les représentent dans l'image courante.



- Figure IV.12.a. -



- Figure IV.12.b. -

Sur le premier exemple (figure IV.12.a.), on associe l'I.P.V. des rectangles 1, 2 et 3 aux rectangles 4 et 5, à condition toutefois que ceux-ci soient aussi à l'intérieur de la zone critique. Sur le second exemple (figure IV.12.b.), l'I.P.V. du rectangle 1 est transféré aux rectangles 2 et 3 uniquement si ceux-ci sont présents dans la zone critique.

Dans le cas où l'objet s'est suffisamment déplacé et n'est plus détecté dans la zone critique, on annule son, ou ses, I.P.V. L'objet, s'il s'arrête, ne présente plus de risque de collision avec le tramway.

#### IV.2.2.2. Stratégie de conduite

##### IV.2.2.2.a. Risque de collision mesuré

Les procédures développées précédemment nous permettent d'avoir une description suffisante de la scène. Les coordonnées du tramway et des objets ainsi que les informations relatives à leur état (en mouvement, arrêté....) sont mémorisées. La traversée du carrefour apparaît comme la

recherche d'objets pouvant présenter un risque de collision pour le tramway. Il faut souligner qu'il ne s'agit pas de poursuivre les voitures, piétons, etc...., afin de déterminer leur trajectoire et vitesse pour piloter le tramway. Le problème consiste à éviter les collisions dans le cas où un objet se rapproche trop du tramway. En effet, la traversée des carrefours est assujettie à un système de feux de circulation qui donne toujours la priorité au tramway. Sa présence dans la scène laisse prétendre l'absence, en théorie, d'autres objets sur son passage. Le système que nous développons a pour but d'éviter la collision avec un objet qui n'a pas répondu aux indications dictées par la réglementation routière.

Nous avons regroupé dans les tableaux IV.1. et IV.2., toutes les situations correspondant à un risque de collision pour un tramway se déplaçant de la droite vers la gauche. On y représente les conditions sur les coordonnées des modèles de l'objet et du tramway, l'intervalle les séparant ainsi que les coordonnées correspondant au tracé de cet intervalle. Ainsi, les distances séparant le tramway des différents objets sont visualisés sur l'écran et permettent d'apprécier la validité des ordres de commande définis ultérieurement. Les coordonnées relatives aux modèles rectangulaires des objets sont données par XMIN et XMAX pour les abscisses minimales et maximales et YMIN et YMAX pour les ordonnées. Celles relatives au tramway sont représentées par XTMIN, XTMAX, YTMIN et YTMAX.

Nous n'avons pas traité le cas où les objets sont situés derrière le tramway. Une collision dans ces conditions ne peut être évitée en freinant ou arrêtant le tramway. Une remarque identique peut être faite dans les situations relevées au tableau IV.2. En effet, aucune mesure visant à éviter une collision ne peut être prise dans ces conditions. Toutefois, afin de garantir le maximum de sécurité, nous tenons compte des situations présentées au tableau IV.2. en stoppant le tramway. Cette fonction est inscrite dans la procédure développée au paragraphe IV.2.2.1.a. Nous avons vu que le positionnement du masque avec la partie antérieure du tramway, permet d'éliminer tous les rectangles compris à l'intérieur du masque. L'élimination de ces rectangles ne peut s'effectuer

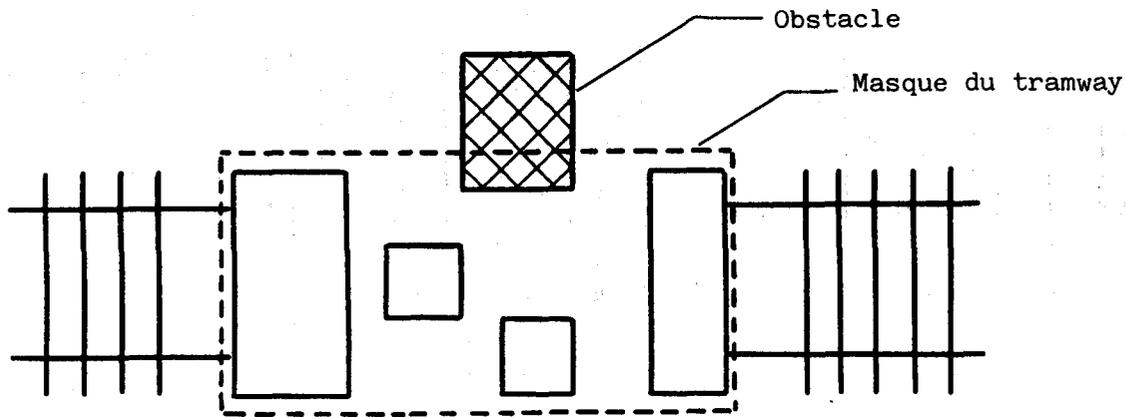
CAS	CONDITIONS	DISTANCES A CALCULER	COORDONNEES DES POINTS POUR LE TRACE DE LA DROITE	FIGURES
<b>1</b>	$X_{MAX} < X_{TMIN}$ $Y_{MAX} < Y_{TMIN}$	$\left( (X_{MAX} - X_{TMIN})^2 + (Y_{MAX} - Y_{TMIN})^2 \right)^{1/2}$	$Y_{MAX}, Y_{MAX}, X_{TMIN}, Y_{TMIN}$	
<b>2</b>	$X_{MAX} < X_{TMIN}$ $Y_{MIN} > Y_{TMAX}$	$\left( (X_{MAX} - X_{TMIN})^2 + (Y_{MIN} - Y_{TMAX})^2 \right)^{1/2}$	$X_{MAX}, Y_{MIN}, X_{TMIN}, Y_{TMAX}$	
<b>3</b>	$X_{MAX} < X_{TMIN}$ $Y_{MIN} \leq Y_{TMIN}$ $Y_{TMIN} < Y_{MAX} < Y_{TMAX}$	$X_{TMIN} - X_{MAX}$	$X_{MAX}, Y_{TMIN}, X_{TMIN}, Y_{TMIN}$	
<b>4</b>	$X_{MAX} < X_{TMIN}$ $Y_{MIN} \leq Y_{TMIN}$ $Y_{MAX} > Y_{TMAX}$	$X_{TMIN} - X_{MAX}$	$X_{MAX}, (Y_{TMAX} - Y_{TMIN})/2,$ $X_{TMIN}, (Y_{TMAX} - Y_{TMIN})/2$	
<b>5</b>	$X_{MAX} < X_{TMIN}$ $Y_{TMIN} < Y_{MIN} \leq Y_{TMAX}$ $Y_{MAX} > Y_{TMAX}$	$X_{TMIN} - X_{MAX}$	$X_{MAX}, Y_{TMAX}, X_{TMIN}, Y_{TMAX}$	
<b>6</b>	$X_{MAX} < X_{TMIN}$ $Y_{TMIN} < Y_{MIN} < Y_{TMAX}$ $Y_{TMIN} < Y_{MAX} < Y_{TMAX}$	$X_{TMIN} - X_{MAX}$	$X_{MAX}, (Y_{TMAX} - Y_{TMIN})/2,$ $X_{TMIN}, (Y_{TMAX} - Y_{TMIN})/2$	

- Tableau IV.1. -

CAS	CONDITIONS	DISTANCES A CALCULER	COORDONNEES DES POINTS POUR LE TRACE DE LA DROITE	FIGURES
<b>7</b>	$X_{MIN} < X_{TMIN}$ $X_{MAX} > X_{TMIN}$ $Y_{MAX} < Y_{TMIN}$	$Y_{TMIN} - X_{MAX}$	$Y_{TMIN}, Y_{MAX}, Y_{TMIN}, Y_{TMIN}$	
<b>8</b>	$X_{TMIN} < X_{MIN} < X_{TMAX}$ $X_{TMIN} < X_{MAX} < X_{TMAX}$ $Y_{MAX} < Y_{TMIN}$	$Y_{TMIN} - Y_{MAX}$	$(X_{MAX} - X_{MIN})/2, Y_{MAX},$ $(X_{MAX} - X_{MIN})/2, Y_{TMIN}$	
<b>9</b>	$X_{TMIN} < X_{MIN} < X_{TMAX}$ $X_{MAX} > X_{TMAX}$ $Y_{MAX} < Y_{TMIN}$	$Y_{TMIN} - Y_{MAX}$	$X_{TMAX}, Y_{MAX}, X_{TMAX}, Y_{TMIN}$	
<b>10</b>	$X_{MIN} < X_{TMIN}$ $X_{MAX} > X_{TMIN}$ $Y_{MIN} > Y_{TMAX}$	$Y_{MIN} - Y_{TMAX}$	$X_{TMIN}, Y_{MIN}, X_{TMIN}, Y_{TMAX}$	
<b>11</b>	$X_{TMIN} < X_{MIN} < X_{TMAX}$ $X_{TMIN} < X_{MAX} < X_{TMAX}$ $Y_{MIN} > Y_{TMAX}$	$Y_{MIN} - Y_{TMAX}$	$(X_{MAX} - X_{MIN})/2, Y_{MIN},$ $(X_{MAX} - X_{MIN})/2, Y_{TMAX}$	
<b>12</b>	$X_{TMIN} < X_{MIN} < X_{TMAX}$ $X_{MAX} > X_{TMAX}$ $Y_{MIN} > Y_{TMAX}$	$Y_{MIN} - Y_{TMAX}$	$X_{TMAX}, Y_{MIN}, X_{TMAX}, Y_{TMAX}$	

- Tableau IV.2. -





- Figure IV.13. -

que si la totalité de leur surface est inscrite dans le masque. Le cas contraire met en évidence la présence d'un objet trop proche d'un des côtés du tramway (rectangle hachuré de la figure IV.13.).

Ainsi, toutes les situations relatives à un choc latéral, et présentées sur le tableau IV.2., sont prises en compte par cette méthode.

La mesure de la distance séparant le tramway d'un objet s'effectue selon le schéma suivant. Pour chaque objet situé à l'avant du tramway, on recherche le cas de figure auquel il se rapporte dans le tableau IV.1., puis on mesure la distance minimale définie comme la distance entre les deux points les plus proches du modèle rectangulaire de l'objet et du masque représentant le tramway. Selon la valeur de cette mesure, on affecte un indice de risque de collision. Cet indice s'étageant du niveau 0 (risque nul) au niveau 4 (risque maximal) apporte une information sur l'importance du danger.

IV.2.2.2.b. Risque de collision estimé

L'indice de collision mesuré précédemment peut être directement utilisé afin de déterminer les ordres de commande du tramway, un faible indice correspondant à un ordre de prévention (action de l'avertisseur sonore), un indice important, à un ordre d'arrêt.

Ainsi, lorsque la distance entre le tramway et l'objet devient faible, on envoie un ordre de commande qui est fonction de cet indice de collision. Mais on imagine facilement que cet ordre ne doit pas être le même si l'objet concerné tend à s'éloigner ou, au contraire, à se rapprocher du tramway. Dans le premier cas, on doit diminuer l'indice de risque de collision et, dans le second cas, l'augmenter.

Cette approche, basée sur une analogie avec le comportement d'un chauffeur, nous amène à développer une procédure de prédiction du sens de déplacement. Connaissant la position du tramway dans l'image précédente et dans l'image courante, il est facile d'estimer sa position dans l'image suivante. Un schéma similaire peut être utilisé pour les objets, bien que rendu plus ambigu du fait des variations importantes pouvant intervenir dans la modélisation du même objet entre deux acquisitions successives.

Il est certes possible d'effectuer une prédiction exacte sur la position des objets en considérant leurs positions dans les différentes acquisitions précédentes et en utilisant des modèles prédictifs appropriés (KALMAN par exemple (LEG 82)). Cependant, une telle approche est trop lourde et trop coûteuse en temps de calcul. Aussi, nous avons choisi d'utiliser uniquement les deux dernières acquisitions et de prolonger le déplacement des objets. Afin de limiter les incertitudes dues aux variations dans la forme des modèles, nous avons moyenné le déplacement en tenant compte de la position des sommets de ces rectangles dans les acquisitions précédentes et courantes.

En posant  $XMAX_k$ ,  $XMIN_k$ ,  $YMAX_k$  et  $YMIN_k$  les coordonnées d'un rectangle à l'instant  $k$  et  $XMAX_{k-1}$ ,  $YMAX_{k-1}$ ,  $YMIN_{k-1}$ ,  $XMIN_{k-1}$  celles du rectangle représentant le même objet à l'instant  $k-1$ , on détermine la position estimée de cet objet dans l'image suivante par :

$$\hat{X}MAX = XMAX_k + \Delta X$$

$$\hat{X}MIN = XMIN_k + \Delta X$$

$$\hat{Y}MAX = YMAX_k + \Delta Y$$

$$\hat{Y}MIN = YMIN_k + \Delta Y$$

avec :

$$\Delta X = \frac{(XMAX_k - XMAX_{k-1}) + (XMIN_k - XMIN_{k-1})}{2}$$

$$\Delta Y = \frac{(YMAX_k - YMAX_{k-1}) + (YMIN_k - YMIN_{k-1})}{2}$$

Dans le cas où le modèle de l'objet se décompose (ou inversement) entre deux acquisitions successives, il est évident que cette estimation ne permettra pas de prévoir avec certitude sa position dans l'image suivante. Nous préférons alors ne pas prendre en compte une prédiction qui risquerait de s'avérer erronée.

La position dans l'image suivante étant ainsi prédite, on estime la distance minimale séparant le tramway de l'objet dans la prochaine acquisition et on affecte un risque de collision estimé. Cet indice, combiné à l'indice mesuré permettra de déterminer les ordres de commande du tramway.

Il faut souligner que pour des raisons évidentes d'optimisation des temps de calcul, la recherche de la position prédite n'est pas systématiquement effectuée pour tous les objets de la scène. En effet, elle n'apporterait pas d'informations utiles dans le cas où l'objet est situé derrière ou sur les côtés du tramway. Les objets à l'avant du tramway, mais éloignés, se verraient attribuer un risque de collision nul, qui resterait inchangé quel que soit le résultat de la prédiction. Seuls les objets

suffisamment proches du tramway méritent une attention particulière, la prédiction pouvant modifier considérablement le risque de collision.

IV.2.2.2.c. Ordres de commande

Les différents facteurs influant sur cette mesure de prédiction ne permettent pas d'affirmer avec certitude que la distance estimée des objets en mouvement sera vérifiée dans l'acquisition suivante. Cependant, elle apporte un complément d'information qu'il ne faudrait pas négliger et qui permet d'ajuster l'indice de risque de collision. La table de décision (tableau IV.3.) est utilisée afin de sélectionner l'indice final, noté  $ic$ , relatif à chaque objet traité.

Indice mesuré $im$	Indice estimé $\hat{ie}$	Indice choisi $ic$
0 à 3	$\hat{ie} > im$	$ic = im + 1$
1 à 4	$\hat{ie} < im$	$ic = im - 1$
0 à 4	$\hat{ie} = im$	$ic = im$



- Tableau IV.3. -

Parmi tous les objets analysés dans la scène, celui qui présente le risque de collision le plus grand est celui qui possède l'indice  $ic$  le plus élevé. On recherche, par conséquent, l'indice  $ic$  maximal et on lui affecte un ordre de commande. Les ordres, au nombre de 5, sont présentés sur le tableau IV.4.

Indice ic	Ordre de commande
0	AVANCER
1	Actionner l'avertisseur
2	Freiner doucement
3	Freiner énergiquement
4	Arrêter immédiatement

- Tableau IV.4. -

Il est important de souligner que les procédures de commande permettent de s'affranchir des problèmes liés à l'inertie du tramway. En effet, la détection de l'arrêt du tramway n'est pas la conséquence instantanée de l'ordre N° 4 (arrêter immédiatement), mais le résultat des procédures de détection des contours des objets en mouvement qui active la poursuite du tramway tant que celui-ci n'est pas réellement arrêté.

#### IV.2.2.3. Structure de commande de la conduite automatique

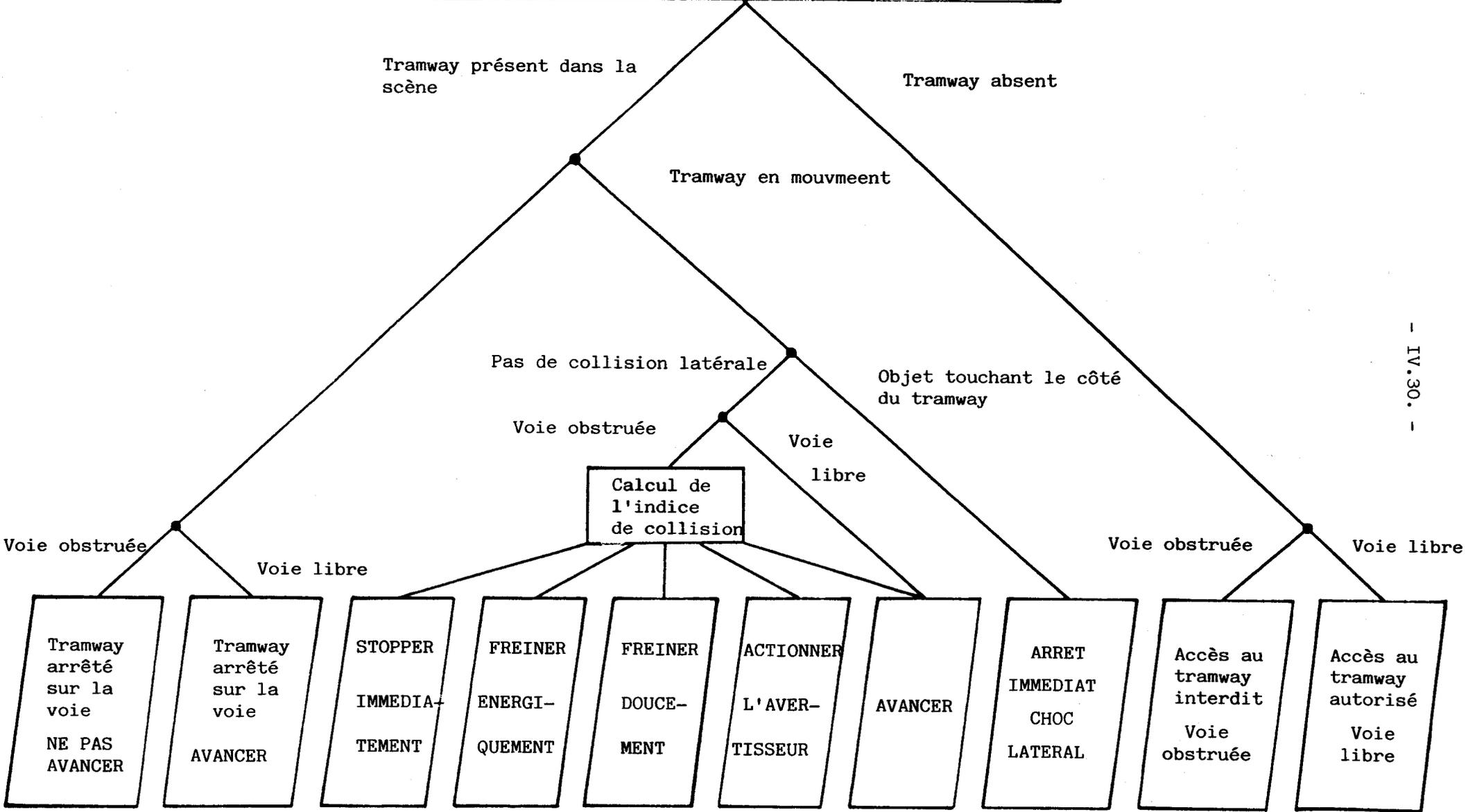
Nous avons présenté dans les paragraphes précédents, les différentes techniques utilisées en vue de donner une représentation complète de la scène analysée.

les différentes situations pouvant intervenir au cours d'une séquence d'images sont résumées sur le tableau IV.5. sous la forme d'un arbre de décision. L'exploration de cet arbre permet de sélectionner les ordres de commande adéquat pour chaque acquisition.

Le principe de la méthode reposant sur la détection des seuls objets en mouvement, il est nécessaire d'analyser la scène "en continu", en vue de mémoriser les coordonnées des objets situés à l'intérieur de la zone critique. Dans le cas où un objet serait présent en début de voie, il apparaît que l'entrée du tramway dans la scène provoquerait une collision immédiate. Cette raison rend nécessaire la surveillance du début de voie en vue d'empêcher, le cas échéant, l'accès du tramway dans la scène. Cette situation est représentée sur la branche droite de l'arbre de décision (tableau IV.5.).

Nous verrons dans le chapitre suivant que ces ordres, au nombre de 10, s'avèrent suffisant pour piloter le tramway au niveau d'un carrefour.

STRUCTURE DE COMMANDE DE LA CONDUITE AUTOMATIQUE



- IV.30. -

- Tableau IV.5. -

## **C H A P I T R E V**

**SIMULATION DE LA TRAVERSEE D'UN CARREFOUR  
EN CONDUITE AUTOMATIQUE.**

### V.1. CONDITIONS EXPERIMENTALES

Ce chapitre illustre le comportement des procédures de conduite automatique de tramway développées précédemment. La maquette de simulation, présentée au paragraphe IV.2.1., est utilisée afin de maîtriser l'évolution des véhicules et du tramway, ce qui permet de créer les situations désirées.

En considérant une période d'échantillonnage de l'ordre du  $1/25^{\circ}$  de seconde, le déplacement observé pour un véhicule se déplaçant à 60 km/h est de 65 cm. Cette distance correspond, en moyenne, au cinquième de la longueur d'un véhicule. De ce fait, sur la maquette de simulation les déplacements des voitures entre deux acquisitions respectent cet ordre de grandeur. Les piétons, quant à eux, sont animés de mouvements difficilement prévisibles afin de se rapprocher des conditions réelles et de tester la fiabilité des algorithmes.

Pour chaque image, nous avons représenté les rectangles encadrant les objets et le tramway, tracé les distances minimales séparant le tramway des obstacles potentiels puis affiché les ordres de commande envoyés par le calculateur. Les rectangles dessinés en traits discontinus indiquent la position estimée de l'objet dans l'image suivante.

A la fin de chaque séquence d'images, nous reproduisons un fichier relatif à chaque prise de vue dans lequel apparaissent les coordonnées de chaque modèle rectangulaire des objets en mouvement ou arrêtés sur les rails, les distances minimales mesurées et estimées ainsi que les numéros se rapportant aux cas de figure du tableau IV.1. La mise à 1 de l'indicateur de présence sur voie (7<sup>e</sup> colonne) indique que l'objet est situé sur les rails. Ces valeurs sont données en pixels, chaque ligne et chaque colonne de l'image en comportant 256. Le point de coordonnées (0,0) est placé en haut à gauche. Les indices de collision, correspondant aux distances séparant les objets du tramway ont été établis de la manière suivante :

<u>Indice de risque de collision</u>	<u>Distance en pixels</u>
1	19
2	14
3	10
4	7

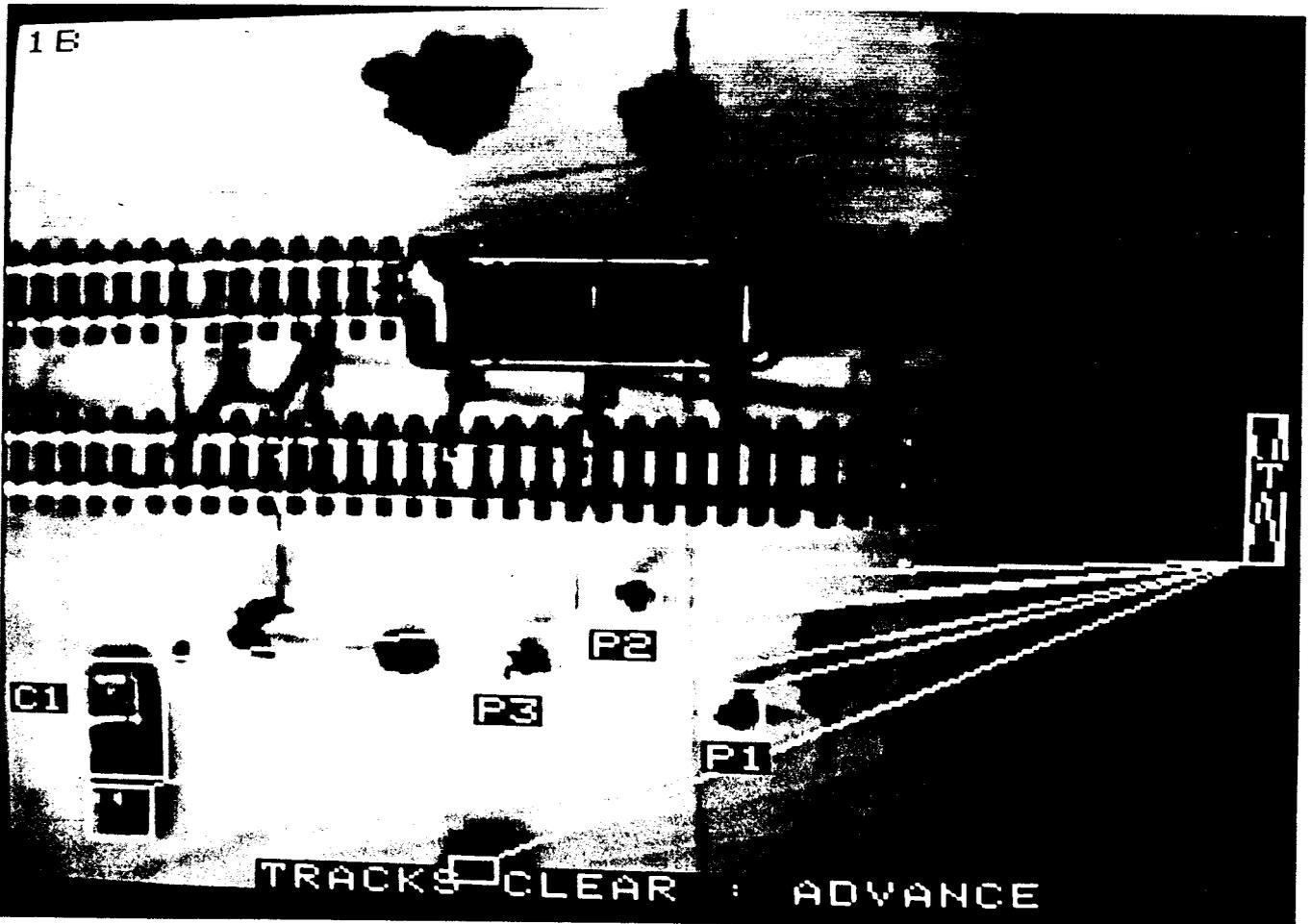
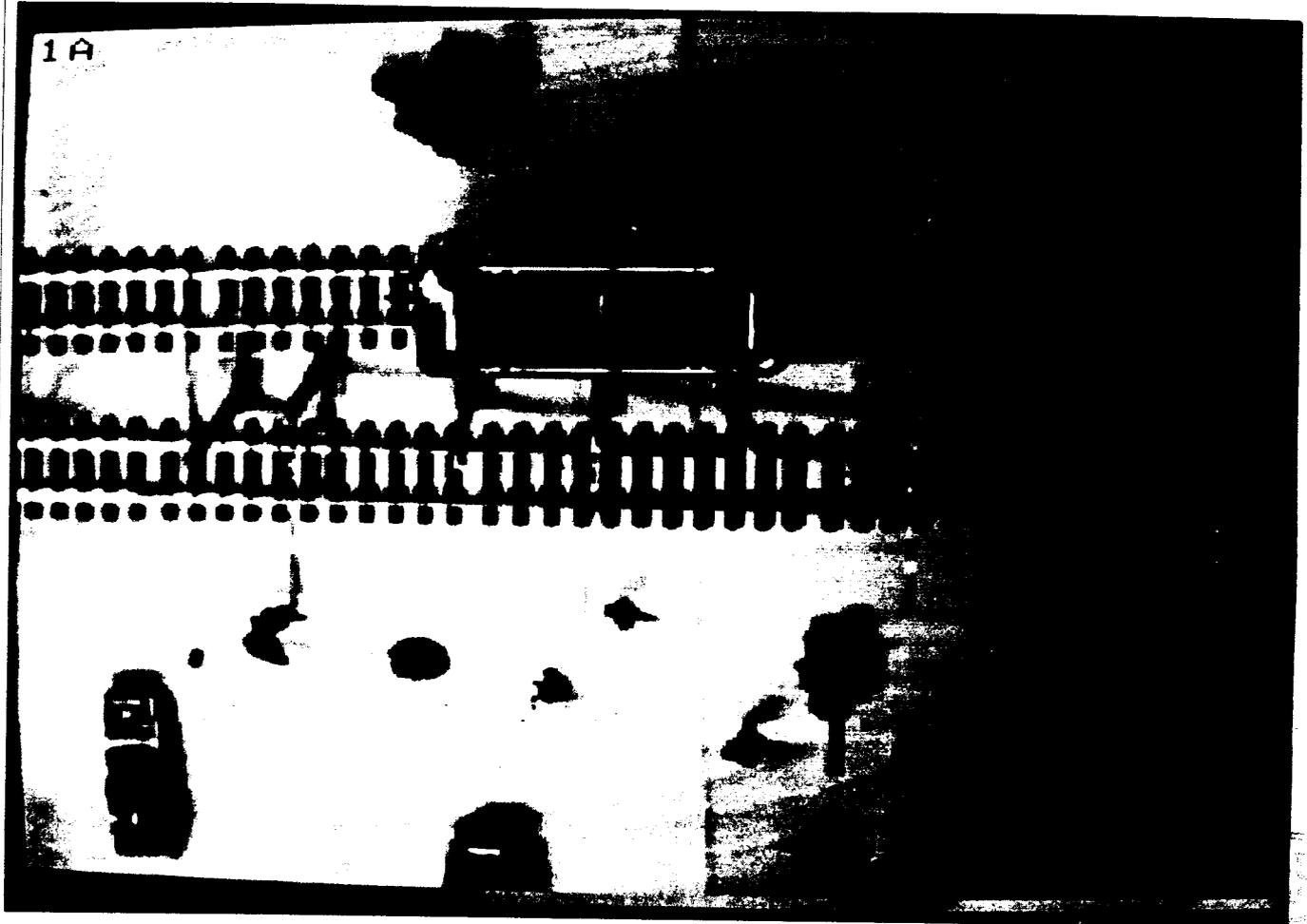
Ces valeurs ont été fixées en fonction des réactions des conducteurs observées lors de voyages effectués dans le poste de pilotage.

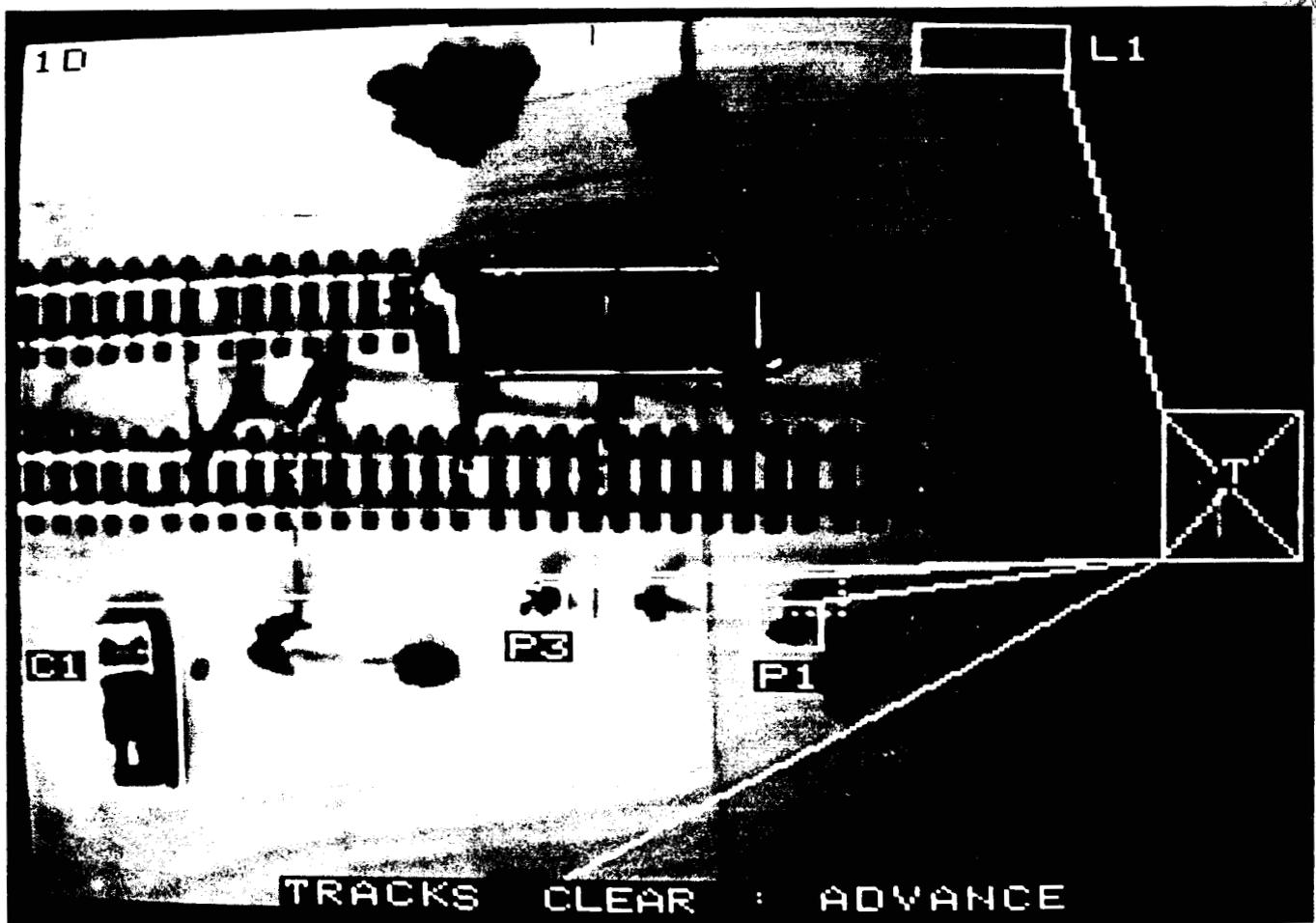
Afin de mieux suivre la traversée du carrefour par le tramway, il nous a semblé utile de commenter les principales situations rencontrées.

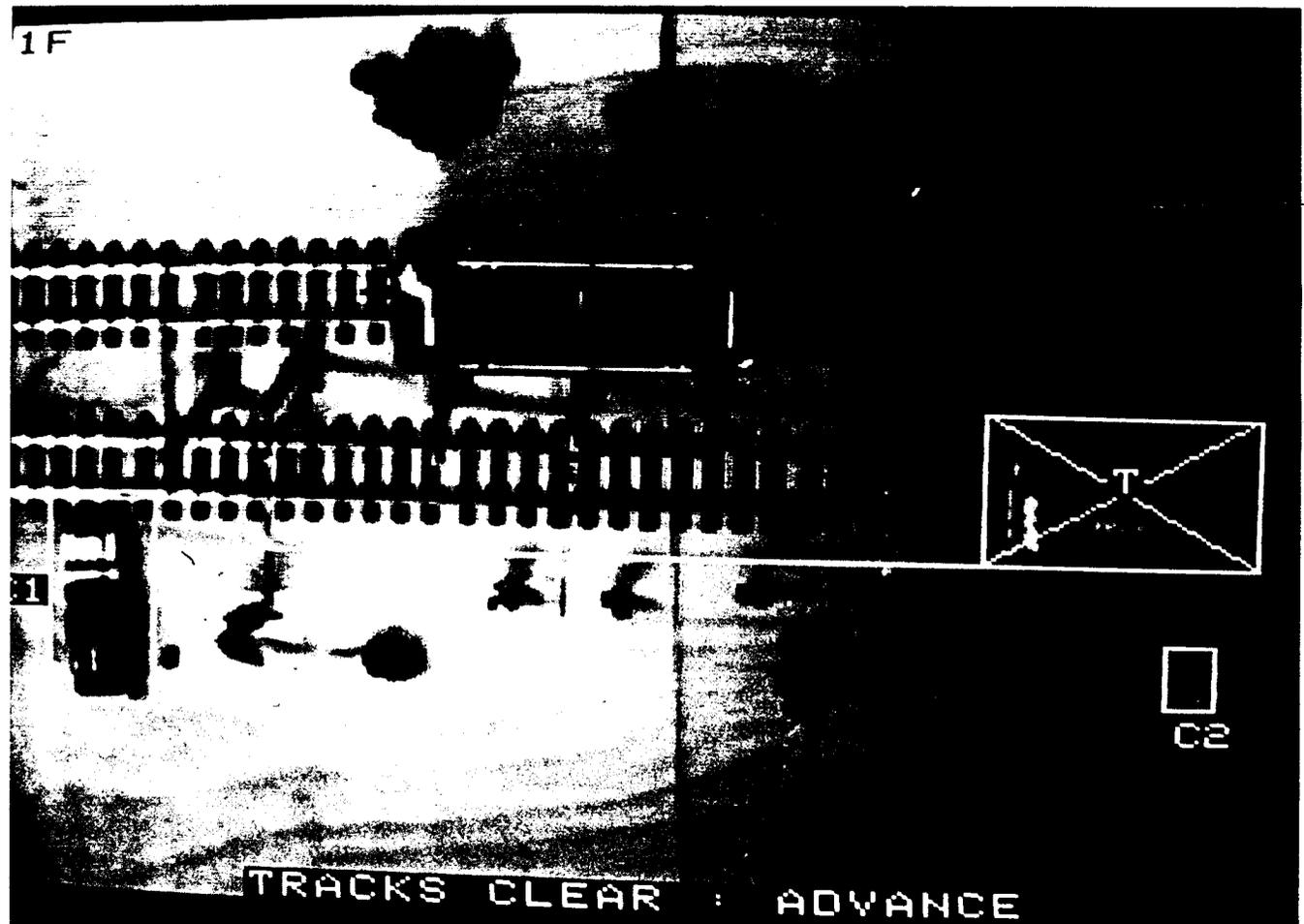
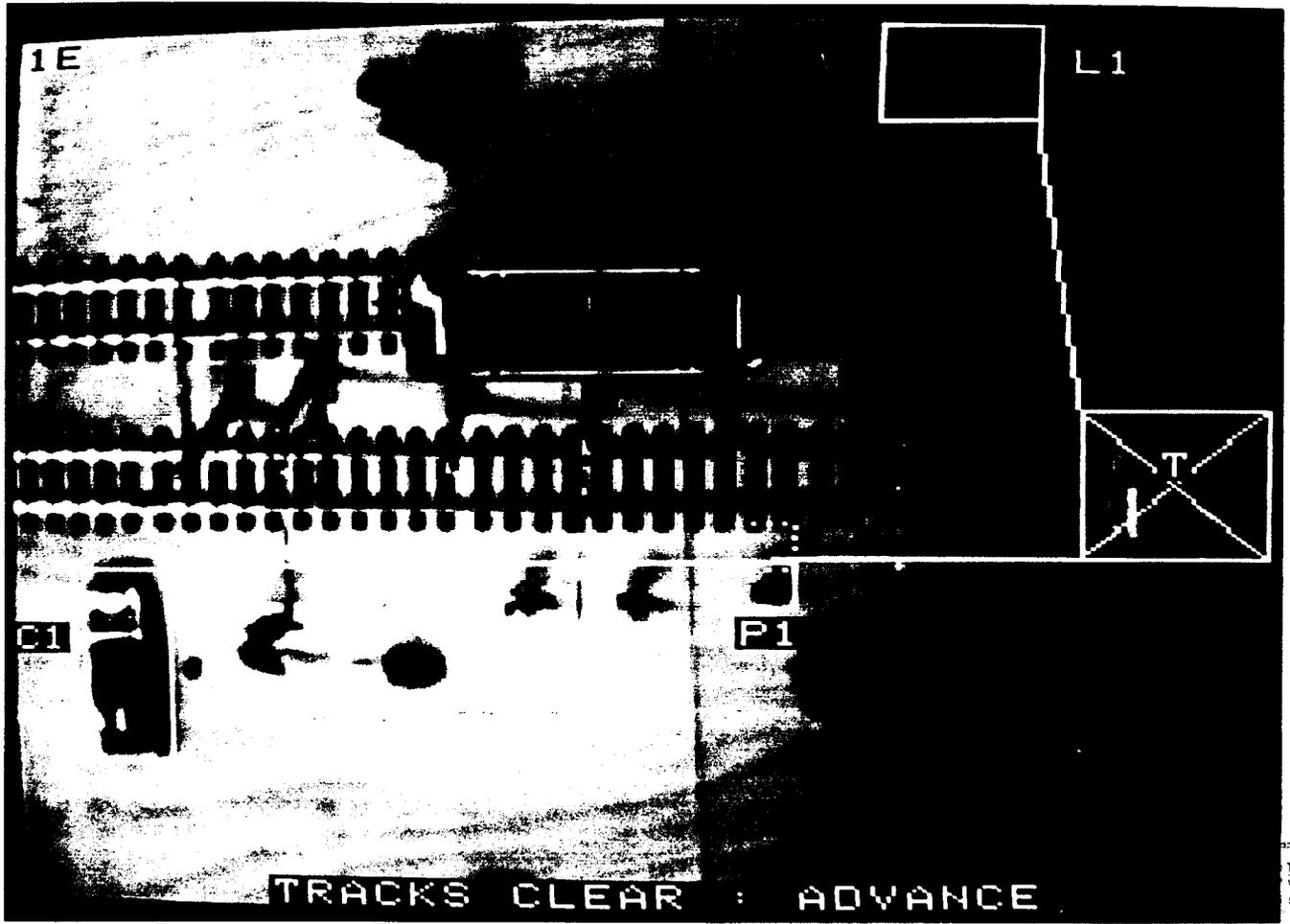
#### V.2. ANALYSE DE LA STRATEGIE DE CONDUITE

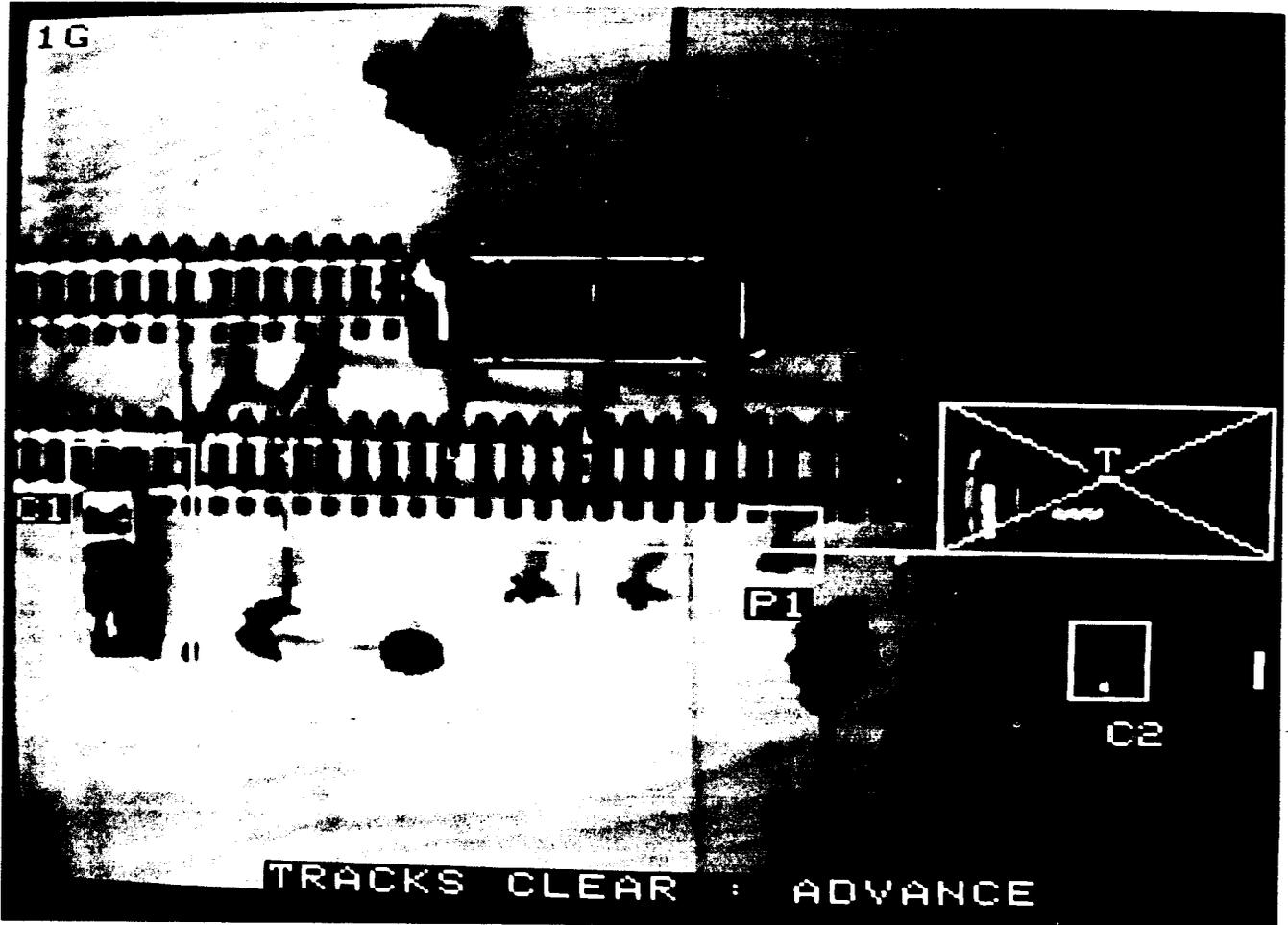
- . Image 1A : Première acquisition de la séquence
- . Image 1B : Les piétons P1, P2 et P3 ainsi que l'automobile C1 se sont déplacés entre les deux acquisitions successives. Ces objets étant éloignés du tramway, ce dernier peut avancer.
- . Image 1C : Le piéton P2 s'est arrêté. N'étant pas sur la voie, il n'est pas nécessaire de mémoriser sa position dans l'image 1B. On remarque que les deux rectangles composant la voiture C1, dans l'image 1B, se sont réunis. Cette modification n'apporte pas de difficultés supplémentaires dans le traitement.
- . Image 1D : Le camion L1 apparaît dans le champ de vision de la caméra. Le rectangle en pointillé près du piéton P1 indique qu'il a tendance à se rapprocher du tramway.
- . Image 1E : Le camion L1 continue de se rapprocher du tramway ainsi que le piéton P1.
- . Image 1F : Le piéton P1 et le camion L1 se sont arrêtés et on détecte la présence d'un nouvel objet dans la scène, à savoir la voiture C2. L'indicateur de présence sur voie relatif à l'objet C1 montre que celui-ci est présent sur les rails. Sa position est mémorisée et sera réactualisée dans les acquisitions suivantes.

- . Image 1G : Les véhicules C1 et C2 continuent leur déplacement. Le piéton P1 s'est brusquement avancé vers le tramway.
  
- . Image 1H : Bien que le piéton P1 ait instantanément inversé le sens de son déplacement et tend maintenant à s'éloigner des rails, il reste cependant trop proche du tramway. Sa position estimée indiquant qu'il a tendance à s'écarter du tramway, celui-ci reçoit l'ordre de freiner doucement.
  
- . Image 1I : Le piéton P1 s'est rapproché du tramway alors que sa position estimée montrait, dans l'image 1H, qu'il s'en éloignait. Cette situation souligne le peu de confiance que l'on peut attribuer à l'estimation sur la position d'un objet dont les mouvements sont peu réguliers. Le danger imminent que présente cette situation a été immédiatement analysé et le tramway reçoit l'ordre de s'arrêter immédiatement.
  
- . Image 1J : Le tramway est arrêté sur la voie, cet état étant affiché sur le haut de l'image. La voiture C1 s'est, elle aussi, arrêtée sur les rails. Comme les coordonnées de ce véhicule étaient réactualisées à chaque acquisition, du fait qu'il se déplaçait sur les rails, nous conserverons sa position dans la scène. La croix à l'intérieur du rectangle indique que l'objet est arrêté. Dans cette image, on aperçoit un nouvel objet qui entre dans le champ de vision de la caméra (piéton P4).
  
- . Image 1K : Bien que les piétons P1, P2 et P3 aient tendance, comme dans l'image 1J, à s'éloigner du tramway, la distance est trop faible pour qu'on donne l'ordre au tramway de repartir.
  
- . Image 1L : Les piétons s'étant arrêtés, on autorise le tramway à redémarrer.
  
- . Image 1M : Aucun obstacle proche ; le tramway peut continuer
  
- . Image 1N : On relève la même situation que pour l'image 1M. Cependant, le tramway devra s'arrêter devant la voiture C1 si elle ne modifie pas sa position dans les acquisitions suivantes.

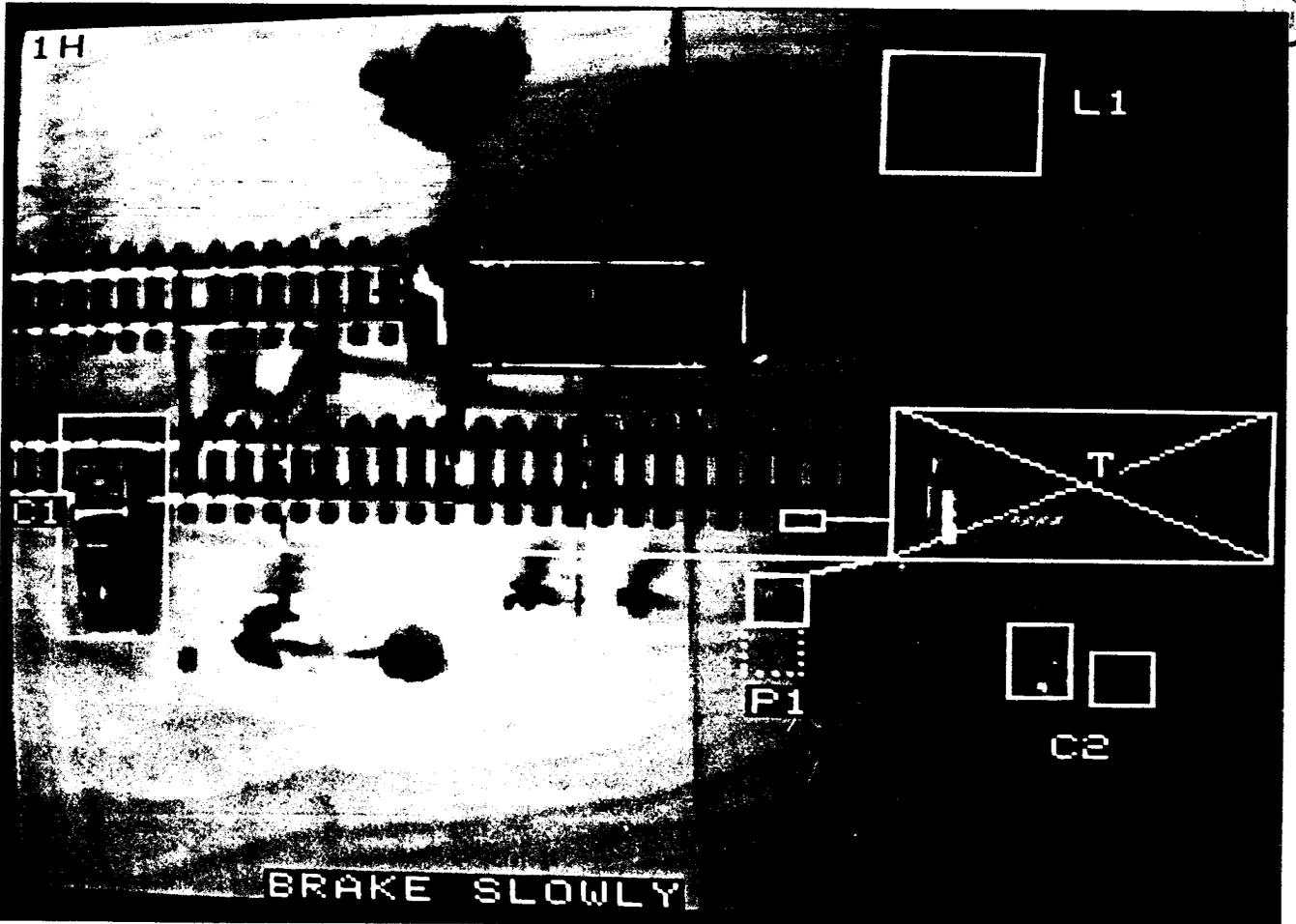


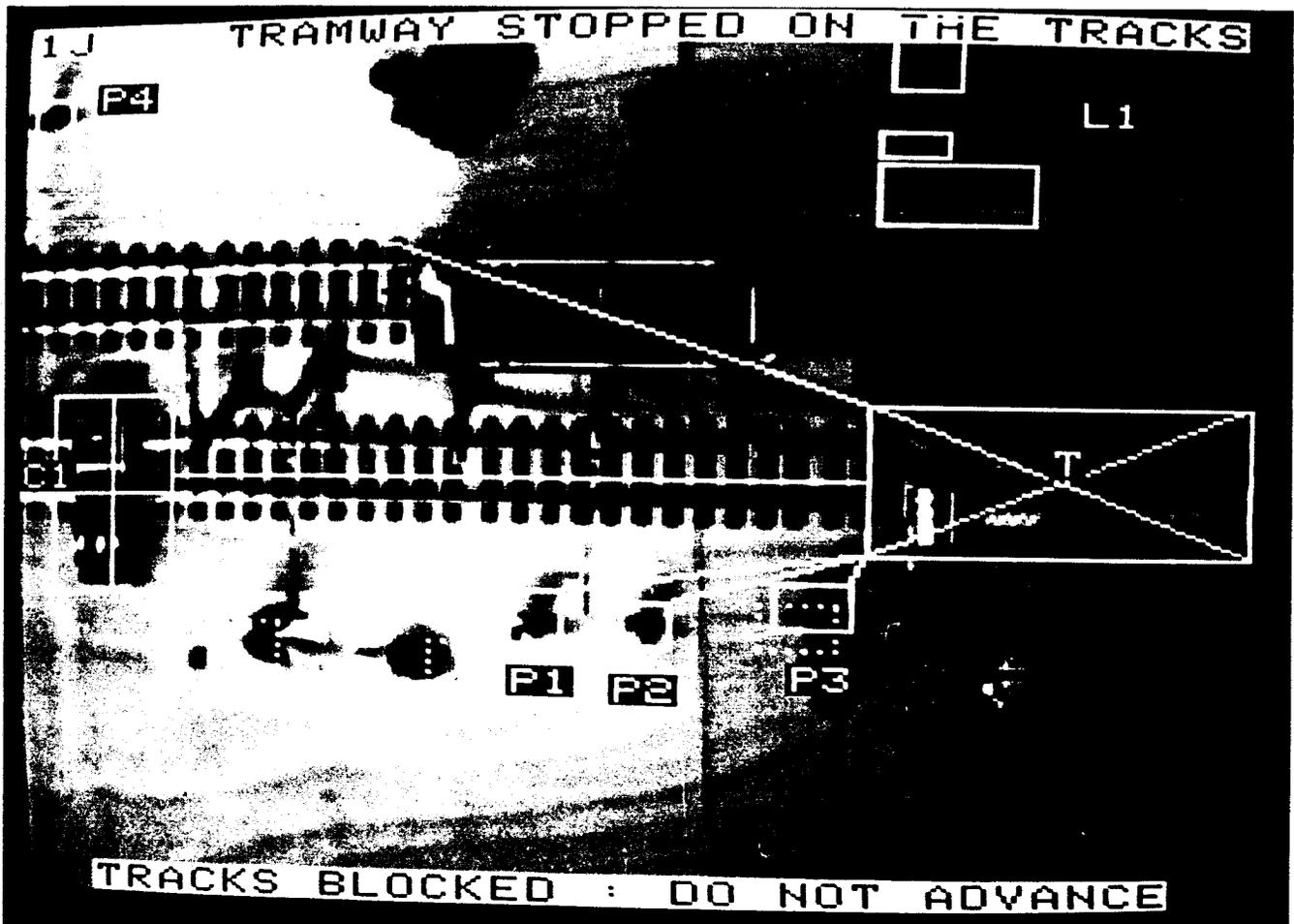
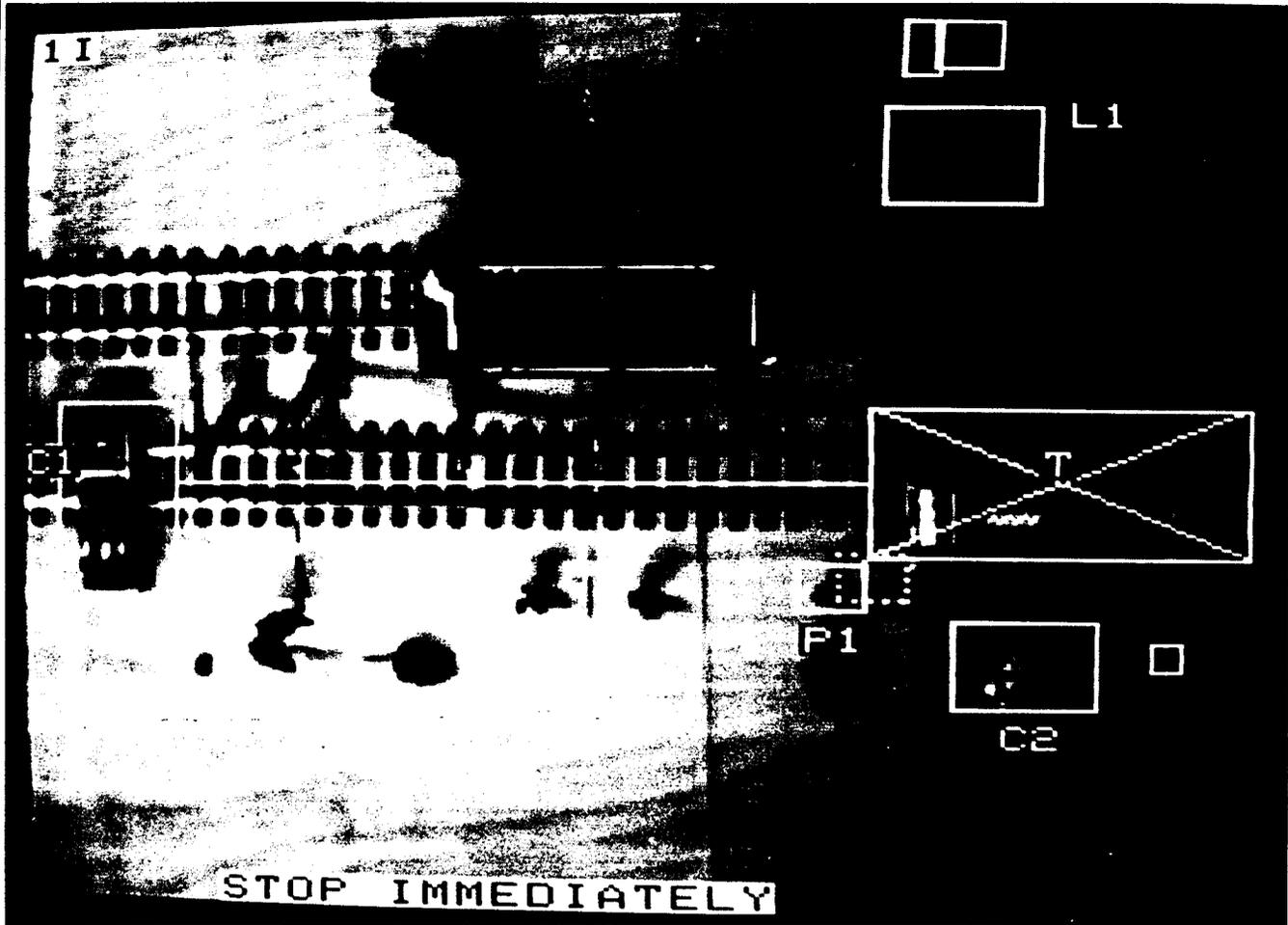






BU  
LE





1K TRAMWAY STOPPED ON THE TRACKS

P4

L1

B1

P3

P2

P1

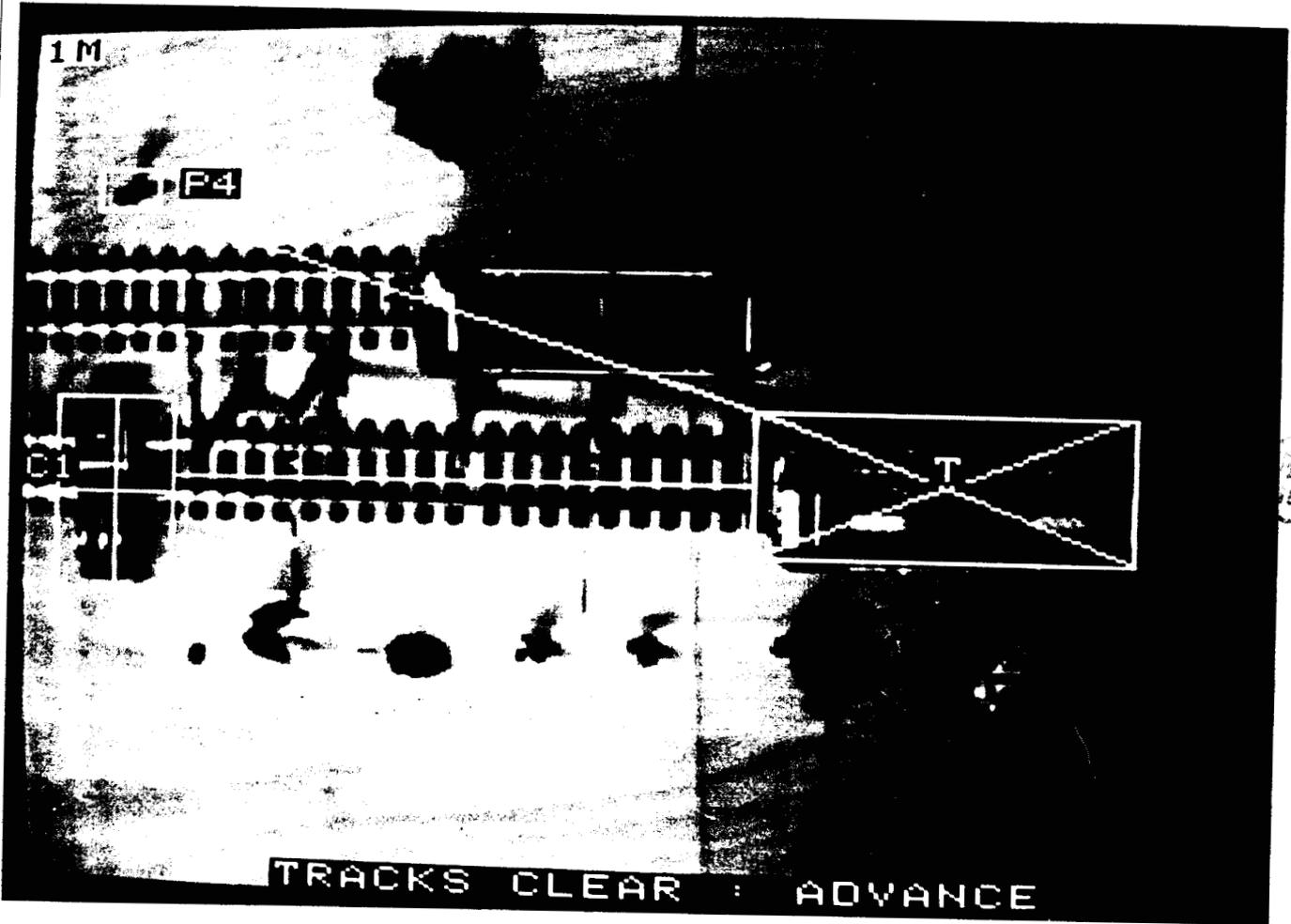
BU  
VILLE

TRACKS BLOCKED : DO NOT ADVANCE

1L

P4

TRACKS CLEAR : ADVANCE



S E Q U E N C E   N U M E R O   1

F            I            C            B

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
124	135	160	172	****	112	2	0
102	115	177	192	****	134	2	0
17	36	188	223	****	213	2	0
144	155	191	205	****	99	2	0
19	33	226	242	****	226	2	0
95	104	244	251	****	169	2	0

F            I            C            C

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
101	113	168	182	****	126	2	0
16	36	175	231	****	204	2	0
146	158	175	189	****	83	2	0

F            I            C            D

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
179	208	2	15	****	97	1	0
102	113	158	171	****	114	2	0
150	162	165	178	61	66	2	0
15	35	166	222	****	192	2	0
92	110	249	251	****	151	2	0

F            I            C            E

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
177	208	2	29	****	82	1	0
150	161	154	167	56	56	2	0
14	35	156	213	****	182	2	0

F            I            C            F

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
13	36	137	199	****	164	5	1
236	246	175	193	****	****	0	0

	F	I	C	G				
	XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
	12	40	125	189	****	149	5	1
	151	166	141	161	****	23	5	1
	215	230	172	194	****	****	0	0
	252	253	180	190	****	****	0	0

	F	I	C	H				
	XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
	176	206	11	45	****	****	0	0
	11	35	114	179	****	144	5	1
	158	166	140	145	****	13	6	1
	152	163	158	172	25	16	2	0
	202	214	171	192	****	****	0	0
	218	230	179	194	****	****	0	0

	F	I	C	I				
	XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
	178	185	2	17	****	****	0	0
	184	197	2	15	****	****	0	0
	175	205	26	53	****	****	0	0
	8	34	109	166	****	139	4	1
	159	172	155	168	-7	2	2	0
	189	217	171	196	****	****	0	0
	228	234	177	185	****	****	0	0

	F	I	C	J				
	XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
	8	34	109	166	****	139	4	1
	176	189	2	22	****	****	0	0
	2	10	22	33	****	180	1	0
	174	187	34	41	****	****	0	0
	174	204	43	60	****	****	0	0
	155	170	161	174	15	8	2	0
	102	114	164	180	117	60	2	0
	124	135	167	180	78	40	2	0

	F	I	C	K				
	XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
	8	34	109	166	****	139	4	1
	175	183	4	31	****	****	0	0
	184	195	20	32	****	****	0	0
	6	19	31	43	****	168	1	0
	173	178	41	63	****	****	0	0
	102	115	170	187	62	60	2	0
	154	163	170	183	29	19	2	0
	124	136	174	187	46	42	2	0

F I C L

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
8	34	109	166	****	127	4	1
14	26	33	45	****	150	1	0

F I C M

XMIN	XMAX	YMIN	YMAX	ESTDIS	DIS	FIGU	VOIE
8	34	109	166	****	118	4	1
16	29	42	54	****	135	1	0



V.3. CONCLUSION

La séquence présentée précédemment prend en compte la majeure partie des situations développées dans le chapitre IV (arrêt des véhicules sur la voie, arrêt du tramway....). Cette séquence, bien que complexe, a mis en évidence les performances de la procédure de conduite qui reproduit assez fidèlement le comportement d'un conducteur de tramway.



**C O N C L U S I O N**

Dans cette étude, nous avons développé une nouvelle méthode d'analyse de scènes dynamiques basée sur la détection des contours des objets en mouvement. L'approche adoptée permet d'analyser des scènes complexes, sous des conditions d'éclairage non contrôlées et sans aucune connaissance a priori sur le nombre, la forme et l'aspect des objets mobiles, ni sur leur environnement.

Dans le cadre de l'application de la méthodologie proposée à la conduite automatique d'un tramway, nous avons associé à chaque objet en mouvement détecté dans la scène, un modèle géométrique rectangulaire simple. Ces modèles ont permis de déterminer, à tout instant, les distances séparant le tramway des objets mobiles constituant son environnement. Cette mesure, associée à une prédiction sur le mouvement, est utilisée en vue de générer la conduite automatique d'un tramway.

L'approche proposée, appliquée dans le cadre de ce travail uniquement à la conduite d'un tramway, peut, moyennant quelques modifications, trouver d'autres champs d'application immédiats parmi lesquels nous citerons :

- La sécurité des robots :

Il convient d'éviter la collision du bras de robot avec un objet (mobile ou stationnaire) ou un opérateur humain pénétrant dans l'espace de travail du robot.

- La sécurité des chariots filoguidés :

La détection d'obstacles pouvant venir encombrer le passage des chariots filoguidés est actuellement confiée à des capteurs exteroceptifs. Le pilotage fonctionne alors souvent en tout ou rien. Un système tel que celui que nous avons développé permettrait de suivre l'évolution des objets entourant le chariot et d'éviter ainsi les arrêts intempestifs tout en augmentant la sécurité.

- La surveillance en site banalisé :

La possibilité de surveiller la pénétration d'objets dans des zones banalisées pour lesquelles il serait difficile d'implanter des capteurs tels que barrière optique, capteur infrarouge...., ouvre un vaste champ d'application aux techniques présentées dans ce travail. Nous pensons, en particulier, aux métros non équipés de portes palières séparant les quais de la voie.

La chute d'une personne sur les rails ou le jet d'un objet (cartable, poubelle....) se présente comme un obstacle qui, malgré l'environnement complexe, doit être immédiatement détecté et repéré en vue d'activer un processus d'alarme.

Ces applications s'avèrent intéressantes du fait qu'elles permettent de s'affranchir de certaines conditions rencontrées dans les scènes extérieures. En effet, bien que les résultats obtenus, sur maquette de simulation ou sur site extérieur, soient encourageants, il convient d'inscrire, dans les procédures de traitement d'images, des techniques permettant de prendre en compte les situations suivantes :

- \* Variations brutales d'éclairage entre deux acquisitions successives
- \* Différentes conditions atmosphériques (neige, pluie, brume...) et variations de l'éclairage durant un cycle journalier (petit jour, crépuscule, nuit...)
- \* Présence d'ombres qui accompagnent les objets dans leurs déplacements et qu'il convient de séparer avec justesse
- \* Importance des déplacements entre deux acquisitions successives. Si le déplacement d'une voiture entre deux prises de vue est élevé, la période d'échantillonnage temporel devra être courte. Mais, pour un tel intervalle de temps, le déplacement d'un piéton sera trop faible pour être détecté.

A ces différentes situations, qui n'ont pas été entièrement traitées dans cette étude, s'ajoute une limite inhérente à la procédure de modélisation. Bien qu'un modèle rectangulaire simplifie énormément la représentation d'un objet, le manque de précision apporte, dans le cas de scènes complexes, telles que celles que nous nous proposons d'étudier, des

ambiguïtés qu'il est difficile de résoudre. Il semblerait intéressant de proposer une modélisation réursive des objets en mouvement, le modèle à chaque instant étant fonction de l'image traitée, mais également des modèles de l'objet considéré dans les images précédentes. Cette poursuite de l'objet permettrait d'affiner le modèle à chaque acquisition d'image.

D'autre part, il est évident qu'une certaine connaissance a priori peut être utilisée pour identifier les objets détectés. Par exemple, la vitesse d'un objet en mouvement détecté peut déjà permettre de distinguer un véhicule d'un piéton. La position des objets détectés donne également une indication sur leur nature, les piétons se déplaçant plutôt sur les trottoirs ou perpendiculairement aux voies de trafic routier, les véhicules suivant, en général, le sens des voies.

Il semble difficile, en utilisant les méthodes de programmation usuelles, d'intégrer toutes ces règles pouvant aider à analyser l'image et de trouver une solution algorithmique pour chaque situation pouvant être rencontrée dans la réalité. Il semble plus intéressant d'utiliser les outils de l'intelligence artificielle en vue de se rapprocher au maximum du comportement d'un conducteur de tramway. Dans cette optique, nous proposons de structurer le système de conduite automatique d'un tramway en différents modules, chaque module étant à lui seul un système expert, qui seraient confiés à un superviseur. Ces modules, au nombre de trois, seraient "experts" en :

- Traitement d'image, pour analyser, identifier et interpréter l'environnement local en vue de donner une modélisation performante des objets en mouvement.
- Analyse de scène, pour analyser, identifier et interpréter l'environnement global en tenant compte des connaissances a priori sur les modèles, les conditions atmosphériques, etc.....
- Conduite, pour exploiter les informations recueillies précédemment en vue de définir une stratégie de conduite.

Malgré l'importance matérielle et logicielle que nécessite un tel système, il semble que cette voie soit, actuellement, la plus prometteuse.

Les récents travaux développés au Centre d'Automatique de l'Université des Sciences et Techniques de Lille 1 sont orientés dans cette voie et devraient avoir des retombées intéressantes pour l'analyse de scènes complexes et la conduite automatique de véhicules autonomes.



BIBLIOGRAPHIE

- ABD 79 - ABDOU I.E. , PRATT W.K  
Quantitative Design and Evaluation of Enhancement/  
thresholding Edge Detectors.  
Proceeding of IEEE, vol 67, n 5, mars 1979.
- AGG 75 - AGGARWAL J.K , DUDA R.O.  
Computer analysis of moving polygonal images.  
IEEE Computer, vol C.24, october 1975.
- AGG 83 - AGGARWAL J.K, MARTIN W.N  
Dynamic scene Analysis;  
Image sequence processing Ed T.S.HUANG, p 40-74, 1983
- AOK 84 - AOKI.M  
Detection of moving objects using line image sequence  
7<sup>th</sup> Conf on Patten Recognition, july 30 august 2  
MONTREAL CANADA, pp 784-786 , 1984.
- BAR 80 - BARNARD S.T , THOMPSON W.B.  
Disparity Analysis of Images.  
IEEE PAMI , vol 2, n 4 , july 1980.
- BAS 79 - BASSEVILLE M.  
Détection de contours :  
Méthodes et études comparatives.  
Ann Telecommunic 34, n 11-12, pp 559-579, 1979.
- BER 80 - BERS K.H., BOHNER M., GERLACH H.  
Object detection in image sequences.  
IEEE 5<sup>th</sup> International Conference on Pattern  
Recognition , MIAMI BEACH , 1980
- BER 83 - BERGMANN H.C.  
Analysis of different Displacement estimation  
Algorithms for Digital Television Signals.  
Image sequence processing , Ed T.S.HUANG, pp 215-235, 83
- BLA 81 - BLANZ W.E., REINHARDT E.R.  
Image Segmentation by Pixel Classification  
Pattern Recognition, vol 13 , n 4 , pp 293-298, 1981.
- BOL 82 - BOLLES R.C., CAIN R.A.  
Recognizing and locating partially visible objects.  
S.R.I Technical note 262 , juin 1982.
- BON 83 - BONNET P.  
Traitement d'images  
Cours de D.E.A , Université des Sciences et Techniques  
de LILLE I , 1983.

- BRI 83 - BRIOT M. ,BOISSIER L.  
Capteurs pour robots automoteurs;  
Les robots automoteurs PARIS ,4 novembre 1983
- BUR 82 - BURT P.J. ,YEN C, XU X.  
Local Correlation Measures for Motion Analysis, a  
comparative study.  
IEEE ,Conf Publication ,vol 14 ,LAS VEGAS ,june 1982
- CAF 76 - CAFFORIO C. ,ROCCA F.  
Methods for Measuring Small Displacement of  
Television Images.  
IEEE Transformations on Information Theory ,vol IT 12  
n 22 ,septmber 1976.
- CAF 83 - CAFFORIO C. ,ROCCA F.  
The differential method for image motion estimation.  
Image Sequence Processing ,Ed T.S.HUANG,pp 104-124,83
- CAM 84 - CAMILLERAP J. , LEPLUMEY Y.  
Codage du contexte et détection au moyen de règles de  
réécritures.  
4<sup>eme</sup> Congrès reconnaissance des formes ,PARIS  
25-27 janvier 1984.
- CHA 81 - CHATILA R.  
Système de navigation pour un robot mobile autonome.  
Thèse de docteur ingénieur ,TOULOUSE ,juillet 1981.
- CHE 80 - CHENG N ,BIRK J.R ,KELLEY R.  
Estimation workpiece pose using the feature points  
method.  
IEEE ,vol AC 25 ,n 6,Decembre 1980
- CHE 83 - CHEBEL B.  
Application de l'analyse de structure et de la  
fonction entropie du traitement du signal.  
Thèse de docteur ingénieur LILLE I ,1983.
- CHO 72 - CHOW C.K ,KANEKO T.  
Boundary detection of radiographic images by a  
threshold method.  
Proc IFIP ,AMSTERDAM, pp 130-140 , 1972.
- CHO 77 - CHOW W.K ,AGGARWAL J.K.  
Computer analysis of planar curvilinear moving  
images.  
IEEE Transaction on computers , vol C 26,february 77.
- COR 83 - CORREIA M.B.  
Contribution à la localisation d'objets polyedrique  
par stéréovision en robotique.  
Thèse de docteur ingénieur ,TOULOUSE,mars 1983.

- DAV 75 - DAVIS L.  
A survey of edge detection techniques.  
Computer vision ,graphics and image processing,vol 4  
pp 248-270 , 1975.
- DAV 83 - DAVIS L ,WU Z ,SUN H.  
Contour-based motion estimation.  
Computer vision ,graphics and image processing ,23,  
pp 313-326 , 1983.
- DEP 84 - DEPARIS J.P ,KUHN F.  
Perspectives de développement des automatismes dans  
les transports guidés de surface.  
IRT/CRESTA ,rapport technique , juin 1984
- DIC 84 - DICKINSON K.W ,WATERFALL R.C.  
Image processing applied to traffic :  
1: a general review.  
Traffic Engineering and control , pp 6-13,janvier 84.
- DOY 62 - DOYLE W.  
Operations useful for similarity-invariant pattern  
recognition.  
J Assoc . Comput. Mach ,n 9, pp 259-267, 1962.
- DRE 78 - DRESCHLER L ,NAGEL H.H.  
Using "affinity" for extracting images of moving  
objects from TV frame sequences.  
Rapport interne B 44/78  
Institut für informatik,D 2000 HAMBURG ,1978.
- DRE 82 - DRESCHLER L ,NAGEL H.H.  
Volumetric model and 3D trajectory of a moving car  
derived from monocular TV frame sequences of a  
street scene.  
Computer Vision graphics and image processing 20,  
pp 199-228,november 1982.
- DRE 83 - DRESCHLER L ,NAGEL H.H.  
On the selection of critical points and local  
curvature.  
Image sequence processing ,Ed T.S.HUANG,pp 457-470.
- DUD 72 - DUDA R ,HART P.  
Use of the HOUGH transformation to detect lines and  
curves in pictures.  
ACM ,vol 15 ,n 1 ,january 1972.
- END 71 - ENDLICH R.M , WOLF D.E ,HALL D.J.  
Use of pattern recognition technique for determining  
cloud motions  
J of applied meteorology 10 , pp 105-117,1971.

- EQU 84 - Equipements de recueil de données de trafic routier.  
Journée d'information DTCS du SETRA ,PARIS,11 oct 84.
- FEN 79 - FENNEMA C.L ,THOMPSON W.  
Velocity determination in scenes containing several  
moving objects.  
Computer vision graphics and image processing 9,  
pp 301-315, 1979
- FER 82 - FERRER M.  
Système multisenseur de perception 3D pour le robot  
mobile hilare.  
Thèse de 3<sup>eme</sup> cycle ,Université P.SABATIER,TOULOUSE,82
- FRE 78 - FREEMAN M.  
Shape description via critical points.  
Pattern recognition 10, pp 156-166 , 1978.
- FRE 79 - FREEMAN M.  
Curved surface representation utilizing data  
extracted from multiple photographic image.  
Workshop on the representation of 3D objects ,mai 79.
- GEU 83 - GEUEN W.  
A fast edge detection algorithm matching visual  
contour perception.  
Image sequence processing, Ed T.S.HUANG ,pp 483-492
- HAY 83 - HAYNES S.M ,JAIN R.  
Detection of moving edges.  
Computer vision graphics and image processing 21,  
pp 345-367 , 1983
- HIL 83 - HILDRETH E.C.  
The detection of intensity changes by computer and  
biological vision systems.  
Computer vision graphics and image processing 22  
pp 1-27 , 1983
- HOR 81 - HORN B.K.P ,SCHUNCK B.G.  
Determining optical flow.  
Artificial intelligence ,vol 17 , 1981 .
- HUA 81 - HUANG T.S.  
Motion estimation  
Image sequence analysis.  
T.S.HUANG ,Springer Verlag ,BERLIN , 1981
- IKE 83 - IKEUCHI K.  
Constructing a depth map from images.  
Technical rep. ,A.I. MEMO n 744 ,MASSACHUSETTS,Aug 83

- JAC 80 - JACOBUS C, CHIEN R ,SELANDER J.  
Motion detection and analysis of matching graphs of  
intermediate level primitives.  
IEEE PAMI ,vol 2 ,n 6 ,november 1980.
- JAI 79 - JAIN R ,MARTIN W.N.  
Segmentation through the detection of changes due  
to motion.  
Computer Vision graphics and image processing n 11 ,  
pp 13-34 , 1979.
- JAI 80 - JAIN R.  
Peripheral processes in machine perception of motion.  
IEEE 5<sup>th</sup> international conference on pattern  
recognition ,MIAMI BEACH ,1980.
- JAI 81a- JAIN R.  
Dynamic scene analysis using pixel-based processes.  
IEEE ,computer 14(8) ,august 1981.
- JAI 81b- JAIN R.  
Extraction of motion information from peripheral  
processes.  
IEEE PAMI ,vol 3 ,n 5 ,september 1981.
- JAI 84 - JAIN R.  
Segmentation of frame sequences obtained by a moving  
observer.  
IEEE PAMI ,vol 6 ,n 5 ,september 1984.
- KEA 82 - KEARNEY J.K ,THOMPSON W.B.  
Gradient based estimation of disparity .  
IEEE conf publication ,LAS VEGAS ,vol 14,juin 1982.
- KIT 82 - KITCHEN L ,ROSENFELD A.  
Gray-level corner detection.  
Pattern recognition letters 1, n 2 ,pp 95-102 ,dec 82
- LAB 82 - LABIT C.  
Estimation de mouvement dans une séquence d'images de  
télévision.  
Thèse de docteur ingénieur ,UNIVERSITE DE RENNES 1,  
Février 1982.
- LAW 83 - LAWTON D.T.  
Processing translational motion sequences.  
Computer vision graphics and image processing 22 ,  
pp 116-144 , april 1983.

- LEC 83 - LECERF J.M.  
Conception d'une structure logicielle redondante en  
vue d'améliorer la sureté des prises de décision en  
temps réel.  
Thèse de 3<sup>ème</sup> cycle , Université des Sciences et  
Techniques de LILLE I , 1983.
- LEE 70 - LEESE J.A , NOVAK C.S , TAYLOR V.R.  
The determination of cloud pattern motions from  
geosynchronous satellite image data.  
Pattern recognition , vol 2 , pp 279-292 , 1970.
- LEG 82 - LEGTERS G.R , YOUNG T.Y.  
A mathematical model for computer image tracking.  
IEEE PAMI , vol 4 , n 6 , pp 583-594 , nov 1982.
- LEG 84 - LE GUILLOUX Y.  
Détermination automatique du mouvement dans une  
séquence d'images. Intérêt pour la compréhension.  
Thèse de docteur ingénieur ENST , juin 1984.
- LEL 79 - LELIEVRE J.L.  
Processus d'acquisition et de traitement d'images en  
temps réel : application à l'inspection automatique.  
Thèse de docteur ingénieur , LILLE I , juin 1979.
- LEN 83 - LENZ R , GERHARD A.  
Image sequence coding using scenes analysis and  
spatio-temporal interpolation.  
Image sequence processing, Ed T.S.HUANG, pp 264-274, 83.
- LEV 83 - LEVINE M.D , NOBLE P.B.  
A rule-based system for characterizing blood cell  
motion.  
Image sequence processing, Ed T.S.HUANG, pp 663-709, 83.
- LIM 75 - LIMB J.O , MURPHY J.A.  
Estimation velocity of moving images in television  
signals.  
Computer vision graphics and image processing 4 ,  
pp 311-327 , 1975.
- LOU 83 - LOUCHET J.  
Détection et classification de cibles en imagerie  
infra-rouge.  
Agard Cong. Proc. , pp 7.1-7.11 , 1983 .
- LUC 81 - LUCAS B , KANADE T.  
An iterative image registration technique with an  
application to stereo vision.  
Proceeding of the 5<sup>th</sup> international joint conference  
on artificial intelligence , august 1981.

- MAI 79 - MAITRE H.  
Système d'entrée/sortie d'images.  
Session méthodes de traitement des images ,sept 1979.
- MAI 84 - MAITRE H ,LOPEZ K.J.  
Transformation de HOUGH:théorie et application à la  
détection de symétries circulaires.  
4<sup>eme</sup> congrès reconnaissance des formes , PARIS ,  
25\_27 janvier 1984.
- MAR 78 - MARTIN W.N ,AGGARWAL J.K.  
Survey-dynamic scene analysis.  
Computer vision graphics and image processing 7,  
pp 356-374 , 1978.
- MAR 79 - MARTIN W.N ,AGGARWAL J.K.  
Computer analysis of dynamic scenes containing  
curvilinear figures .  
Pattern recognition ,vol 11 ,pp 169-178, 1979.
- MIC 78 - MICLET,GRENIER,KALISZEWSKI.  
Reconnaissance de formes.  
ENST 1978.
- MOR 77 - MORAVEC H.P.  
Towards automatic visual obstacle avoidance.  
5<sup>th</sup> IJCAI ,CAMBRIDGE ,August 1977.
- MOR 79 - MORAVEC H.P.  
Visual mapping by a robot rover.  
Procedure IJCAI, TOKYO ,august 1979
- MOR 83 - MORAVEC H.P.  
The stanford cart and the CMU rover.  
Proceedings of the IEEE ,vol 71 ,n 7 ,july 1983.
- NAG 83 - NAGEL H.H.  
Displacement vectors derived from second order  
intensity variations in image sequences.  
Computer vision graphics and image processing 21 ,  
pp 85-117 , 1983 .
- OSH 79 - OSHIMA M ,SHIRAI Y.  
A scene description method using three dimensional  
information.  
Pattern recognition ,vol 11 ,pp 9-17 , 1979.
- PRA 80 - PRAZDNY K.  
Egmotion and relative depth map from optical flow.  
Biological cybernetics 36 ,pp 87-102, 1980.

- PRA 81 - PRAZDNY K.  
A simple method for recovering relative depth map in  
the case of a translating sensor.  
Procedure 7<sup>th</sup> IJCAI VANCOUVER , aout 1981.
- RAS 80 - RASHID R.F.  
Towards a system for the interpretation of moving  
light displays.  
IEEE PAMI , vol 2 , n 6 , pp 574-581 , nov 1980 .
- ROA 79 - ROACH J.W , AGGARWAL J.K;  
Computer tracking of objects moving in space.  
IEEE PAMI , vol 1, n 2 , april 1979.
- ROS 70 - ROSENFELD A.  
Connectivity in digitized pictures;  
J.Assoc. computing machine, vol 17 , pp 146-160, 1970.
- ROS 75 - ROSENFELD A , WESZKA J.S.  
An improved method of angle detection on digital  
curves.  
IEEE computer , september 1975.
- ROS 76 - ROSENFELD A , HUMME P.R , ZUCKER S.W.  
Scene labeling by relaxation operation.  
IEEE trans. on syst. man. and cyb. , 6.6 ,  
pp 420-430 , 1976 .
- ROS 82 - ROSENFELD A.  
Vision par ordinateur.  
Rocquencourt , juin 1982.
- SAM 84 - SAMY R.  
Traitement de séquence d'images TV .Application  
à la poursuite de cibles.  
Thèse de 3<sup>eme</sup> cycle , Université d'AIX -MARSEILLE ,  
mai 1984 .
- SCH 78 - SCHACHTER B.J , ROSENFELD A.  
Some new methods of detecting step edge in digital  
picture.  
Communication of the ACM , vol 1 , n 2 , fevrier 1978.
- SNY 80 - SNYDER W.E , RAJALA S.A , HIRZINGER G.  
Image modeling the continuity assumption and tracking  
ICPR , vol 1 , december 1980.
- SPO 83 - SPOER P.  
Déplacement estimation for objects on moving  
background .  
Image sequence processing , Ed T.S.HUANG ,  
pp 425-437 , 1983.

- SQU 83 - SQUALI A.  
Méthodes de détermination d'un champ de vitesse sur un modèle hydraulique par analyse de séquences d'images.  
Thèse de docteur ingénieur ,PARIS SUD , 1983.
- STU 80 - STUCK F.  
Réalisation d'un système adaptatif de traitement d'images pour l'identification et la localisation de pièces en robotique.  
Thèse de 3<sup>eme</sup> cycle ,TOULOUSE ,déc 80.
- TAN 82 - TANG I.S ,SNYDER W.E ,RAJALA S.A.  
Extraction of moving objects in textured dynamic scenes .  
IEEE conf.publication , vol 14 ,june 1982,LAS VEGAS.
- THO 80 - THOMPSON W.B.  
Combining motion and contrast for segmentation .  
IEEE PAMI ,vol 2, n 6, november 1980.
- THO 81a- THOMPSON W.B ,BARNARD S.T.  
Lower-level estimation and interpretation of visual motion.  
IEEE computer , august 1981 .
- THO 81b- THORAVALL Y.  
C.C.D. des registres à transfert de charges.  
Micro systeme ,sept/oct 1981.
- TSU 80 - TSUJI S ,OSADA M ,YACHIDA M.  
Tracking and segmentation of moving objects in dynamic line images.  
IEEE PAMI ,vol 2 ,n 6 ,november 1980.
- WAD 78 - Wide Area Detection System  
Report n° FHWA RD 77.86  
Federal highway administration ,WASHINGTON, 1978.
- WAL 75 - WALTZ D;  
Understanding line drawings of scene with shadows.  
Psychology of computer vision ,WINSTON Ed ,  
Mc GRAW HILL , 1975.
- WAT 84 - WATERFALL R.C ,DICKINSON K.W.  
Image processing applied to traffic .2 :  
practical experience.  
Traffic engineering and control ,pp 60-67 ,  
feb 1984 .
- WIL 80 - WILLIAMS T.D.  
Depth from camera motion in a real world scene.  
IEEE PAMI ,vol 2 , n 6, november 1980.

- WIL 81 - WILLIAMS T.D.  
Computer interpretation of a dynamic image from a moving vehicle.  
Coins tech. rep. 81.22.
- WIL 83 - WILLIAMS T.D ,GLAZER F.  
Comparison of feature operators for use in matching pairs.  
Image sequence processing ,Ed T.S.HUANG ,  
pp 395-423 ,1983.
- WOH 83 - WOHN K ,DAVIS L.S ,THRIFT P.  
Motion estimation based on multiple local constraints and nonlinear smoothing.  
Pattern recognition , vol 16 ,n 6 ,1983.
- YAC 78 - YACHIDA M ,ASADA.M, TSUJI S.  
Automatic motion analysis system of moving objects from the records of natural processes.  
Proc.int. joint conf. on pattern recognition ,KYOTO ,  
pp 726-730 , JAPAN 1978.
- YAC 81 - YACHIDA M ,ASADA M, TSUJI S.  
Automatic analysis of moving images.  
IEEE PAMI ,vol 3 , n 1, january 1981.
- YAC 83 - YACHIDA M.  
Determining velocity maps by spatio-temporal neighborhoods from image sequences.  
Computer vision graphics and image processing 21,  
pp 262-279 ,february 1983.
- YAL 80 - YALAMANCHILI S ,MARTIN W.N ,AGGARWAL J.K.  
Differencing operations for the segmentation of moving objects in dynamic scenes.  
IEEE 5<sup>th</sup> international conference on pattern recognition ,MIAMI BEACH , 1980.
- YAL 82 - YALAMANCHILI S ,MARTIN W.N ,AGGARWAL J.K.  
Extraction of moving object descriptions via differencing.  
Computer vision graphics and image processing 18,  
pp 188-201 ,february 1982.
- YAK 78 - YAKIMOVSKY Y ,CUNNINGHAM R.  
A system for extracting three-dimensional measurements from a stereo pair of TV cameras.  
Computer vision graphics and image processing 7,  
pp 195-210 , 1978.

- R E M E R C I E M E N T S -

---

Ce court métrage a été filmé dans les studios du Centre d'Automatique de l'Université des Sciences et Techniques de Lille 1.

La mise en scène a été orchestrée par Monsieur P. VIDAL, dont il convient de souligner le haut degré de professionnalisme.

La production a été gérée par Monsieur Y. DAVID, de l'Institut de Recherche des Transports, assisté pour les problèmes techniques de Monsieur J.P. DEPARIS. Ils ont toujours affiché une confiance totale quant à la réussite commerciale de cette oeuvre.

Le lancement de ce film a été effectué par Monsieur B. DUBUISSON, de l'Université de Compiègne, qui a su trouver, en avant-première, un public intéressé.

La censure a été confiée à Monsieur J.G. POSTAIRE. Maniaque du ciseau et de la colle, il a su imposer ses directives tout en laissant croire à une indépendance absolue. Sa disponibilité permanente et de grandes qualités humaines ont largement contribué à la réalisation de ce film.

Monsieur P. BONNET, spécialiste de l'éclairage, occupait aussi le poste d'intendant. Dommage que ses économies d'énergie nous aient contraint à filmer plusieurs scènes dans l'obscurité.

La prise de son a été assurée par B. CEURSTEMONT et M. POLLART qui m'ont habitué à parler devant un micro (...ordinateur).

Derrière les caméras se trouvaient M. AMEZIANE, H. DELEPLANQUE, F. DELBARRE. Ils ont toujours su se focaliser avec précision sur le sujet principal.

Le script a été réalisé par A. PIGNON aidée, pour le sous-titrage en anglais, par C. RAJAGOPALAN.

Le spécialiste de la chambre noire, Monsieur J. HOUZE, s'est occupé de la reproduction des films en grande série.

Que toutes ces personnes, ainsi que celles qui n'ont pu être nommées mais qui ont contribué à créer une ambiance décontractée sur le plateau, se voient remerciées.

