

THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

pour l'obtention du titre de

DOCTEUR

En productique : Automatique et Informatique Industrielle

par

Louahdi KHOUDOUR

**ANALYSE SPATIO-TEMPORELLE DE SEQUENCES
D'IMAGES LIGNES. APPLICATION AU COMPTAGE DES
PASSAGERS DANS LES SYSTEMES DE TRANSPORT**

Soutenu le 21 janvier 1997 devant la commission d'examen :

MM.

P. VIDAL

Président

Professeur à l'USTL

A. CHEHIKIAN

Rapporteur

Professeur à l'INPG

J. JACOT

Rapporteur

Professeur à l'EPFL

M. BEAUBAT

Examineur

Directeur du centre de Lille-Villeneuve d'Ascq de l'INRETS

J-P. DEPARIS

Examineur

Chargé de recherches à l'INRETS

L. DUVIEUBOURG

Co-directeur de thèse

Maitre de conférence Université du littoral

J-G. POSTAIRE

Co-directeur de thèse

Professeur à l'USTL



AVANT PROPOS

Les personnes qui m'ont aidé et à qui j'adresse des remerciements appartiennent à deux institutions : Le centre d'automatique de Lille (CAL) et l'Institut National de Recherches sur les Transports et leur Sécurité (INRETS).

AU CAL :

Monsieur Pierre Vidal, professeur à l'USTL, directeur du Centre d'Automatique de Lille, qui me fait l'honneur de présider le jury de cette thèse.

Monsieur J-G. POSTAIRE, professeur à l'USTL, responsable de l'équipe Image et Décision, pour la confiance immédiate qu'il m'a témoignée, pour sa disponibilité et son optimisme permanent dans la direction de cette thèse.

Luc DUVIEUBOURG, maître de conférence à l'Université du Littoral, pour son soutien et son suivi dans la co-direction de ce travail.

Yuefeng WAN pour sa gentillesse et l'aide précieuse apportée dans la programmation des algorithmes.

Tous les membres de l'équipe Image et Décision dont j'ai fait la connaissance pendant ce travail de thèse.

A L'INRETS :

Jean-Pierre DEPARIS, chargé de recherches, pour son encadrement, son suivi, sa confiance en moi ainsi que son infinie patience à certaines occasions.

Marc HEDDEBAUT, directeur du LEOST et mon supérieur hiérarchique, pour l'accueil de cette thèse au sein de son unité de recherches.

Lionel BARANES, directeur général adjoint, pour ses encouragements et l'intérêt qu'il a porté à mon travail.

Michel BEAUBAT, directeur régional, d'accepter de siéger au jury de cette thèse.

Bruno MEUNIER, pour sa gentillesse, son intérêt et sa très grande implication dans l'accomplissement de ce travail.

Enfin tous mes collègues de l'INRETS qui m'on très gentiment accueilli et supporté.

Enfin mes remerciements vont à Mrs Alain CHEHIKIAN, professeur à l'Institut National Polytechnique de Grenoble et Jacques JACOT, professeur à l'Ecole Polytechnique Fédérale de Lausanne, pour avoir accepté d'être les rapporteurs de cette thèse.

Table des matières

TABLE DES MATIERES

<i>Chapitre I</i>	3
LE COMPTAGE AUTOMATIQUE DES PIETONS PAR VISION ARTIFICIELLE : UN OUTIL D'AIDE A L'EXPLOITATION DANS LES TRANSPORTS PUBLICS	3
1.1 Le cahier des charges du système.....	5
1.2 Les systèmes de comptage automatique existants.....	6
1.2.1 Détection par contact.....	6
1.2.2 Détecteurs actifs de proximité.....	7
1.2.3 Détecteurs passifs de proximité.....	8
1.2.4 Vision et analyse de l'image	9
1.3 Organisation de la thèse	9
<i>Chapitre II</i>	13
LES SYSTEMES DE COMPTAGE EXISTANTS	13
2.1 Introduction	13
2.2 Les systemes de comptage à base de caméras matricielles	16
2.2.1 Traitements macroscopiques	16
2.2.1.1 Estimation de la densité et de la vitesse d'une foule par S. Velastin	17
Calcul de la densité de foule.....	17
Estimation du mouvement de la foule.....	18
2.2.1.2 Les travaux de A. Rourke et M.G. Bell.....	18
2.2.1.3 Prise en compte des piétons dans la régulation du trafic routier	19
Principe de l'algorithme.....	19
2.2.2 Traitements microscopiques.....	20
2.2.2.1 Un système de comptage de piétons dans les couloirs du métro parisien	20
Présentation générale du système	20
Algorithme	21
La compression	22
La détection et la segmentation	22
Le suivi des zones et le décompte des jetons.....	22
Conclusion.....	23
2.2.2.2 Les travaux de A. Rourke et M.G. Bell.....	23
2.2.2.3 Les travaux de A. Mecocci.....	24
Détection	25
Validation et évaluation de la direction.....	25
2.3 Les applications basées sur des capteurs linéaires.....	26
2.3.1 Applications dans le domaine des transports.....	27
2.3.1.1 Détection d'intrusions et de chutes de voyageurs sur les voies	27
2.3.1.2 Détection d'obstacles à l'avant des véhicules routiers	29
2.3.2 Applications dans le domaine industriel.....	30
2.3.2.1 Inspection automatique par vision linéaire.....	30
2.3.2.2 Numérisation de documents	31
2.3.2.3 Télédétection	31
2.3.2.4 Métrologie	31
2.3.2.5 Spectroscopie	31
2.4 Conclusion.....	32
<i>Chapitre III</i>	35
LE DISPOSITIF EXPERIMENTAL	35
3.1 Description du détecteur de base.....	35
3.1.1 Principe	35
3.1.2 Qualité de la réponse	36
3.1.2.1 Focalisation	36
3.1.2.2 Niveau des minima.....	37
3.1.2.3 Limites du détecteur	37
3.1.3 Représentation spatio-temporelle de l'intrusion.....	38
3.1.3.1 Principe	38

3.1.3.2 Lien entre l'objet réel détecté et sa représentation spatio-temporelle	40
Projection inverse d'un pixel de la barrette CCD de la caméra	40
Description de l'objet.....	41
3.2 La camera active.....	42
3.2.1 Inconvénients du système de base.....	42
3.2.2 Description du dispositif à caméra active.....	42
3.2.3 Propriétés liées aux matériaux rétro-réfléchissants	44
3.2.4 Utilisation de diodes électroluminescentes infrarouges pulsées.....	46
3.2.4.1 Utilisation de sources infrarouge.....	46
3.2.4.2 Utilisation de lumière pulsée	47
3.3 Utilisation d'un double plan de surveillance	47
3.3.1 Principe	47
3.3.2 Méthodologie du calcul de la vitesse	48
3.4 Dispositif de comptage des piétons	51
3.4.1 Description de l'installation	51
3.4.1.1 Dispositif sur site.....	51
3.4.1.2 Dispositif complet	51
3.4.2 Représentation spatio-temporelle du mouvement	53
3.4.2.1 Allure des images obtenues.....	53
3.4.2.2 Influence de la vitesse des piétons sur les images obtenues	55
3.5 Le calcul de la vitesse.....	56
3.5.1 Calcul de la vitesse des piétons au niveau des pixels affectés par le mouvement.....	56
3.5.2 Vitesses obtenues dans le cas d'un piéton.....	58
3.6 Conclusion.....	59
<i>Chapitre IV</i>	65
LE PRETRAITEMENT DES DONNEES	65
4.1 Le filtrage des vitesses	65
4.1.1 Représentation du signal vitesse.....	66
4.1.2 Choix du type de filtrage.....	67
4.1.3 Filtres morphologiques.....	67
4.1.3.1 Résultats obtenus avec le filtre morphologique.....	68
4.1.4 Filtre conditionnel	70
4.1.4.1 Calcul des minima conditionnels.....	70
4.1.4.2 Calcul des maxima conditionnels.....	71
4.1.4.3 Résultats obtenus avec le filtre conditionnel	71
4.2 Classification finale des vitesses	72
4.3 Sélection de l'intersection des formes	74
4.4 Séquences d'images utilisées pour le test des algorithmes.....	76
4.4.1 Séquence No 1.....	76
4.4.2 Séquence No 2.....	77
4.4.3 Séquence No 3.....	78
4.4.4 Séquence No 4.....	79
4.4.5 Séquence No 5.....	80
4.4.6 Séquence No 6.....	81
4.5 Conclusion.....	82
<i>Chapitre V</i>	85
LES ALGORITHMES DE COMPTAGE	85
5.1 Introduction.....	85
5.1.1 Choix de la séparation des formes.....	87
5.2 Procédure de séparation des formes	88
5.2.1 Procédure de séparation par la vitesse.....	88
5.3 Principes de base de la morphologie mathématique.....	91
5.3.1 Transformations en tout ou rien par un élément structurant.....	91
5.3.2 Transformations morphologiques élémentaires	92
5.3.2.1 Transformation par érosion	92
5.3.2.2 Transformation par dilatation.....	93
5.3.2.3 Ouverture et fermeture morphologiques	94
Ouverture.....	94

Fermeture	95
5.4 Traitements inspirés de la morphologie mathématique avec un élément structurant déformable dynamique	95
5.4.1 Elément structurant rectangulaire dynamique	96
5.4.1.1 Utilisation de la vitesse	97
5.4.1.2 Largeur de l'élément structurant	97
5.4.1.3 Longueur de l'élément structurant	98
5.4.1.4 Définition de la largeur de l'élément structurant	98
Largeur d'élément structurant trop réduite	99
Largeur d'élément structurant trop élevée	99
Largeur d'élément structurant conseillée	99
5.4.1.5 Résultats	100
Séquence 1	101
Séquence 2	102
Séquence 3	103
Séquence 4	104
Séquence 5	105
Séquence 6	106
5.4.2 Elément structurant ellipsoïdal dynamique	107
Utilisation de la vitesse	107
5.4.2.1 Largeur et longueur de l'élément structurant	107
Largeur de l'élément structurant	107
Longueur de l'élément structurant	108
5.4.2.2 Application de l'élément structurant ellipsoïdal	108
5.4.2.4 Résultats obtenus	110
Séquence 1	110
Séquence 2	111
Séquence 3	112
Séquence 4	113
Séquence 5	114
Séquence 6	115
5.4.3 Elément structurant directionnel dynamique	116
5.4.3.1 Classification des trajectoires des piétons	116
5.4.3.2 La série d'éléments structurants utilisée	117
Application de l'élément structurant le plus adapté aux formes	117
Largeur de l'élément structurant	117
Longueur de l'élément structurant	118
Direction de l'élément structurant	118
5.4.3.3 Résultats obtenus	120
Séquence 1	120
Séquence 2	121
Séquence 3	122
Séquence 4	123
Séquence 5	124
Séquence 6	125
5.5 Suivi des formes	126
5.5.1 La procédure de suivi des formes	126
Le suivi des formes	128
5.5.1.1 Scénario 1	128
5.5.1.2 Scénario 2	129
5.5.2 La procédure de filtrage des formes	131
5.6 Conclusion	131
Chapitre VI	135
L'ALGORITHME ASSESA	135
6.1 Introduction	135
6.2 première étape	136
6.2.1 Principe	136
6.2.2 Application aux images	139
6.3 deuxième étape	140

6.3.1 Principe	140
6.3.2 Applications aux images	142
6.4 Effet progressif de l'algorithme	143
6.5 Choix de la taille des rectangles $[N1, M1]$ et $[N2, M2]$	143
6.6 Intérêt du seuillage adaptatif	144
6.7 Resultats	146
Séquence 1.....	146
Séquence 2.....	147
Séquence 3.....	148
Séquence 5.....	150
Séquence 6.....	151
6.8 Conclusion.....	152
<i>Chapitre VII</i>	155
LES RESULTATS DU COMPTAGE	155
7.1 contexte	155
7.1.1 L'environnement de l'expérimentation	155
7.1.2 Le matériel utilisé.....	156
7.2 Performances des algorithmes inspirés de la morphologie mathématique.....	157
7.2.1 Les paramètres généraux.....	157
7.2.2 Elément structurant rectangulaire dynamique.....	159
7.2.2.1 Les résultats globaux.....	159
7.2.2.2 L'analyse des erreurs	160
7.2.3 Elément structurant ellipsoïdal dynamique.....	161
7.2.3.1 Les résultats globaux.....	161
7.2.3.2 L'analyse des erreurs	162
7.2.4 Elément structurant directionnel dynamique.....	162
7.2.4.1 Les résultats globaux.....	162
7.2.4.2 L'analyse des erreurs	163
7.2.5 Aspects temps réel.....	164
7.3 Resultats de l'algorithme ASSESA.....	164
7.3.1 Les paramètres utilisés	164
7.3.2 Les résultats globaux	165
7.3.3 L'analyse des erreurs	166
7.3.4 Aspects temps réel.....	167
7.4 Comparaison des deux types d'algorithmes.....	167
7.5 Conclusion.....	168
<i>Chapitre VIII</i>	171
CONCLUSION ET PERSPECTIVES	171
<i>Références bibliographiques</i>	177

RESUME

Ce travail de thèse, réalisé à l'Institut National de Recherches sur les Transports et leur Sécurité (INRETS), s'intègre dans un programme de recherche et de développement d'outils favorisant l'automatisation des moyens de transport. Cette recherche a consisté à développer un système permettant d'effectuer un comptage en temps réel de piétons à l'aide de vision artificielle linéaire. Le système complet comprend une partie matérielle (dispositif) et une autre logicielle qui contient les algorithmes de comptage.

Le **dispositif** sur site comprend :

- deux caméras linéaires actives, en général fixées au plafond et déterminant deux plans de surveillance perpendiculaires à la direction de déplacement des piétons,
- deux bandes rétro-réfléchissantes collées au sol, dont la longueur correspond à la largeur du site de mesures (couloir, entrée, etc...).

La paire de caméras est connectée par une liaison filaire à un micro-ordinateur pour l'acquisition et le traitement des données.

Les données disponibles pour procéder au comptage sont :

- l'information spatiale qui correspond à l'axe défini par l'orientation des barrettes CCD,
- l'information temporelle liée à la fréquence d'acquisition des lignes,
- l'information vitesse obtenue grâce aux deux plans de surveillance.

Quatre **algorithmes** de comptage en temps réel ont été développés :

- les trois premiers s'inspirent de méthodes basées sur la morphologie mathématique qui a été élargie et enrichie par l'utilisation d'éléments structurants déformables dynamiques,
- le quatrième algorithme est une nouvelle méthode de segmentation utilisant des sous-échantillonnages des images et des seuillages successifs.

Plusieurs évaluations du système sur site réel ont été effectuées. Les résultats, en termes de comptage, montrent qu'une précision de l'ordre du pourcent est obtenue.

SUMMARY

This thesis work was done at INRETS, French National Institute on Transport and Traffic safety and was defined as an exploitation help in public transport.

The goal of the work carried out was to develop an automatic and real-time pedestrians counting system in public transport such as subways. For this counting system, a new technology was developed based on two linear cameras defining two surveillance planes.

The thesis report contains two main parts : the technology of the system and the algorithms developed for counting.

Technology proposed

In field conditions, under daylight or artificial light, and due to the reflecting characteristics of any pedestrian crossing the surveillance planes, the output signal is not as simple as the theoretical one. The efficiency of the detection is improved by :

- using RetroReflecting Material (RRM) and adjusting the illumination level according to the reflecting material characteristics,
- using infrared diodes and rejecting visible light with appropriate filters on the lens.

An "active camera" has been designed, with Light Emitting Diodes set close to the axis of the camera lens, so that the light sent to the retroreflecting stripe is reflected back towards the lens without dispersion.

Counting algorithms

Four algorithms were developed to count the pedestrians crossing the surveillance planes :

- Three of them are inspired from mathematical morphology methods,
- the fourth is a new segmentation procedure.

The proposed algorithms have been implemented and applied to several sequences of line images. These sequences were collected in real sites. The counting system accuracy is about 1%.

Chapitre I

*Le comptage automatique des piétons
par vision artificielle : un outil d'aide à
l'exploitation dans les transports publics.*

Chapitre I

LE COMPTAGE AUTOMATIQUE DES PIETONS PAR VISION ARTIFICIELLE : UN OUTIL D'AIDE A L'EXPLOITATION DANS LES TRANSPORTS PUBLICS

Le travail présenté dans ce mémoire de thèse s'intègre dans un programme de recherche et de développement d'outils favorisant l'automatisation des moyens de transport. L'automatisation d'un processus quelconque consiste essentiellement à remplacer l'homme par un système susceptible d'agir à sa place.

L'objet de cette thèse est de proposer un système automatique de comptage de piétons dans les transports publics à l'aide d'un système de vision artificielle linéaire.

Dans toutes les grandes villes à travers le monde, de grandes masses de personnes se déplacent quotidiennement pour des trajets domicile-travail, pour leurs loisirs ou pour leurs

achats. En période de pointe, le déplacement de toutes ces personnes conduit à d'inévitables situations de congestion et d'inconfort, que ce soit sur la route ou à l'intérieur de réseaux de transport public. Pendant les périodes creuses de trafic, principalement le soir et pendant la nuit, les rames et les couloirs sont désertés, créant ainsi un sentiment de peur et d'insécurité chez les voyageurs et favorisant la criminalité et le vandalisme.

Depuis quelques années, toutes les politiques, au niveau régional comme au niveau national, consistent à essayer d'amener un maximum de personnes, utilisant leur voiture au quotidien, à emprunter les réseaux de transports publics. Cela allègerait les congestions permanentes sur les routes, réduirait significativement les taux de pollution urbaine ainsi que le nombre de victimes sur ces routes.

Un certain nombre de facteurs sont susceptibles d'amener ce changement en accroissant l'attractivité des transports publics. Il s'agit de facteurs améliorant la qualité de service, en termes de temps passé, de confort et de sécurité :

- information suffisante pour l'orientation des passagers,
- réduction des temps d'attente,
- réduction de la durée des voyages,
- propreté des lieux,
- adaptation de l'offre à la demande pour éviter des situations de congestion.

Tous ces facteurs peuvent être améliorés par l'utilisation d'outils télématiques faisant usage de l'informatique, de la vision artificielle ou des moyens de transmission modernes.

Les projets sont nombreux dans ces domaines, notamment dans le cadre du PREDIT et des projets européens du 4^e Programme Cadre de Recherche et Développement. Le projet CROMATICA, qui exploite le traitement d'images et les transmissions pour améliorer la sécurité dans les transports publics, en est un exemple [DEP 96/1].

En ce qui concerne l'adaptation de l'offre à la demande, pour laquelle la conduite automatique des trains peut être un élément très positif, l'exploitant doit avoir une connaissance aussi précise que possible sur la population qui emprunte son système de transport [DEP 96/2]. Autrement dit, il est nécessaire de mesurer en permanence les flux d'usagers en différents points du réseau de transport afin de connaître les fréquentations des différents trajets empruntés par les usagers. La réalisation d'un dispositif de comptage directionnel aux points d'accès constitue donc un objectif important en matière de gestion des déplacements.

Pour l'exploitant lui-même, la possession d'un système de comptage automatique des flux de voyageurs, répond à un certain nombre de besoins :

- par exemple à Paris, pour le métro et le RER : partage des recettes avec la SNCF ou/et entre centres de résultats de la RATP,
- amélioration de la qualité de service : mesure de fréquentation pour projeter la réalisation d'accès supplémentaires et l'installation d'escaliers mécaniques,
- meilleur diagnostic sur les caractéristiques de la fraude : l'efficacité des moyens de lutte dépend de l'évaluation de la fraude,
- pour le bus et le tramway : partage des recettes et amélioration de l'offre grâce à une mesure du flux de passagers empruntant les voitures,
- pour tous les modes de transport : un support d'études commerciales ou de prévisions de trafic.

1.1 LE CAHIER DES CHARGES DU SYSTEME

Du point de vue des exploitants de réseaux de transports publics, un système de comptage automatique doit, pour être acceptable, répondre à un certain nombre de critères :

- ne pas entraver les déplacements des voyageurs,
- s'adapter à une très grande diversité, tant des voyageurs eux-mêmes que des conditions de comptage. En effet, si compter des objets est relativement facile, compter les voyageurs l'est beaucoup moins, compte tenu de la multiplicité des formes, des vitesses de déplacement, surtout dans une diversité d'infrastructures et d'environnements (champs électriques ou magnétiques, chaleur),
- s'adapter à la taille des sites de mesures : en effet, la taille des stations pouvant être variable, des aménagements éventuels du système devront être possibles sans remettre en question le principe de fonctionnement,
- s'intégrer facilement dans les infrastructures existantes : il n'est pas concevable de modifier l'architecture d'une station pour installer le système de comptage. Celui-ci devra nécessiter le moins possible de travaux de génie civil,
- être mis en place rapidement : l'exploitation du réseau ne doit pas être perturbée pendant un intervalle de temps trop important. L'idéal serait de pouvoir réaliser les travaux d'installation pendant les interruptions normales du trafic.

D'autre part, quelque soit le type de réseau, le système devra être insensible aux conditions d'éclairage des sites de mesure. Dans ce cas, on peut classer les sites en deux catégories :

- sites souterrains,
- sites à ciel ouvert.

Pour les sites souterrains, seul l'allumage ou l'extinction d'une ou plusieurs sources de lumière artificielle provoque des variations lumineuses. Par contre, pour les sites extérieurs, les passages de nuages provoquent des variations soudaines de lumière et le cycle jour/nuit provoque une variation plus lente de luminosité. De plus, pour ces sites, les conditions atmosphériques (pluie, neige, vent, brouillard....) peuvent perturber le fonctionnement du système. Celui-ci devra être insensible à tous ces facteurs.

Enfin, en ce qui concerne les performances du système, celui-ci devra :

- compter avec la plus grande précision possible : une précision de $\pm 3\%$, avec un taux de confiance de l'ordre de 95%, est un objectif souhaité,
- fournir cette qualité de comptage de façon fiable et durable,
- permettre diverses modalités de recueil de données. Pour les besoins ponctuels, un mode d'exploitation manuel doit être envisagé avec du matériel portable, alors que pour les besoins permanents, il faut concevoir un mode entièrement automatique avec une transmission, puis un traitement sur site dédié,
- offrir le meilleur rapport qualité/prix.

1.2 LES SYSTEMES DE COMPTAGE AUTOMATIQUE EXISTANTS

1.2.1 Détection par contact

Par définition, la détection nécessite un contact entre le capteur et le corps à détecter [SAU 94]. Les marches sensibles des autobus sont basées sur ce principe. Des systèmes plus sophistiqués existent, comme les tapis "intelligents" (Figure 1.1).

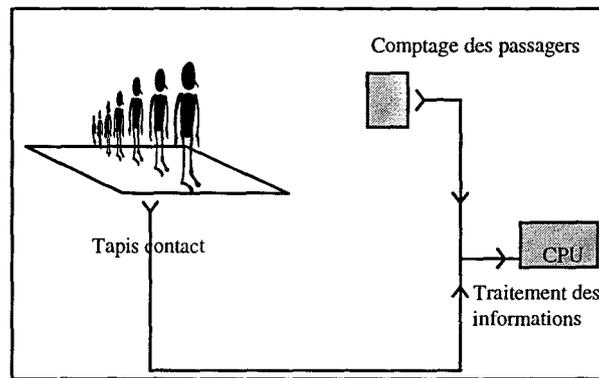


Figure 1.1 : détection par tapis

Les systèmes usuels ont des performances très moyennes. Ils nécessitent une maintenance rigoureuse sous peine de voir leurs performances se dégrader et ne peuvent assurer correctement leur fonction lorsqu'il n'y a plus de marche sur le lieu de passage. Ils ne permettent pas de différencier le sens des passages dans le cas des bus en libre-service. Par contre, le coût du système traditionnel est d'autant plus intéressant qu'il assure, aussi, souvent des fonctions de sécurité.

1.2.2 Détecteurs actifs de proximité

Ces types de capteurs "tout ou rien" sont les plus répandus. Ils sont généralement basés sur deux principes de détection : les systèmes de type "barrière" fonctionnent par coupure du faisceau réfléchi par un réflecteur optique. Ceux du type "réflex" procèdent par détection directe, l'objet faisant office de réflecteur (Figure 1.2). Deux grandes familles de capteurs permettent de mettre en œuvre cette dernière technique. Il s'agit de capteurs qui travaillent dans le domaine de l'infrarouge ou des ultrasons.

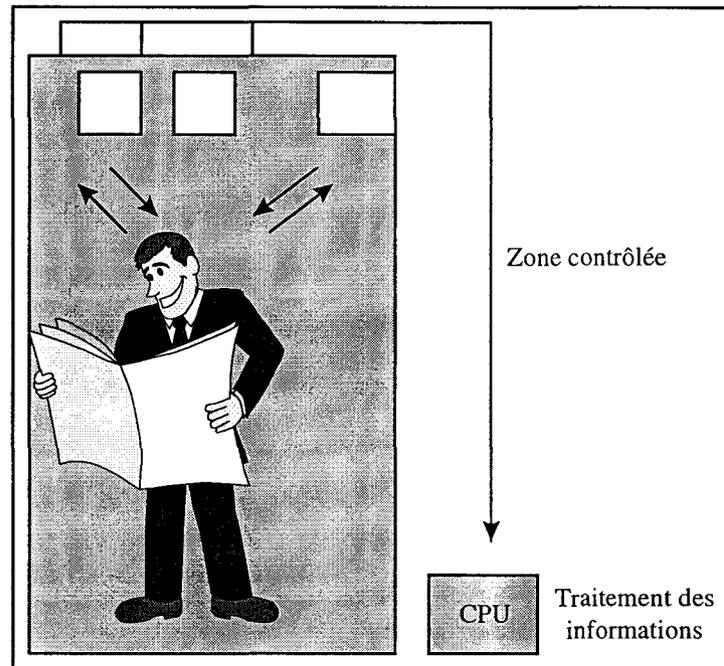


Figure 1.2 : détecteur actif de proximité

L'infrarouge est sensible à la lumière ambiante et permet, si l'on utilise une double rangée de sources lumineuses de ce type, de déterminer le sens de passage. Cependant, le réglage est laborieux et le capteur est très influencé par la couleur et la matière des vêtements des voyageurs. Les précisions annoncées par les constructeurs sont souvent remises en cause lors des tests expérimentaux.

Les ultrasons permettent de s'affranchir des problèmes de couleur, mais se heurtent à des difficultés de réglage, à des problèmes liés à la présence de vibrations ou de bruits dans l'environnement et à la non détection de certains textiles.

1.2.3 Détecteurs passifs de proximité

Les détecteurs passifs de proximité, qui fonctionnent dans une plage de rayonnement dans le spectre de l'infrarouge, font appel à la pyrométrie infrarouge. Cette technique permet de réaliser des systèmes à encombrement réduit, faciles à installer, pouvant différencier les sens de passage et dont le coût est relativement intéressant.

Le principe de ces systèmes est très proche de ceux basés sur la détection active. Si la faible hauteur est un atout pour les comptages dans les couloirs bas de plafond et pour le comptage dans les autobus, elle peut en revanche être un handicap esthétique dans des couloirs très hauts.

1.2.4 Vision et analyse de l'image

L'association des images et l'étude des algorithmes a permis le développement rapide de systèmes à base de vision.

Dans cette technique, on utilise des caméras linéaires ou matricielles. Le positionnement des caméras et les principes d'analyse des images conditionnent souvent les performances de ces systèmes. Ces derniers, d'abord utilisés pour des tâches de surveillance, sont devenus, au fil des années, des outils très performants de diagnostic grâce aux techniques de traitement de l'image qui les accompagnent. Nous reviendrons très largement sur ce type de systèmes, opérant par vision artificielle, tout au long de ce mémoire.

1.3 ORGANISATION DE LA THESE

Aucun des dispositifs de comptage actuellement disponibles ne satisfait complètement le cahier des charges défini au § 1.1.

Les systèmes de vision artificielle offrent, *à priori*, une possibilité réelle de détection et de comptage des piétons en mouvement. En effet, les caméras peuvent être utilisées pour obtenir des informations sur la taille des piétons, sur leur position dans la zone contrôlée, sur leur sens et vitesse de déplacement.

C'est dans ce sens que nous présentons, au chapitre **II** de ce mémoire, une synthèse sur les techniques de détection par vision artificielle, ainsi que les outils de traitements d'images qui les accompagnent.

Au chapitre **III**, nous exposons le principe de détection visuelle basé sur l'utilisation d'une caméra linéaire comme capteur actif grâce à des adaptations technologiques. La combinaison de deux caméras de ce type permet d'obtenir des informations très précises sur la taille et la position des piétons traversant le champ visuel des caméras. On débouche ainsi sur le concept de double plan de surveillance.

Le chapitre **IV** est dédié à un certain nombre de prétraitements portant sur les données brutes afin d'éliminer le bruit entâchant l'information recueillie avant l'application des algorithmes de comptage des piétons. Une méthode originale de filtrage des vitesses de déplacement est développée dans ce chapitre.

Les chapitres **V** et **VI** constituent le corps principal de ce mémoire de thèse, puisqu'ils sont dédiés au développement des algorithmes de comptage. Le chapitre **V** présente 3

algorithmes de traitement des données, s'inspirant tous les trois de méthodes de morphologie mathématique qui ont été élargies et enrichies. Le chapitre **VI** présente un algorithme supplémentaire de comptage, baptisé "ASSESA" pour Algorithme de Segmentation par Sous-Echantillonnage et Seuillage Adaptatif, complètement différent des trois précédents dans la façon de traiter les images.

Au chapitre **VII**, sont présentés des résultats expérimentaux qui permettent de comparer les différents algorithmes proposés en termes de performance et fiabilité.

Les limites du capteur sont exposées en conclusion et des améliorations sont proposées pour y remédier, tant sur le plan algorithmique que sur le plan purement technologique.

Chapitre II

Les systèmes de comptage existants

Chapitre II

LES SYSTEMES DE COMPTAGE EXISTANTS

2.1 INTRODUCTION

Le sens habituel du mot image sous-entend le concept d'image "visuelle" ou encore "optique" [OKS 89]. La signification technique de ce mot est en fait beaucoup plus générale : dans le domaine du traitement des images, on rencontre aussi bien des "images visuelles" (par exemple dans les applications de robotique ou encore dans l'exploitation des photographies aériennes ou spatiales), que des images représentant en fait d'autres grandeurs que des grandeurs photométriques classiques.

On peut ainsi appeler image toute fonction $Q(x,y)$ de deux variables quelconques, x et y correspondant à des coordonnées spatiales. Pour représenter une telle fonction, on utilise, par exemple :

- une surface,
- des coordonnées spatiales x et y discrétisées $(x_1, x_2, \dots, x_i, \dots, x_p)$ et $(y_1, y_2, \dots, y_i, \dots, y_n)$,
- une image visuelle, dans laquelle la valeur de Q en un point est codée par un niveau photométrique, par exemple un niveau de gris, qui s'étend du noir pour une valeur minimale au blanc pour une valeur maximale.

Souvent, une autre dimension intervient : il s'agit du temps. En effet, les phénomènes représentés, à un instant donné, par une image, peuvent évoluer dans le temps. On peut alors

les représenter par une image dépendant elle-même du temps ou, ce qui revient au même, par une fonction des trois variables x , y , t .

Chaque fois que l'on voudra représenter une image par un tableau de valeurs, il faudra obligatoirement utiliser des coordonnées discrètes. Une image sera alors un ensemble de valeurs :

$$Q_{x_i, y_j} = f(x_i, y_j)$$

Le triplet (Q_{x_i, y_j}, x_i, y_j) est un élément de l'ensemble formé par une image. Il porte le nom de pixel. Sur une image visuelle, il peut être considéré comme un point lumineux de cette image. Très souvent, les points (x_i, y_j) dont la disposition définit le maillage de l'image, sont disposés régulièrement dans un repère cartésien. Si n est le nombre de lignes de l'image et p le nombre de colonnes, on parle alors d'image de dimension $n \times p$ (exemple image "256 x 256"). Ceci est le principe de représentation des images saisies avec des **caméras matricielles**.

La zone observée par une caméra matricielle est représentée par une image à deux dimensions. Des séquences d'images peuvent être recueillies à une cadence donnée dans le temps (Figure 2.1).

La cadence d'acquisition des caméras vidéos, par exemple, est de 25 images par seconde.

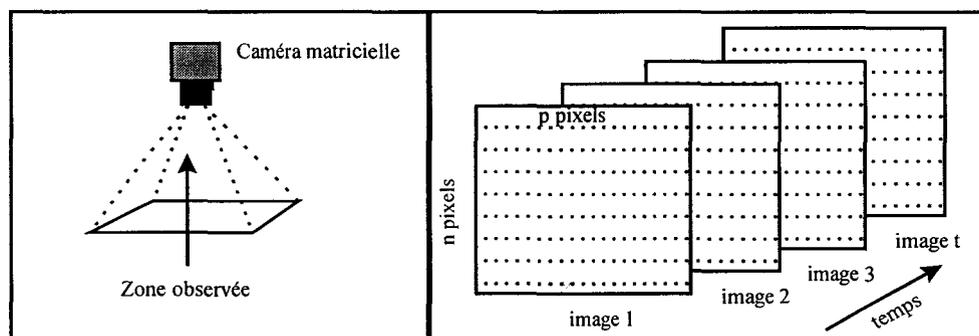


Figure 2.1 : séquence d'images 2D saisies par une caméra matricielle

Le principe des **caméras linéaires** est différent de celui des matricielles dans la mesure où l'image acquise par ce type de caméras est unidimensionnelle. Le capteur linéaire est une version simplifiée des capteurs CCD (Charge Coupled Devices) matriciels utilisés dans les caméras vidéo. Il comporte essentiellement une rangée de photocapteurs ou photosites. La zone observée par une caméra linéaire est représentée par une image unidimensionnelle ou image ligne (Figure 2.2). Comme pour les caméras matricielles, la cadence d'acquisition des

images lignes peut être variable. Elle peut atteindre plusieurs centaines d'images lignes par seconde.

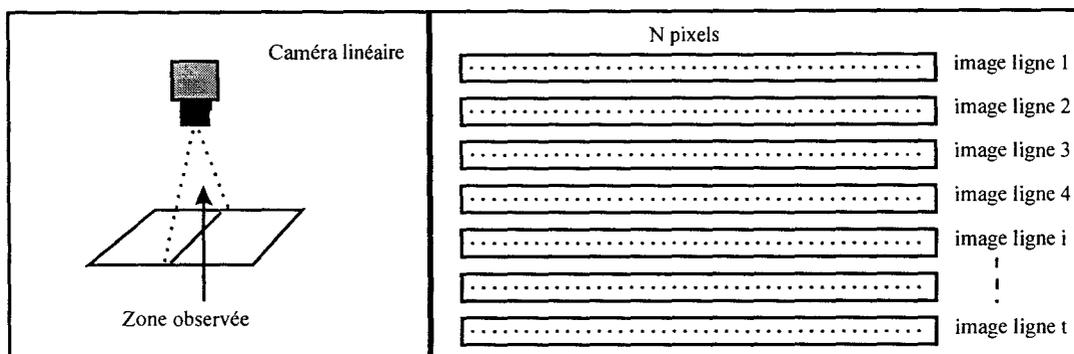


Figure 2.2 : séquence d'images lignes saisies par une caméra linéaire

La principale différence entre ces deux familles de caméras réside donc dans la dimension des images :

- image unidimensionnelle pour une caméra linéaire,
- image bidimensionnelle pour une caméra matricielle.

Cependant, il est possible avec une caméra linéaire de reconstituer une image bidimensionnelle en empilant les images lignes acquises. A l'inverse avec une caméra matricielle, il est possible de n'extraire et de ne traiter qu'une seule ligne de l'image. Tout dépend de l'objectif recherché.

Les systèmes de comptage d'objets ou de personnes à base de vision artificielle, qui peuvent satisfaire le cahier des charges présenté au chapitre I, peuvent utiliser les deux familles de caméras.

On peut voir dans la littérature qu'un certain nombre de systèmes de comptage à base de caméras matricielles ont été développés ces dernières années. La première partie de ce chapitre s'attachera à présenter ceux qui se rapprochent le plus de notre problématique. Pour chacun d'entre eux, nous situerons leur contexte d'utilisation, la méthodologie de traitement des données, leurs performances ainsi que leurs avantages et inconvénients.

Actuellement, à notre connaissance, il n'existe pas de système de comptage de personnes ou d'objets à base de caméras linéaires. Par contre, les applications utilisant ce type de capteur sont très nombreuses. La deuxième partie de ce chapitre en présentera quelques unes.

2.2 LES SYSTEMES DE COMPTAGE A BASE DE CAMERAS MATRICIELLES

On distingue pour les systèmes de comptage à base de caméras matricielles deux types d'analyse :

- la première a pour vocation de réaliser une analyse globale ou macroscopique des scènes où des événements se produisent. Il est possible, par exemple, de définir l'évolution spatio-temporelle de personnes ou d'objets dans une scène surveillée. Ces systèmes, qui sont en réalité des outils de surveillance, sont installés de façon à contrôler une zone importante. On les installe dans des halls, sur des quais de transport ferroviaire ou encore dans des zones stratégiques. Pour cette analyse, la composante comptage est moins importante que pour les systèmes qui cherchent à individualiser les personnes ou les objets. Il s'agit, le plus souvent, d'estimer uniquement une densité globale qui donne une idée du taux d'occupation, exprimé en nombre de personnes ou d'objets présents dans la zone observée,
- la deuxième analyse s'intéresse à des événements précis qui se produisent dans la scène observée. C'est un traitement microscopique car l'analyse effectuée est plus fine que celle qui consiste à analyser la globalité d'une scène. On les installe à des points d'accès à un site, comme une entrée ou une sortie de couloir, une entrée d'un grand magasin. Il s'agit, dans ce cas, d'individualiser les événements qui se produisent : par exemple dénombrer aussi exactement que possible les personnes qui entrent ou sortent d'un magasin.

2.2.1 Traitements macroscopiques

Les systèmes mis au point pour les traitements macroscopiques sont dédiés à la surveillance de zones particulières où ils cherchent à évaluer la densité et le flux de personnes ou d'objets. Il s'agit de caractériser automatiquement des mouvements par des techniques de traitement d'images et de relever des indices sur l'évolution spatiale et temporelle de la foule. On peut aussi détecter des regroupements anormaux et soudains de personnes à un endroit précis ou des changements anormalement rapides de la vitesse d'évolution d'une foule.

2.2.1.1 Estimation de la densité et de la vitesse d'une foule par S. Velastin

Dans ses travaux, S.Velastin [VEL 93] montre qu'il est possible de faire une estimation de la densité et du mouvement d'une foule, sans connaissances à priori sur les personnes en mouvement, ni sur leur taille ou leur vitesse.

Calcul de la densité de foule

La densité de foule est calculée selon deux méthodes différentes.

La première méthode est basée sur une détection des contours des personnes présentes dans la scène. Pour évaluer cette méthode, l'auteur utilise un certain nombre d'images pour lesquelles il effectue les traitements suivants :

- pour chaque image, un comptage manuel du nombre de piétons est effectué,
- pour chaque image, une détection automatique des contours est effectuée avec un comptage des pixels formant ces contours.

Par la suite, l'auteur trace la courbe qui représente le nombre de piétons comptés manuellement dans toutes les images traitées en fonction du nombre de pixels des contours détectés. Cette courbe mise en place est approchée par une droite.

Pour chaque nouvelle image, on associe le nombre de piétons au nombre de pixels des contours détectés grâce au modèle précédent. Les résultats, en termes de comptage, donnent 23% d'erreur pour des images contenant en moyenne 15 piétons et prises à intervalles de temps réguliers de 10 secondes.

Dans la seconde méthode, l'auteur utilise une image de référence. Il s'agit de l'image du fond, obtenue lorsqu'aucun piéton ou objet n'est présent dans la scène. La détection se fait par soustraction pixel à pixel et en niveaux de gris de l'image courante et de l'image de référence. Par la suite, l'auteur trace la courbe qui représente le nombre de piétons comptés manuellement en fonction du nombre de pixels de l'image de référence dont le niveau de gris est supérieur à un seuil. Cette courbe est aussi approchée par une droite. Cette méthode permet d'abaisser l'erreur de comptage à 6%, résultat obtenu par moyennage des erreurs sur une dizaine d'images.

S. Velastin montre que les performances du système dépendent :

- de la position de la caméra,

- des effets de perspective,
- de la répartition des piétons dans la scène.

Estimation du mouvement de la foule

Les estimations de mouvement dans la scène vont permettre de déduire la direction et la densité du flux de piétons. Ces estimations peuvent se faire au niveau des pixels, ou au niveau d'un voisinage de chaque pixel :

- Calcul du flot optique en utilisant l'algorithme de Horn et Schunck [HOR 81]. Cette méthode est utilisée par l'auteur pour évaluer le mouvement. De manière à réduire le temps de traitement, les calculs sont effectués sur des pixels d'intérêt. Ces pixels, d'un niveau de gris supérieur à un seuil, sont obtenus en faisant la différence entre l'image de fond sans piéton et l'image courante.
- Corrélation des niveaux de gris sur des voisinages de pixels d'intérêt de deux images successives. Les voisinages des pixels sont appelés fenêtres. La méthode proposée établit la correspondance entre deux fenêtres par la corrélation des niveaux de gris entre deux images successives. Ceci permet d'associer un vecteur mouvement à cette fenêtre. S. Velastin en déduit, ensuite, une direction et une amplitude du mouvement global de la foule.

L'auteur spécifie que, quels que soient les résultats obtenus, il est très difficile de les comparer à des observations manuelles. Cependant, les caractéristiques du mouvement d'une foule se déplaçant globalement dans un sens sont correctement déterminées.

2.2.1.2 Les travaux de A. Rourke et M.G. Bell

A.Rourke et M.G. Bell [ROU 92] proposent, dans le cas des foules, de mettre en relation la densité des piétons présents dans une scène avec la quantité de mouvement.

Pour caractériser la quantité de mouvement pendant une période de surveillance fixée, les opérations suivantes sont menées :

- des différences d'images successives isolent les zones en mouvement,
- pour chaque pixel dans les images différences, les auteurs comptent le nombre de fois où ce pixel a appartenu à une zone en mouvement,

La densité de piétons est ensuite déterminée comme étant la somme pondérée des pixels ayant appartenu à une zone en mouvement. Dans cette somme, le comptage de chaque pixel est pondéré par le nombre de fois où ce pixel a appartenu à une zone en mouvement.

2.2.1.3 Prise en compte des piétons dans la régulation du trafic routier

Les piétons ont souvent, dans leurs déplacements en ville, le sentiment d'être les laissés pour compte de la gestion des feux de circulation.

Si ceux-ci leur assurent une certaine sécurité de traversée aux carrefours, c'est en effet souvent au prix de temps d'attente longs, d'autant plus insupportables que les temps de traversée qui leur sont alloués dans l'exploitation du carrefour n'ont aucun lien avec le nombre de piétons engagés dans cette traversée.

En effet, une donnée importante et pourtant généralement ignorée dans les stratégies de régulation des feux tricolores est le nombre de piétons empruntant les passages cloûtés aux carrefours [COH 90]. Le séquençement ne tient compte que du débit de véhicules mesurés sur les tronçons du carrefour [KHO 88].

C'est la raison pour laquelle une équipe de l'INRETS, en collaboration avec le Centre d'Automatique de Lille (CAL), a décidé d'entreprendre une recherche qui consiste à procéder à une estimation du nombre de piétons traversant un passage pour piétons [VAN 96].

Une caméra matricielle fixe et visant la zone de traversée permet de relever des données pour quantifier la densité de piétons empruntant le passage, grâce à un algorithme original d'analyse du mouvement.

Principe de l'algorithme

La réponse d'un filtre différentiel appliqué sur les images en entrée donne une mesure de la probabilité pour un pixel d'appartenir à un contour. En comparant les contours d'une image de référence périodiquement réactualisée avec les contours des images successives, il est possible de déterminer les zones affectées par le mouvement. Ces zones correspondent à des déplacements de piétons.

Le fait de travailler avec les contours d'une image de fond réactualisée présente deux avantages :

- le premier est que les contours sont des caractéristiques peu sensibles aux variations lumineuses qui peuvent survenir dans un environnement extérieur,

- le deuxième est que cette approche, basée sur une analyse bas niveau, est indépendante de la forme et du nombre d'objets ou de personnes en mouvement.

2.2.2 Traitements microscopiques

Dans les traitements microscopiques, les évènements sont individualisés. Pour le cas du comptage de piétons, il s'agit de les dénombrer, de déterminer la direction de leurs déplacements et leurs vitesses respectives. En vision artificielle, un piéton en mouvement est assimilé à un objet déformable.

Dans ce cas, les systèmes de comptage consistent à suivre aussi exactement que possible la trajectoire de chaque piéton. Le problème devient vite compliqué lorsque les piétons évoluent dans une foule. En effet, nous sommes alors confrontés au problème des occlusions, les piétons se masquant les uns les autres. A cela, il faut ajouter les problèmes posés par les variations d'intensité lumineuse dans la scène.

Malgré toutes ces difficultés, un certain nombre de systèmes ont fait l'objet de recherches approfondies ces dernières années. Dans le cadre de ce mémoire, nous présentons quelques unes de ces recherches.

2.2.2.1 Un système de comptage de piétons dans les couloirs du métro parisien

Une collaboration entre la RATP et une équipe de l'INRETS a permis le développement d'un système de comptage de piétons à des endroits précis à l'intérieur du métro parisien [GLA 95].

Les exigences de la RATP pour ce système de comptage sont les suivantes : le système doit pouvoir effectuer un comptage en temps réel avec une précision de $\pm 3\%$ sur des échantillons de 200 personnes avec un taux de confiance de 95%. Outre les contraintes de coût, l'encombrement du système sur le terrain doit être minimum pour ne pas gêner l'écoulement normal des voyageurs, être discret et s'intégrer facilement aux installations existantes.

Présentation générale du système

Dans le système présenté, le piéton est considéré comme un objet déformable. Il n'est comptabilisé que s'il franchit une ligne imaginaire tracée au sol.

Le suivi de cet objet déformable s'effectue grâce à une extraction de plusieurs caractéristiques liées au mouvement du piéton. Le mouvement est détecté à l'aide d'un marquage au sol en forme de damier qui engendre une distribution de transitions noires et blanches (Figure 2.3). La présence d'un piéton en mouvement entraîne l'altération momentanée du maillage. Il est alors possible de définir le gabarit du piéton et son sens de circulation.

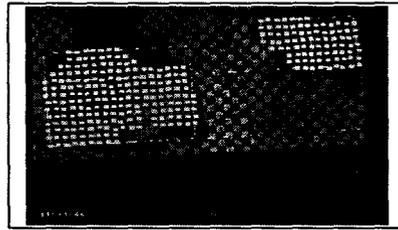


Figure 2.3 : damier de détection

En cas de dégradation du damier (usure ou petits objets posés dessus), un mécanisme auto-adaptatif, exécuté en tâche de fond, permet de revenir à une image de référence correspondant au damier non dégradé. Ce mécanisme permet aussi de palier les inconvénients liés à de légers décentrages des caméras.

Une ou plusieurs caméras matricielles sont installées en fonction de la largeur du site à surveiller, à raison d'une caméra pour 1,20 mètre de large. Ces caméras sont disposées à la verticale de la scène.

Algorithme

L'algorithme est construit autour de la notion de jeton. A chaque individu entrant dans la zone de comptage, on fait correspondre un jeton. Dès que cet individu sort de cette zone, le jeton correspondant est détruit.

Cette mise à jour des jetons ainsi créés s'effectue à chaque nouvelle image.

L'analyse de l'ensemble de ces jetons permet le comptage des personnes.

Sur chaque jeton, des informations spatiales sur la personne à laquelle il est rattaché sont déterminées grâce à des prétraitements morphologiques [SER 82].

Les différentes étapes de l'algorithme sont les suivantes :

- compression des informations,
- détection et segmentation des images,

- suivi des zones de présence et décompte des jetons.

La compression

La compression permet de diminuer la masse d'informations disponibles dans l'image pour ne garder que celles nécessaires au comptage. Elle est composée :

- d'une compression temporelle pendant laquelle une seule image est acquise à intervalle régulier,
- d'une compression spatiale où on remplace le niveau de gris d'un pixel par la valeur moyenne des niveaux de gris sur un voisinage centré en ce même pixel. La taille de ce voisinage correspond à celle d'un carré du maillage au sol.

La détection et la segmentation

La technique de détection utilisée est une technique coopérative. Le damier permet de réaliser une cartographie des relations locales entre les carrés blancs et noirs en l'absence d'individus en mouvement.

En présence d'un individu, ces relations sont altérées. Il est alors possible de repérer les carrés modifiés et de localiser les zones affectées par le mouvement. La détection utilise cette propriété.

La segmentation permet de regrouper les carrés modifiés par le mouvement d'un même individu. Son objectif est de séparer tous les individus présents, même dans le cas d'une foule dense. Cette segmentation utilise des outils de la morphologie mathématique. Des hypothèses sur la taille des individus sont retenues afin de faciliter cette segmentation. Chaque individu est donc décrit dans la scène par un ensemble de carrés bien distincts. A tout instant, chacun de ces ensembles est lié à un jeton. Par la suite, ces jetons peuvent être fusionnés en un seul jeton ou au contraire être divisés en plusieurs.

Le suivi des zones et le décompte des jetons

Sur chaque jeton est inscrit un certain nombre de caractéristiques de l'objet auquel il est assigné et qu'il faut retrouver dans les images suivantes afin de déterminer sa trajectoire.

Une prédiction du mouvement est effectuée en tenant compte, de manière statistique, du déplacement moyen (δx , δy) de chaque piéton :

$$\begin{cases} x_k = x_{k-1} + \delta x \\ y_k = y_{k-1} + \delta y \end{cases}$$

où x_k et y_k représentent les coordonnées de la position du piéton à l'instant t_k .

Ce calcul permet de faire des hypothèses sur les corrélations entre les jetons et les zones affectées par des mouvements de piétons. Un graphe est utilisé afin de représenter l'ensemble des appariements. A chaque arc est associé un critère quantifiant la vraisemblance de l'appariement qui lui est associé. Une minimisation du coût de chaque chemin du graphe permet de déduire les meilleurs appariements.

Conclusion

Par rapport aux exigences de la RATP en termes de comptage, ce système ne répond pas à la précision souhaitée. Il nécessite la présence d'un damier au sol. Une influence de la dégradation du damier sur le pourcentage de bon comptage a été observée. Les résultats, en termes de comptage correct, sont très différents d'une scène à une autre :

- dans un couloir, on obtient une erreur de comptage de moins de 4% pour 96% des échantillons observés,
- dans un escalier, on obtient une erreur de comptage de moins de 10% également pour 96% des échantillons,
- dans un escalier mécanique, on obtient une erreur de comptage de moins de 5% pour 96% des échantillons.

Grâce à la compression spatio-temporelle, le temps de traitement obtenu permet une implantation de l'algorithme pour compter en temps réel, à une cadence de dix images par seconde.

2.2.2.2 Les travaux de A. Rourke et M.G. Bell

Nous avons déjà vu, dans le § 2.2.1.2, que A.Rourke et son équipe ont développé un système permettant une estimation du nombre de personnes dans une zone en fonction de la quantité de mouvements détectés [ROU 94].

Leur système, grâce à des fonctionnalités supplémentaires, permet de procéder au comptage de personnes en utilisant des fenêtres prédéfinies dans l'image.

Afin d'identifier les zones des fenêtres dans lesquelles un mouvement est détecté, on procède à une différence entre deux images successives saisies à des instants très rapprochés. Par la suite, on applique à cette image différence une procédure de détection et de suivi de contours. Dans cette image différence, des zones de mouvement sont présentes et peuvent être

associées à des parties du corps (bras, jambe, tête). L'assemblage de ces zones permet d'identifier la trace d'un ou plusieurs individus dans l'image. Cette association nécessite un modèle standard de piéton (Figure 2.4).

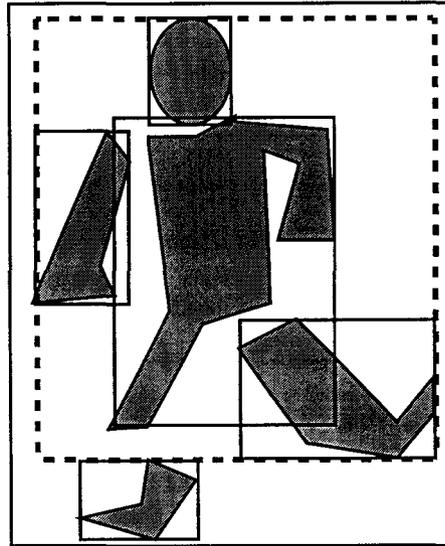


Figure 2.4 : découpage de l'espace en fenêtres

Le découpage de l'image en fenêtres permet de ne manipuler qu'une faible partie de l'information fournie par la caméra. Toutefois le manque d'information, et donc de précision, peut réduire l'efficacité de la détection et du comptage en cas de foules denses.

De plus, la diversité des individus fait qu'un modèle universel est très difficile à mettre en place. Ainsi, cet algorithme consiste à définir le rectangle exinscrit contenant chacune des zones de mouvement présentes dans l'image. Les rectangles se chevauchant sont alors fusionnés dans un rectangle de taille plus importante, correspondant à un individu selon le modèle de piéton prédéfini.

Cette procédure fournit des résultats moyens dans le cas de quelques piétons et médiocres lorsque la densité de piétons est élevée, conduisant ainsi à de multiples occlusions.

2.2.2.3 Les travaux de A. Mecocci

Dans le système de comptage mis au point par A. Meccoci [MEC 94], les problèmes de changement de luminosité sont traités de manière approfondie. Le but est de compter les voyageurs qui entrent ou qui sortent d'un bus, sachant que plusieurs personnes peuvent monter ou descendre simultanément. La caméra est placée à la verticale de chaque entrée. L'algorithme de traitement se décompose en deux étapes :

- détection des individus en mouvement,
- validation et estimation de la direction de déplacement pour différencier les passagers entrants de ceux qui sortent.

Détection

Pour ce système, les conditions expérimentales sont très difficiles compte tenu des variations spatio-temporelles de la luminosité ambiante. En effet, le bus circulant, les sources lumineuses extérieures tels que le soleil ou les lumières artificielles sont très variables et peuvent être l'objet d'occlusions par un ou plusieurs voyageurs entrant ou sortant du bus. Ainsi le système doit être le moins sensible possible à ces incessantes variations lumineuses.

Afin de ne pas analyser toute l'image fournie par la caméra qui comporte des informations inutiles pour le comptage, l'opérateur peut définir une zone d'intérêt autour de l'accès du bus.

Afin d'insensibiliser le système aux variations de luminosité brutales, la méthode utilisée est basée sur la détermination d'un contour de référence qui est construit en remarquant que les discontinuités de l'image dues aux variations lumineuses peuvent être alignées avec les lignes de l'image en positionnant la caméra de manière adéquate.

Cette référence représentant l'état de la scène à un instant donné est continuellement remise à jour. La détection d'un voyageur résulte de la modification de cette référence.

Validation et évaluation de la direction

Un voyageur ayant été détecté, une seconde fenêtre de traitement est placée autour du contour de référence modifié. Par la suite, les calculs sont réduits à cette fenêtre afin de diminuer le temps de traitement.

L'estimation du flot optique [HOR 81] est alors réalisée pendant tout le temps de détection des voyageurs :

$$\frac{\partial I(x,y,t)}{\partial x} \times u + \frac{\partial I(x,y,t)}{\partial y} \times v + \frac{\partial I(x,y,t)}{\partial t} = 0$$

où $I(x,y,t)$ est l'intensité d'un pixel (x,y) de l'image, et u et v les deux composantes de la vitesse en ce même pixel (x,y) dans le plan image.

Or la composante en y de la vitesse est plus importante que la composante en x , car le mouvement de montée ou de descente d'un voyageur est vertical dans l'image. Ainsi, en négligeant la composante en x , l'équation précédente se réduit à :

$$v = \frac{\frac{\partial I(x, y, t)}{\partial t}}{\frac{\partial I(x, y, t)}{\partial y}}$$

L'estimation de v est réalisée pendant toute la durée T de présence du ou des passagers sous la caméra. On dispose donc, à la fin de T , de la vitesse $(0, v)$, associée à chaque pixel, à chaque instant t de T . On définit alors le tableau à deux dimensions $v^{(k)}(x)$ et le vecteur \underline{v} de la manière suivante :

$$v^{(k)}(x) = \frac{1}{N} \sum_{y=1}^N v(x, y, k) \quad \text{avec } (M, N) \text{ la fenêtre de calcul, } x \in [1, M] \text{ et } k \in [1, T].$$

$$\underline{v} = [\mu^{(T)}(1), \dots, \mu^{(T)}(M)] \quad \text{avec } \mu^{(k)}(x) = [(k-1) \times \mu^{(k-1)}(x) + v^{(k)}(x)] \times \frac{1}{k}$$

L'analyse du vecteur \underline{v} permet, de dénombrer les zones en mouvement, le sens du mouvement et la vitesse. Mecocci détermine alors la longueur L_i de chaque zone en mouvement i identifiée, puis génère le coefficient \underline{V} suivant :

$$\underline{V} = \frac{\sum_{i=1}^R L_i V_i}{\sum_{i=1}^R L_i^2}$$

Le signe de \underline{V} indique si le voyageur entre ou sort du bus.

La performance de ce système est de 98% de comptage exact.

2.3 LES APPLICATIONS BASEES SUR DES CAPTEURS LINEAIRES

On pourrait croire à priori que l'image linéaire n'apporte rien par rapport à l'image matricielle, et que son intérêt se limite à des cas exceptionnels. Mais il n'en est rien car les applications en traitement d'images utilisant des caméras linéaires comme capteurs sont très nombreuses. Ces caméras sont en général utilisées pour des applications où le temps de traitement doit être aussi bref que possible, c'est à dire pour des systèmes où le traitement des événements est réalisé en temps réel. Les caméras linéaires fournissent en effet moins d'information que les caméras matricielles car les images acquises sont unidimensionnelles. Elles permettent cependant de développer des processus de traitement plus rapides. Elles sont utilisées dans des domaines variés et sont très appréciées dans des applications pour lesquelles la résolution des caméras matricielles est trop faible pour être utilisable. Dans ce chapitre,

nous présentons quelques applications dans lesquelles les caméras linéaires remplacent avantageusement les matricielles.

2.3.1 Applications dans le domaine des transports

2.3.1.1 Détection d'intrusions et de chutes de voyageurs sur les voies

Une application intéressante dans le domaine de la sécurité dans les transports a été menée par L. Duvieubourg [DUV 91]. Il s'agissait de concevoir un système fonctionnant en temps réel permettant de détecter des intrusions dans des zones dangereuses à l'intérieur de systèmes de transports collectifs, comme les métros. L'application a porté sur le traitement de la sécurité au niveau des interfaces quais-voies.

En effet, de nos jours encore, plusieurs métros entièrement automatiques en service ou en construction, comme ceux de Vancouver, Miami, Détroit et la ligne D du métro de Lyon ont été conçus sans portes palières, de telle sorte qu'aucun dispositif ne peut prévenir les chutes sur les voies en station. Avec la disparition d'une présence humaine et en l'absence de portes palières, des dispositifs assurant que la voie est libre, de façon sûre et fiable, sont devenus indispensables. Le système développé par L. Duvieubourg permet de répondre à cette demande.

Deux configurations possibles pour disposer les caméras afin de surveiller l'interface quai-voie ont été envisagées.

La première consiste à disposer deux caméras linéaires en voûte visant une même rampe lumineuse installée en nez de quai (Figure 2.5.a).

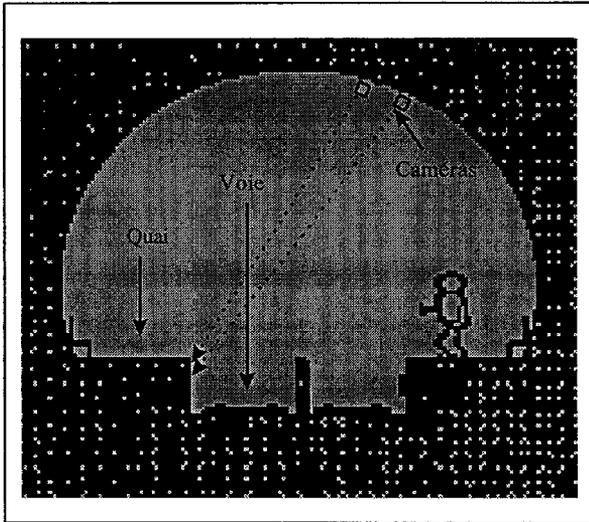


Figure 2.5.a : configuration de surveillance

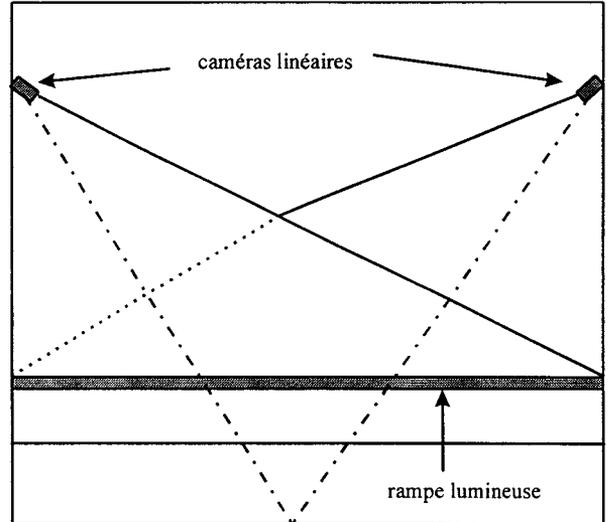


Figure 2.5.b : type de stéréoscopie

Les deux caméras visant la même rampe lumineuse, il est possible, par stéréoscopie, d'évaluer la distance et la taille des objets ou personnes traversant le plan de surveillance commun aux deux caméras (Figure 2.5.b).

La deuxième configuration envisagée consiste à disposer une caméra linéaire à chaque extrémité du quai, chacune d'elles visant une rampe lumineuse installée à l'autre extrémité (Figure 2.6.a).

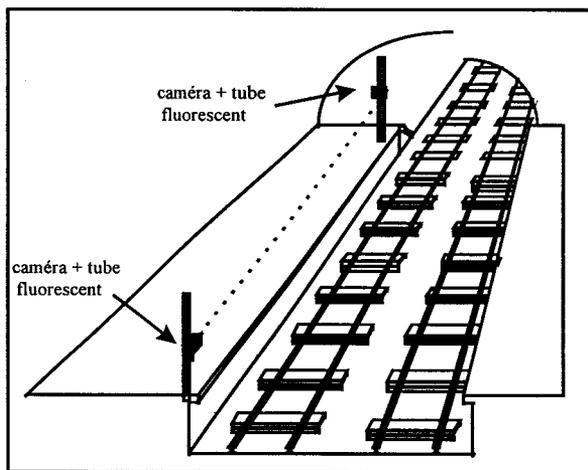


Figure 2.6.a : configuration de surveillance

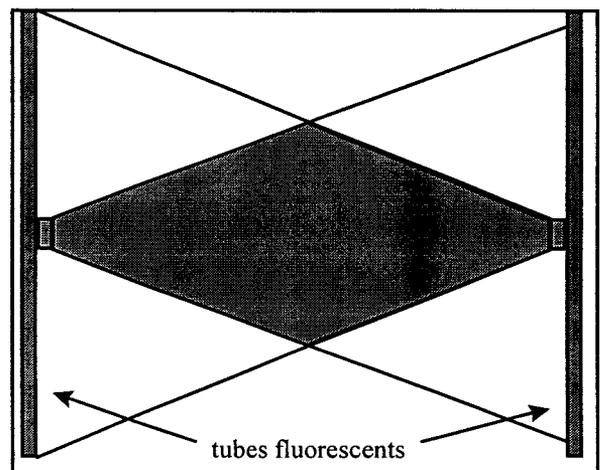


Figure 2.6.b : type de stéréoscopie

De la même manière, et par stéréoscopie, il est possible d'évaluer la position et la taille des objets ou personnes évoluant au bord du quai (Figure 2.6.b).

Les résultats en termes de qualité de détection sont excellents : tous les objets ou personnes de taille supérieure à un seuil sont correctement détectés.

Par contre, l'estimation de la taille et de la position des objets ou personnes détectés est plus délicate car elle dépend de plusieurs paramètres :

- la position de l'objet ou de la personne détectée sur le quai,
- la taille de l'objet ou de la personne.

2.3.1.2 Détection d'obstacles à l'avant des véhicules routiers

La détection d'obstacles à l'avant des véhicules routiers, menée dans le cadre du programme Eurêka PROMETHEUS avec l'aide des constructeurs automobiles PSA et Renault, vise à détecter automatiquement les obstacles qui se présentent sur la chaussée devant un véhicule, les localiser et déterminer leur nature, afin d'améliorer la sécurité routière [BRU 94][BUR 95]. L'objectif est de développer un détecteur embarqué sur une voiture de tourisme, basé sur la vision artificielle passive, fonctionnant en temps réel avec les moyens technologiques existants (Figure 2.7).

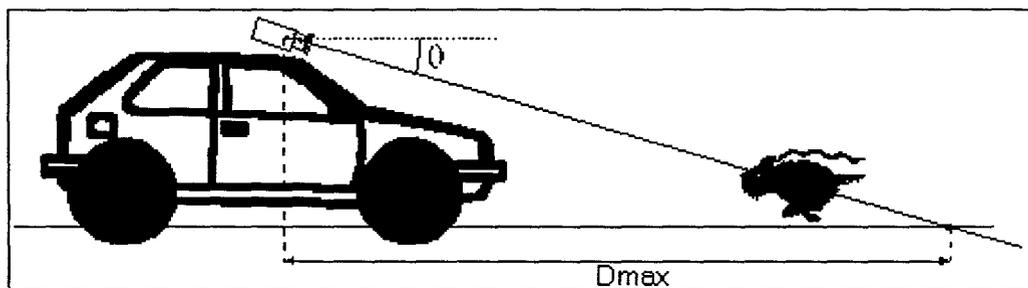


Figure 2.7 : système embarqué de détection d'obstacles

Ces exigences ont conduit à l'introduction d'un dispositif de prise de vue stéréoscopique utilisant des caméras linéaires. Pour cette application, ces capteurs ne comportant qu'une ligne de photocapteurs, réduisent de beaucoup la quantité d'information à traiter tout en offrant une résolution horizontale bien meilleure que celle offerte par des caméras matricielles.

Les caractéristiques particulières du capteur stéréoscopique linéaire ont nécessité la mise en œuvre d'une procédure de calibrage originale, destinée notamment à assurer que les champs de vision des deux caméras linéaires sont confondus (Figure 2.8).

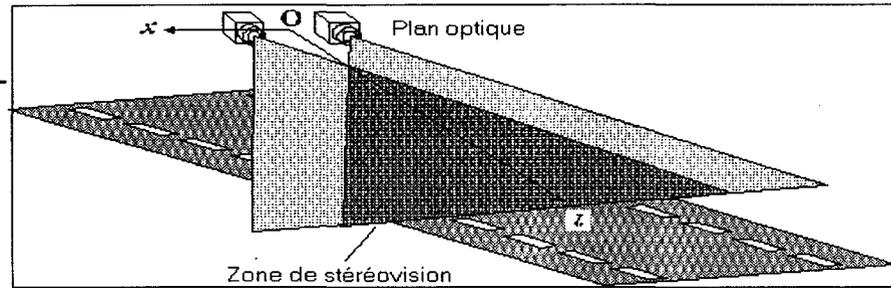


Figure 2.8 : mise en correspondance des lignes des deux caméras

Des méthodes de traitement d'images spécifiques à la stéréovision linéaire et plus particulièrement un algorithme de mise en correspondance des lignes des deux caméras, ont été développés.

Les résultats obtenus dans des conditions réelles d'utilisation montrent une bonne fiabilité de la détection d'obstacles. Par contre, le système est très sensible au mouvement de tangage du véhicule.

2.3.2 Applications dans le domaine industriel

2.3.2.1 Inspection automatique par vision linéaire

L'inspection automatique des matériaux plans en défilement continu présente un très grand intérêt dans le domaine textile [TAN 90], l'habillement [CAR 96], la sidérurgie [MAC 93], l'industrie des matières plastiques, la papeterie.

L'inspection de produits plats en défilement s'effectue, en général, en utilisant des caméras linéaires. On profite du mouvement des produits à inspecter sous la caméra, dont le capteur est placé dans la direction perpendiculaire au défilement pour reconstituer, ligne par ligne, la surface du matériau (Figure 2.9).

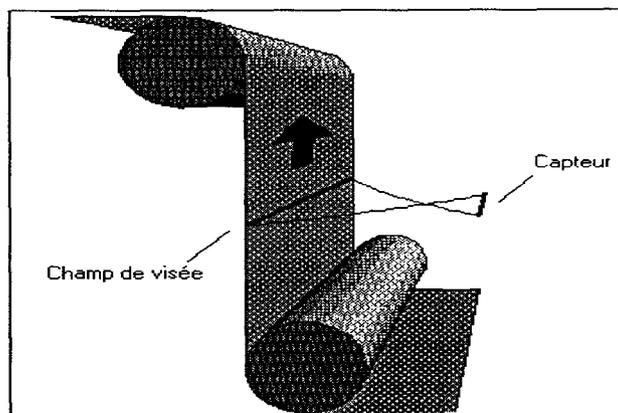


Figure 2.9 : inspection d'un produit plan en défilement

2.3.2.2 Numérisation de documents

Les capteurs linéaires sont la pièce essentielle des scanners, dans lesquels un document est filmé avec une très grande résolution pour être intégré, en tant qu'image, dans la composition d'un document, par exemple en Publication Assistée par Ordinateur.

2.3.2.3 Télédétection

Les satellites d'observation terrestre SPOT observent la terre grâce à des capteurs linéaires de grande résolution, plus fiables et plus rapides que les radiomètres à balayage mécanique qui équipaient les satellites plus anciens tels que les LANDSAT [COU 81]. C'est le mouvement orbital du satellite qui fournit le balayage.

2.3.2.4 Métrologie

Dans ce domaine, l'intérêt de la vision artificielle linéaire réside dans sa précision ($\pm 0,05\%$ dans le cas d'un capteur de 2048 points), mais aussi et surtout dans l'absence de tout contact mécanique susceptible de perturber la mesure. Un exemple de ce type d'application est le suivi des déformations d'une éprouvette soumise à des contraintes mécaniques [DUV 93]. La caméra vise l'éprouvette dans le sens de sa longueur. Deux repères tracés sur cette éprouvette permettent de suivre l'évolution de la déformation. Cette mesure se fait simplement en évaluant la distance entre les deux repères par des techniques d'analyse sub-pixel.

2.3.2.5 Spectroscopie

Dans le domaine de la spectroscopie, un capteur linéaire saisit l'image du spectre et la fournit à l'ordinateur qui en assure l'interprétation. Des capteurs spécifiques ont été développés pour cette application, qui exige une grande résolution, une absence totale de

distorsion géométrique, une grande sensibilité, une réponse spectrale très étendue, et une excellente uniformité de la réponse des différents photosites [BUI 89].

2.4 CONCLUSION

Nous avons vu, dans ce chapitre, que les systèmes de comptage de personnes ou d'objets pouvaient utiliser des données images issues de deux types de capteurs : caméras linéaires ou matricielles.

Loin d'avoir été exhaustifs dans l'énumération des systèmes présentés, nous avons cité un certain nombre d'applications dans différents domaines, utilisant ces capteurs.

Dans plusieurs des applications évoquées, les capteurs linéaires et matriciels peuvent être mis en concurrence : c'est le cas par exemple du système de comptage de piétons dans le métro parisien.

Les images relevées par des caméras matricielles offrent une vision globale de la scène à étudier. Par conséquent, elles génèrent une quantité d'information très importante qui peut, à première vue, apparaître comme un atout dans la prise de décision. Mais il s'avère que l'extraction de l'information pertinente nécessaire à cette prise de décision n'est que très rarement simple, augmentant ainsi considérablement les temps de traitement et limitant les possibilités d'implantations en temps-réel.

Les caméras linéaires à l'inverse fournissent une quantité d'information moins importante mais au bénéfice de traitements plus rapides permettant facilement des implantations en temps-réel sur des outils informatiques standard. Nous avons vu que les applications utilisant des capteurs linéaires sont nombreuses.

Pour le développement du système décrit dans ce mémoire de thèse, nous avons fait le choix du capteur linéaire à des fins de traitements en temps réel. Ainsi, un nouveau capteur linéaire a été développé, incluant notamment plusieurs améliorations technologiques permettant de très bonnes détections.

Ce nouveau capteur est présenté dans le chapitre suivant.

Chapitre III

Le dispositif expérimental

Chapitre III

LE DISPOSITIF EXPERIMENTAL

3.1 DESCRIPTION DU DETECTEUR DE BASE

3.1.1 Principe

Le système formé par une caméra linéaire visant un élément lumineux parallèle définit un plan de surveillance (Figure 3. 1).

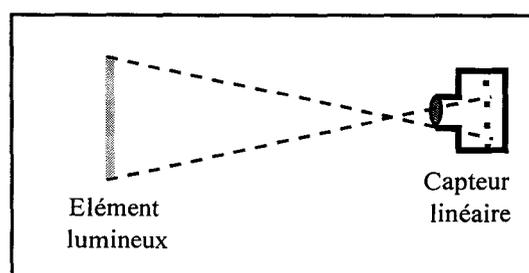


Figure 3.1 : plan de surveillance

Si aucun objet ne coupe le plan, l'élément lumineux est entièrement vu par la caméra dont le signal de sortie est une image ligne fidèle (Figure 3.2).

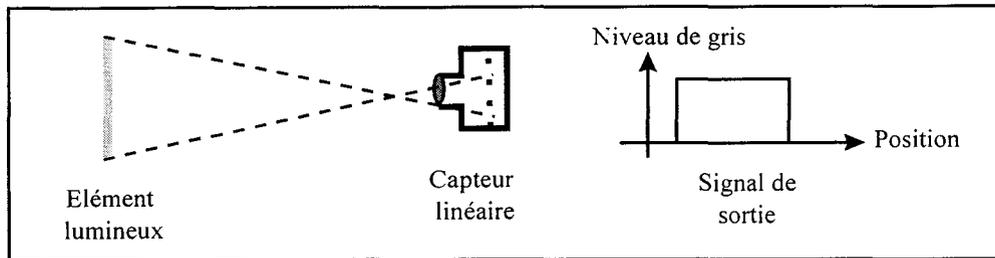


Figure 3.2 : formation d'une image ligne

La caméra disposant d'un capteur de N pixels, l'image ligne est représentée par une fonction $G(n)$ traduisant le niveau de gris du pixel n :

$$0 \leq G(n) \leq 255 \text{ avec } 1 \leq n \leq N$$

Compte tenu de la géométrie de la scène et de la distance focale retenue, l'élément lumineux n'occupe qu'une partie du capteur correspondant à une fenêtre F du signal de sortie (Figure 3.2). En général, nous cherchons à ce que la partie utile de l'élément lumineux corresponde à la longueur de la barrette CCD.

Si un objet opaque s'interpose entre l'élément lumineux et la caméra, celui-ci provoque une ombre sur le capteur qui va générer une diminution locale du signal de sortie qui permet de détecter sa présence (Figure 3.3).

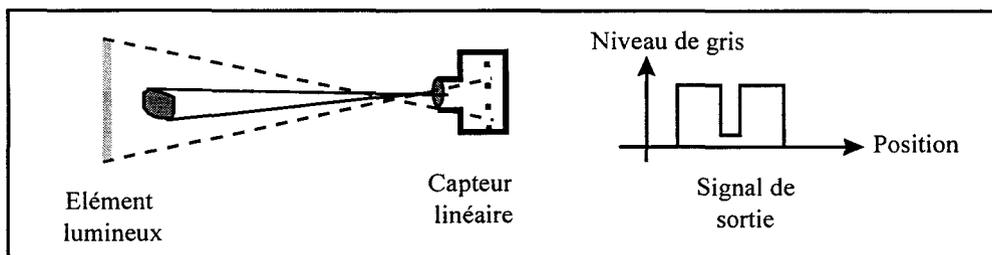


Figure 3.3 : détection d'une intrusion

3.1.2 Qualité de la réponse

3.1.2.1 Focalisation

En fonction de la position relative de l'objet, de la distance de mise au point de l'objectif de la caméra et de son ouverture, les bords de l'objet et de la fenêtre correspondant à l'élément lumineux vont être vus de façon plus ou moins nette, entraînant une pente plus ou moins raide au niveau de la réponse.

Une mise au point sur l'élément lumineux va donner des bords nets pour la fenêtre et des bords moins nets pour des objets plus proches de la caméra.

Par contre, une mise au point à une distance intermédiaire entre l'élément lumineux et la caméra, va donner, pour un objet à distance moyenne, des bords francs et une pente élevée pour la réponse alors que l'élément lumineux sera vu de façon plus floue, se traduisant par des pentes plus faibles pour les limites de la fenêtre d'observation (Figure 3.4).

3.1.2.2 Niveau des minima

La présence d'un objet coupant le plan de surveillance se traduit par une ombre sur le capteur linéaire de la caméra. La lumière, issue de l'élément lumineux, ne parvient plus à la caméra et on pourrait s'attendre à une réponse nulle au niveau des pixels correspondants.

En réalité, ce cas de figure théorique ne se présente jamais dans une situation réelle car le dispositif est placé dans une ambiance lumineuse. L'objet coupant le plan de surveillance reçoit inévitablement de la lumière dont il renvoie une partie vers la caméra. Les pixels correspondant à l'objet ne donnent donc pas une réponse nulle sur l'image ligne, mais un minimum (Figure 3.4) dont la valeur dépend de :

- la luminosité ambiante,
- la réflectance de l'objet présent dans le plan de surveillance,
- la sensibilité de la caméra, réglable par l'ouverture du diaphragme et le temps d'intégration du capteur.

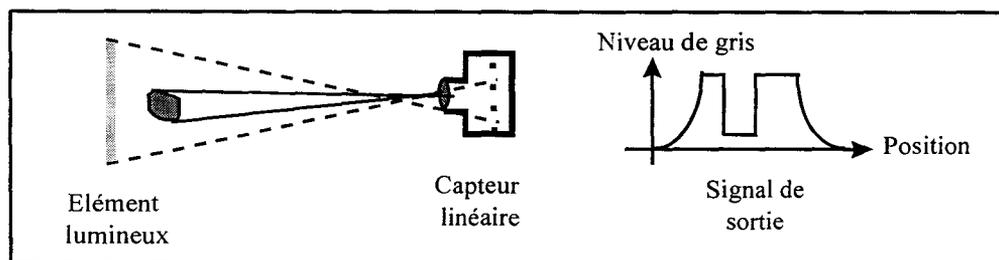


Figure 3.4 : signal de sortie dans des conditions réelles

3.1.2.3 Limites du détecteur

La quantité de lumière envoyée par l'élément lumineux, visé par la caméra, combinée à l'ouverture du diaphragme de l'objectif, fixe l'énergie lumineuse reçue par la barrette linéaire de photodétecteurs CCD. Compte tenu de cette énergie, on détermine le temps d'intégration

nécessaire pour atteindre un niveau proche du niveau maximum (255) dans la zone de l'élément lumineux vue par la caméra.

Ce temps d'intégration ΔT fixe la fréquence d'acquisition maximale envisageable. Il détermine également la vitesse maximale d'un objet détectable par le dispositif. En effet, pour être détecté, il doit occulter la lumière reçue par des pixels pendant au moins la durée d'une période d'intégration. Théoriquement, un objet d'épaisseur E traversant le plan à la vitesse V doit satisfaire :

$$\frac{E}{V} > \Delta T$$

Avec un temps d'intégration moyen de 20 millisecondes et une épaisseur d'objet de 10 centimètres, on trouve une vitesse maximale de 5 mètres par seconde, soit 18 km/h.

En réalité, dans les traitements, les pixels doivent être occultés pendant plusieurs acquisitions successives pour que la détection soit prise en compte car un filtrage, indispensable pour s'affranchir du bruit, oblige à éliminer toute détection isolée.

3.1.3 Représentation spatio-temporelle de l'intrusion

3.1.3.1 Principe

Les réponses présentées précédemment donnent l'image ligne captée par la caméra à un instant donné. Cette présentation de la réponse est peu commode pour l'analyse de l'intrusion. Il est intéressant de lui substituer une représentation bidimensionnelle obtenue en mettant en évidence le déroulement de l'intrusion. Nous réalisons ceci en :

- binarisant les réponses $G(t,n)$ successives avec un seuil convenablement choisi en fonction du niveau des minima recueillis en présence d'un objet comme indiqué au paragraphe précédent (Figure 3.5),

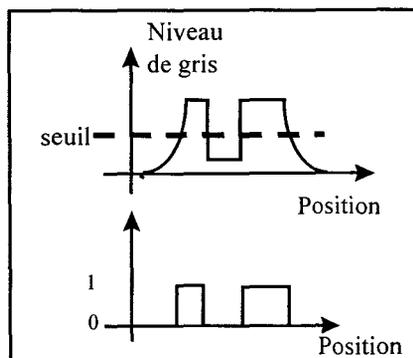


Figure 3.5 : seuillage des réponses

- représentant les images des pixels de la ligne d'acquisition par des lignes de pixels noirs ou blancs (Figure 3.6),
- en empilant les lignes successives $G(t, n)$, $G(t + \Delta T, n)$, $G(t + 2\Delta T, n)$...où ΔT représente la période d'acquisition. L'empilement des images successives crée ainsi une image bidimensionnelle.

Cette image binaire est en réalité une matrice d'éléments pixels-temps. La dimension horizontale de cette image est associée aux pixels de la barrette de la caméra alors que l'autre, verticale, est associée aux acquisitions successives séparées de la période ΔT .

Par extension, nous garderons le nom de pixels pour les éléments de cette image binaire, bien que la deuxième dimension ne soit pas représentative de l'espace, mais du temps.

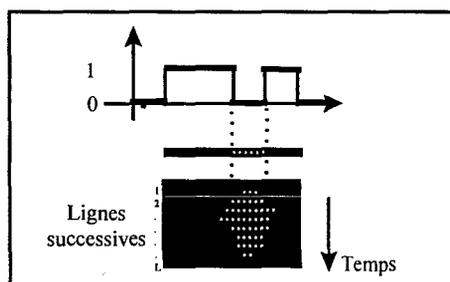


Figure 3.6 : représentation spatio-temporelle d'une séquence d'images lignes

Comme on peut le constater sur la figure 3.7, cette représentation permet non seulement de détecter la présence d'un objet traversant le plan de surveillance mais aussi, dans le cas d'une traversée à vitesse constante, d'avoir une bonne idée de la forme de cet objet.

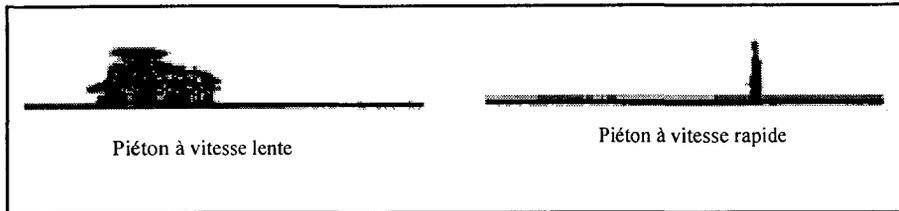


Figure 3.7 : détection d'intrusions avec des vitesses constantes

Sur la figure 3.7, un piéton est détecté de profil par la caméra linéaire. Sur la partie gauche de la figure, le piéton a une vitesse constante relativement lente. Son image est dilatée en largeur, ce qui signifie qu'il est resté longtemps dans le champ de la caméra compte tenu de sa vitesse faible. A l'inverse, sur la partie droite, le piéton détecté traverse le plan de surveillance à une vitesse plus rapide et la forme recueillie est plus étroite.

3.1.3.2 Lien entre l'objet réel détecté et sa représentation spatio-temporelle

Projection inverse d'un pixel de la barrette CCD de la caméra

On appelle "ombre" d'un objet traversant le plan de surveillance, la projection de l'objet sur le plan de l'élément lumineux.

La traversée du plan de surveillance par un objet fait apparaître une zone sombre sur le capteur CCD.

Le système optique utilisé associe à tout point de l'élément lumineux, un point (pixel) de la barrette CCD. Ainsi, à tout point Z de l'élément lumineux correspond un pixel i de la barrette CCD (Figure 3.8).

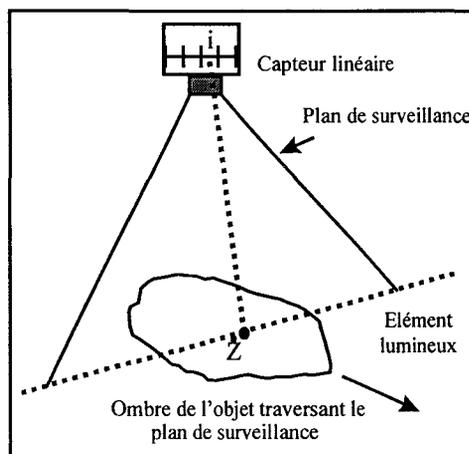


Figure 3.8 : projection inverse d'un pixel de la barrette CCD

Lors du mouvement d'un objet traversant le plan de surveillance, au point Z, on voit défiler un segment de l'ombre de l'objet. (Figure 3.9).

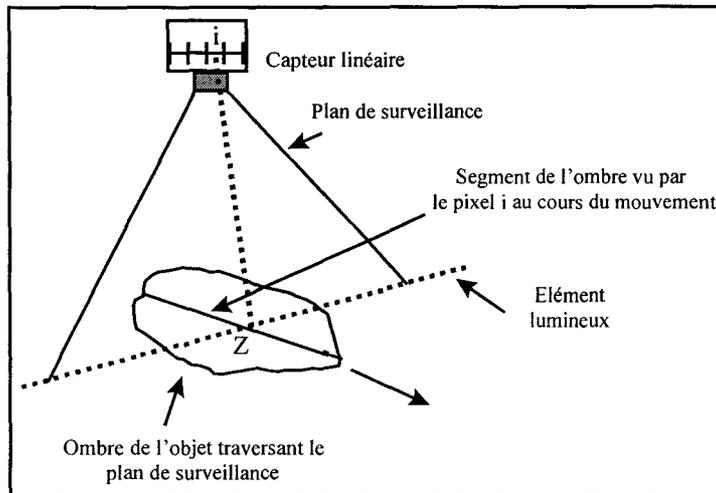


Figure 3.9 : positions de l'ombre de l'objet par rapport à un pixel

Description de l'objet

On décrit alors l'ombre de l'objet comme la juxtaposition de segments de droite qui sont vus par les différents pixels adjacents "touchés" par l'ombre (Figure 3.10).

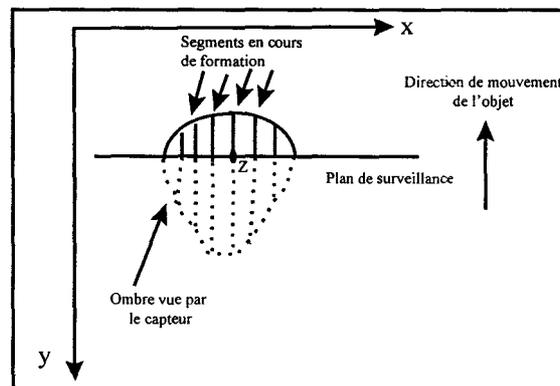


Figure 3.10 : définition des segments du mouvement sur l'ombre d'un objet

Nous verrons plus loin que pour chaque segment de l'ombre de l'objet on fait correspondre un autre segment dans le diagramme spatio-temporel.

Il est à noter qu'aucune information de vitesse n'est disponible sur les points internes de chaque segment, vue la nature du capteur utilisé. Les informations sont obtenues uniquement sur les transitions, c'est à dire les points extrêmes des segments.

3.2 LA CAMERA ACTIVE

3.2.1 Inconvénients du système de base

La faisabilité d'un capteur utilisant une caméra linéaire avait été menée avec le concours de la RATP dans le but de réaliser un détecteur de chutes sur les voies (Cf. chapitre II). Pour ces essais, les éléments lumineux visés par la caméra étaient des tubes fluorescents disposés :

- soit horizontalement en nez de quai, la caméra étant installée en voûte,
- soit verticalement à un bout du quai, la caméra étant en vis à vis à l'autre extrémité du quai.

Cette phase d'expérimentation a permis de mettre en évidence à la fois le grand intérêt du capteur pour la détection des intrusions et les faiblesses du système utilisé vis à vis des contraintes d'exploitation et des coûts d'installation. Ces défauts, liés à l'entretien et au remplacement des tubes fluorescents et au câblage de l'alimentation, étaient liés à la nature des éléments lumineux. Il a donc semblé intéressant de les remplacer. Nous avons proposé ce que nous avons appelé "la caméra active".

3.2.2 Description du dispositif à caméra active

Par rapport au système de base présenté précédemment à la figure 3.1, dans le dispositif avec une caméra active (Figure 3.11), l'élément lumineux est remplacé par une bande de matériau rétro-réfléchissant, tandis qu'une source de lumière éclairant cette bande a été ajoutée à la caméra. Celle-ci devient active et évite le câblage d'une alimentation au niveau de la "cible" rétro-réfléchissante. Cette source de lumière est constituée d'une ou plusieurs diodes électroluminescente (DEL).

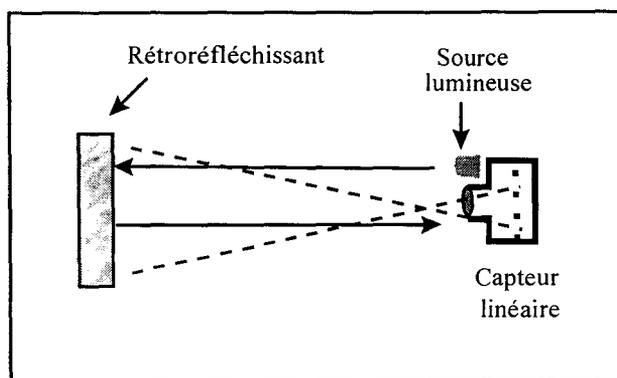


Figure 3.11 : caméra active

Mis à part un aller et retour de la lumière à la place d'un simple aller avec la configuration utilisant un tube, le fonctionnement est le même que celui du système de base. En absence d'objet coupant le plan de surveillance, la lumière issue de la caméra active frappe la cible rétro-réfléchissante qui la renvoie dans la direction incidente. La lumière retourne donc vers la caméra qui fournit un signal de niveau élevé en sortie (Figure 3.12).

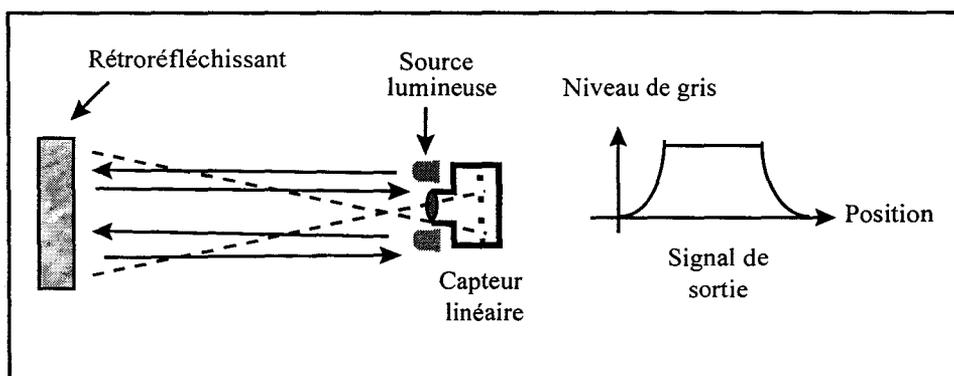


Figure 3.12 : signal en l'absence d'objet

Si un objet coupe le plan de surveillance, une partie de la lumière issue de la source active de la caméra n'atteint plus directement la cible (Figure 3.13).

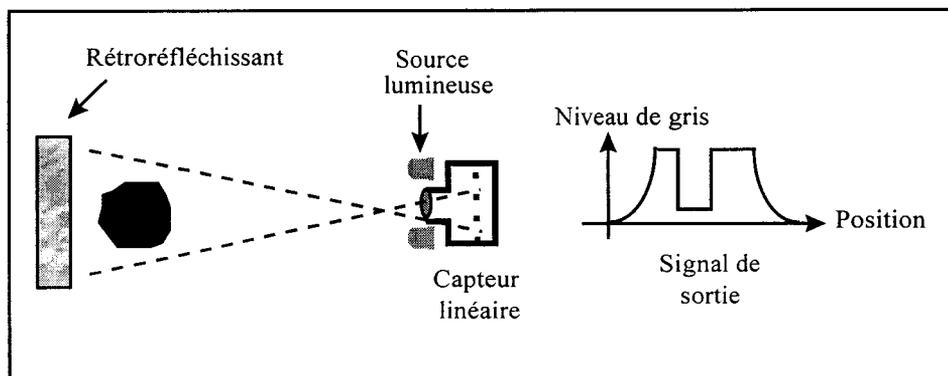


Figure 3.13 : détection d'objet

La caméra ne reçoit plus la lumière rétro-réfléchi par la cible mais seulement une faible énergie lumineuse réfléchi, de façon spéculaire, par l'objet qui vient s'ajouter à la contribution due à la réflexion de la lumière ambiante présentée précédemment.

Le dispositif est réglé en tenant compte de l'énergie de la source et de la conjugaison de la sensibilité des photo-capteurs, du temps d'intégration et de l'ouverture du diaphragme de l'objectif. Dans ce cas, un objet coupant le plan de surveillance, quelle que soit sa nature et l'intensité de l'éclairage ambiant (soleil, brouillard, lumière artificielle), la réponse des pixels occultés doit être notablement plus faible que celle des pixels recevant l'énergie renvoyée par le rétro-réfléchissant.

Cette caractéristique du dispositif est obtenue aisément grâce au type de source de lumière et aux qualités intrinsèques des matériaux rétro-réfléchissants présentés ci-dessous.

3.2.3 Propriétés liées aux matériaux rétro-réfléchissants

Les matériaux rétro-réfléchissants utilisés, mis au point pour la signalisation routière, ont la propriété de renvoyer une très haute énergie lumineuse. Dans la direction incidente. Ceci est obtenu grâce aux propriétés optiques remarquables des micro-prismes ou micro-billes utilisés pour réaliser ces matériaux.

Ces éléments, incrustés sur un support autocollant, sont protégés par une couche transparente assurant des qualités de résistance du produit vis à vis des intempéries et de l'abrasion. Le matériau renvoie une énergie lumineuse très importante, un rapport de 2000 comparée à un matériau "normal" (spéculaire) blanc étant facilement atteint. On comprend, dans ces conditions, qu'il est possible d'ignorer les effets de la réflexion spéculaire sur un objet coupant le plan et de la lumière ambiante.

Dans le cas d'un corps à réflexion spéculaire, l'énergie apportée par un rayon incident est renvoyée sur un demi-espace avec une intensité faible dans chaque direction (Figure 3.14).

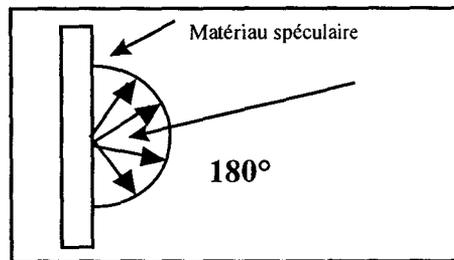


Figure 3.14 : angle de renvoi de lumière avec un matériau spéculaire

Dans le cas du matériau rétro-réfléchissant, cette énergie n'est renvoyée que dans un angle solide très faible, de l'ordre d'une fraction de degré, mais avec une intensité de l'ordre d'un millier de fois plus forte (Figure 3.15).

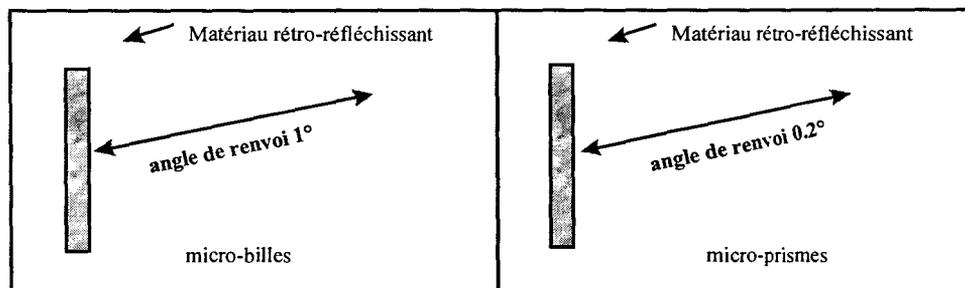


Figure 3.15 : angle de renvoi de lumière avec rétro-réfléchissant : micro-billes ou micro-prismes

Cette caractéristique favorable du point de vue de la quantité d'énergie renvoyée va de paire avec une difficulté de conception de la caméra active. Comme l'angle de réflexion de l'énergie lumineuse est très faible, celle-ci est renvoyée vers la source, c'est à dire les DEL. Il faut donc que les DEL soient installées le plus proche possible de l'objectif de la caméra pour qu'un maximum d'énergie soit renvoyé à la caméra.

Une implantation de la source lumineuse trop éloignée de l'objectif ne ramènerait que peu d'énergie sur l'objectif et par conséquent sur la barrette CCD. En effet, le maximum d'énergie retournée se trouve centré sur l'axe de la source lumineuse (Figure 3.16). Dans ces conditions, la disposition qui paraît à priori préférable consiste à placer les DEL à la périphérie de l'objectif pour ne pas le masquer.

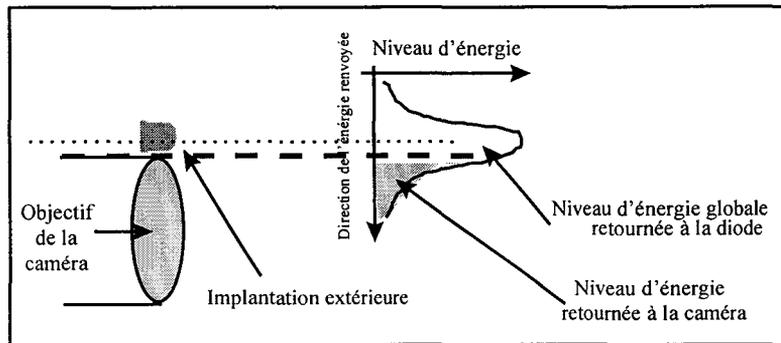


Figure 3.16 : implantation de la source lumineuse en périphérie de l'objectif

Cette disposition doit cependant être écartée pour choisir une disposition qui masque partiellement l'objectif mais ramène néanmoins plus d'énergie (Figure 3.17). Nous avons dit, auparavant, que l'énergie est renvoyée par le rétro-réfléchissant dans un angle très faible, donc vers la source émettrice. Si celle-ci est placée très près de l'objectif de la caméra elle le masquera mais ramènera quand même plus d'énergie à la caméra que dans le cas où la source émettrice est placée plus loin de l'objectif.

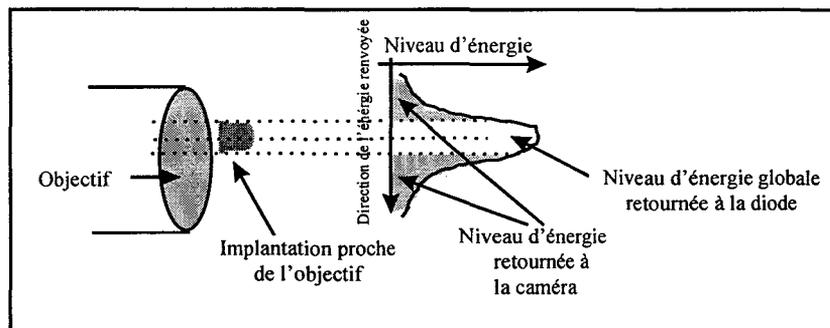


Figure 3.17 : implantation préférable

Le choix du nombre de diodes, et de leur position, fait entrer en ligne de compte :

- la distance et l'angle sous lequel est vue la bande rétro-réfléchissante,
- l'intensité lumineuse fournie par une diode,
- la sensibilité de la caméra.

3.2.4 Utilisation de diodes électroluminescentes infrarouges pulsées

3.2.4.1 Utilisation de sources infrarouge

L'utilisation de sources infrarouge présente deux avantages :

- la discrétion du dispositif qui n'est pas perçu par les passagers,

- une amélioration de la réjection de la contribution due à la lumière ambiante par utilisation de filtres éliminant la lumière visible.

3.2.4.2 Utilisation de lumière pulsée

L'utilisation de lumière pulsée apporte deux améliorations :

- elle permet, grâce à l'exploitation des qualités des diodes, de fournir une puissance instantanée plus importante,
- elle permet, en alimentant la source de lumière une acquisition sur deux, d'éliminer l'influence de l'éclairage ambiant en réalisant la différence des deux images successives; l'une en activant la source de lumière et la suivante sans l'utiliser.

3.3 UTILISATION D'UN DOUBLE PLAN DE SURVEILLANCE

3.3.1 Principe

La faible quantité d'information disponible avec notre système à base de caméra linéaire, comparée à celle fournie par une caméra matricielle, s'est avérée pénalisante pour l'analyse du mouvement des objets traversant le plan de surveillance. Une évolution du système de base a permis de s'affranchir de ce défaut. Il s'agit de l'ajout d'un second plan de surveillance, parallèle au premier.

Avec ce dispositif, deux types d'informations supplémentaires deviennent disponibles : le sens de traversée et la vitesse de l'objet.

La figure 3.18 fournit une représentation du dispositif permettant de déterminer la vitesse et le sens de traversée. On peut trouver, dans le cas d'une intrusion "normale" d'un objet d'épaisseur supérieure à l'écartement des plans, les phases suivantes :

- entrée de l'objet dans le premier plan,
- entrée de l'objet dans le second plan,
- sortie de l'objet du premier plan,
- sortie de l'objet du second plan.

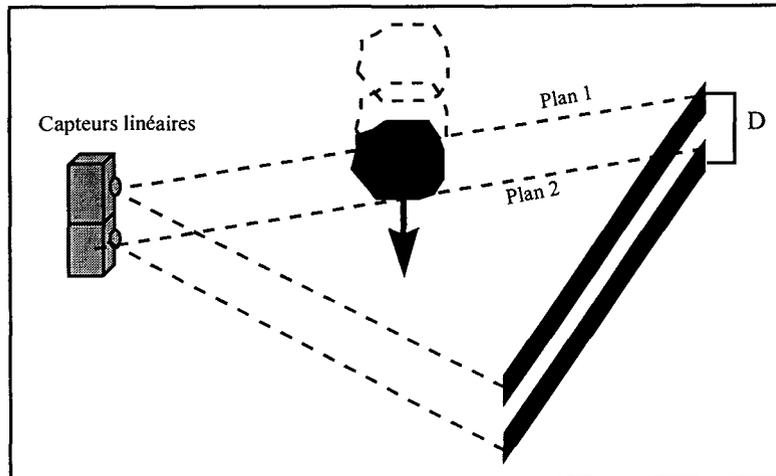


Figure 3.18 : connaissance du sens et de la vitesse de traversée

Du point de vue des représentations spatio-temporelles des images lignes acquises par les caméras, on trouve, si l'objet n'est pas déformable et si sa vitesse est constante, des formes identiques décalées dans le temps d'un intervalle :

$$\Delta T = D / V$$

où D est l'écartement entre les plans de surveillance parallèles et V la vitesse de l'objet.

3.3.2 Méthodologie du calcul de la vitesse

En réalité, les phases de pénétration de l'objet, citées plus haut, d'un point de vue global doivent être analysées au niveau des pixels de la barrette linéaire. Chaque pixel "touché" par le passage de l'objet va voir son niveau de gris seuillé passer du maximum au minimum puis revenir au maximum (Figure 3.19).

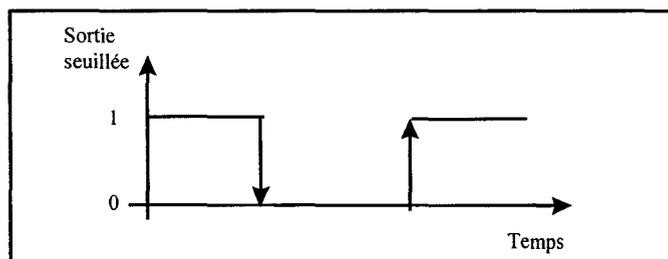


Figure 3.19 : évolution temporelle du niveau de luminance reçu par un pixel

Vue la géométrie du système, et la méconnaissance de la direction exacte de déplacement des objets traversant les plans, on ne s'intéresse qu'à la composante

longitudinale de la vitesse, perpendiculaire aux plans. On étudie donc l'évolution temporelle des niveaux de luminance reçus par les pixels homologues des deux barrettes.

Le mouvement est analysé en observant le phénomène d'occultation-désoccultation respectivement aux points $Z1$ et $Z2$ correspondant aux pixels i des deux barrettes CCD (Figure 3.20).

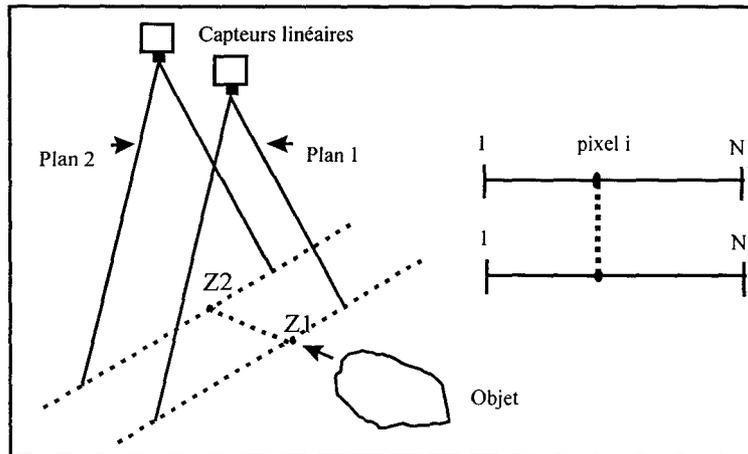


Figure 3.20 : projection inverse des pixels avec deux barrettes

Les signaux de l'évolution du niveau de luminance des pixels font apparaître des fronts caractéristiques des entrées et sorties de l'objet par rapport aux deux plans. Les courbes sur la figure 3.21 montrent les quatre événements de base.

- A : entrée dans le plan 1
- B : sortie du plan 1
- C : entrée dans le plan 2
- D : sortie du plan 2

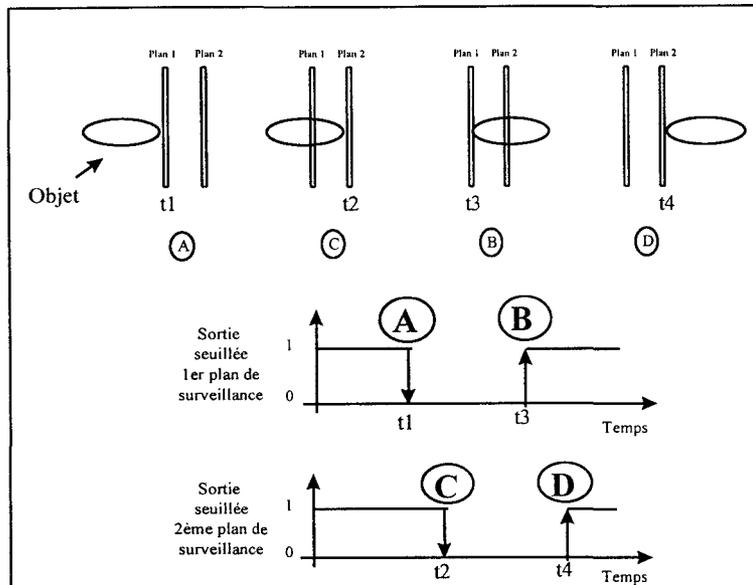


Figure 3.21 : chronologie des intensités lumineuses reçues par des pixels homologues

La succession de ces différents fronts permet de connaître le déroulement de l'intrusion.

Les décalages temporels entre ces différents événements donnent des informations sur l'objet traversant les plans :

- A à C : décalage temporel $T_{AC} = t_2 - t_1$ donne la vitesse d'entrée au pixel i ,
- B à D : décalage temporel $T_{BD} = t_4 - t_3$ donne la vitesse de sortie au pixel i ,
- A à B : décalage temporel $T_{AB} = t_3 - t_1$ lié à la longueur réelle du segment vu par le pixel i de la première barrette CCD au cours du déplacement et à la vitesse de traversée,
- C à D : décalage temporel $T_{CD} = t_4 - t_2$ lié à la longueur réelle du segment vu par le pixel i de la deuxième barrette CCD au cours du déplacement et à la vitesse de traversée.

Dans le cas d'un objet indéformable et en déplacement *uniforme* :

- $T_{AB} = T_{CD}$, traduit l'égalité des mesures de la longueur du segment par les deux plans,
- $T_{AC} = T_{BD}$, traduit l'égalité entre les vitesses d'entrée et de sortie.

De la même façon, ce calcul, mené sur tous les pixels homologues des deux barrettes, donne le même résultat en termes de vitesse d'entrée et de sortie pour la traversée d'un objet qui ne se déforme pas pendant cette traversée.

3.4 DISPOSITIF DE COMPTAGE DES PIETONS

3.4.1 Description de l'installation

3.4.1.1 Dispositif sur site

Dorénavant, nous ne parlerons plus de détection d'objets mais de détection de piétons. Le dispositif sur site, qui permet de compter les piétons comporte (Figure 3.22) :

- les deux caméras linéaires actives, qui sont en général fixées au plafond et déterminent deux plans de surveillance perpendiculaires à la direction de déplacement des piétons,
- les bandes rétro-réfléchissantes collées au sol, dont la longueur correspond à la largeur du site de mesures (couloir, entrée...).

La paire de caméras est connectée par une liaison filaire à un micro-ordinateur pour l'acquisition et le traitement des données. Une description détaillée est fournie ci-dessous.

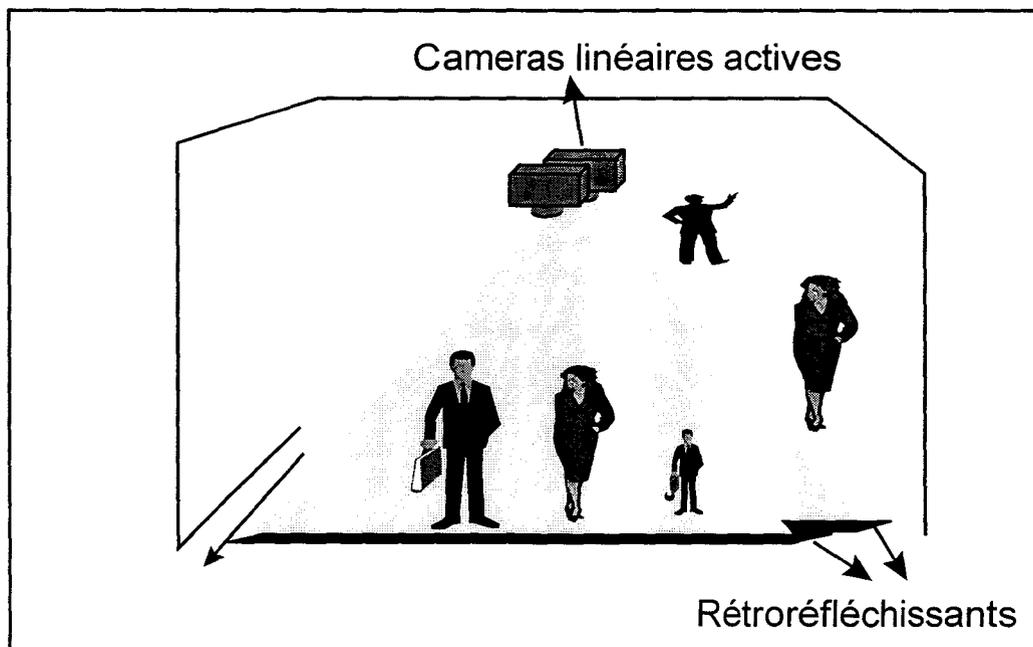


Figure 3.22 : dispositif sur site

3.4.1.2 Dispositif complet

Le dispositif complet, matériel et logiciel, comprend les composantes suivantes (Figure 3.23) :

- les bandes rétro-réfléchissantes collées au sol et placées de façon à être vues par les caméras,
- les deux caméras linéaires IDC100 de marque I2S. Les capteurs CCD utilisés comprennent 1024 pixels dont les niveaux de luminance sont codés sur 256 niveaux de gris,
- les diodes électroluminescentes qui émettent une lumière non visible dans le proche infrarouge (880 nm) avec un angle d'émission de 20°. Les diodes sont montées de façon à recevoir le maximum d'énergie rétro-réfléchi,
- des filtres placés sur les objectifs des caméras, qui permettent de supprimer une quantité non négligeable de l'énergie lumineuse parasite (soleil, éclairage public,...),
- les cartes d'acquisition de marque I2S. Leur rôle consiste à transmettre les ordres du logiciel d'acquisition vers les caméras, à faire transiter les lignes acquises des caméras vers la mémoire vive du PC via la liaison caméra-carte, puis le BUS du PC. Les données sont des lignes de 1024 pixels dont les niveaux de gris sont codés de 0 à 255,
- les liaisons filaires parallèles caméras-cartes d'acquisition,
- le logiciel de pilotage des cartes d'acquisition qui, pour un cycle, effectue les tâches suivantes:
 - allumage des DEL,
 - lancement de l'acquisition pendant un temps ΔT ,
 - temps d'attente pour le traitement de la ligne précédente,
 - fin de l'attente,
 - extinction des DEL,
 - lecture des images lignes et mise en mémoire vive du PC.

Une fois les lignes mises en mémoire, il est possible de les stocker sur disque pour des traitements ultérieurs, ou de les traiter directement, en temps réel. Les logiciels qui réalisent les traitements sont écrits en langage C.

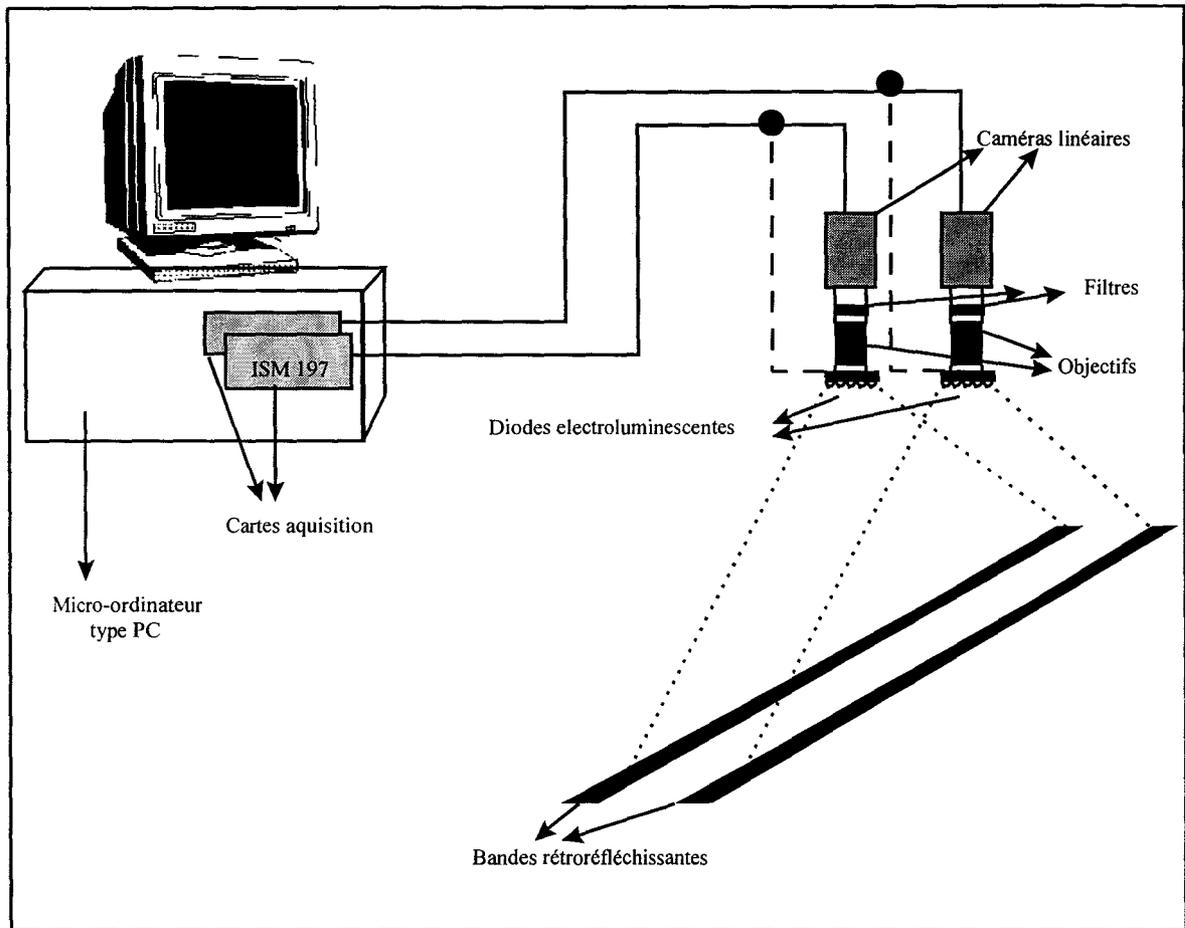


Figure 3.23 : dispositif complet

3.4.2 Représentation spatio-temporelle du mouvement

3.4.2.1 Allure des images obtenues

Après binarisation et empilement des lignes d'acquisitions successives, nous obtenons deux images binaires, analogues à celle décrite au §3.1.3.1. Ces images représentent chacune une séquence d'images lignes saisies à la période d'acquisition de ΔT . Les deux dimensions des deux images binaires représentent (Figure 3.24) :

- en horizontal, la largeur du site de mesures contrôlé par la caméra à l'aide de la barrette CCD,
- en vertical, l'information temporelle qui correspond à l'empilement des images lignes dans le temps. L'ajustement de la fréquence d'acquisition des lignes sera abordé plus loin dans ce mémoire.

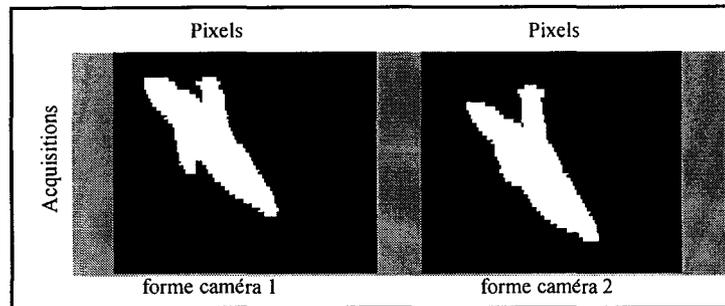


Figure 3.24 : un exemple de deux images binaires

Terminologie et vocabulaire

Un certain nombre de termes vont être utilisés, très souvent, tout au long du mémoire. Les termes en question sont : **trace**, **forme**, **silhouette** et **noyau** dont nous donnons ci-dessous la définition.

Les images lignes binarisées donnent une image bi-dimensionnelle (§ 3.1.3.1) que nous appelons **trace**.

Nous appelons **forme**, la trace laissée dans la représentation espace-temps (Figure 3.25) par un ou plusieurs piétons. Sur cette figure, on ne sait pas effectivement, si la forme, dans la représentation espace-temps (caméras 1 et 2) correspond à un ou plusieurs piétons.

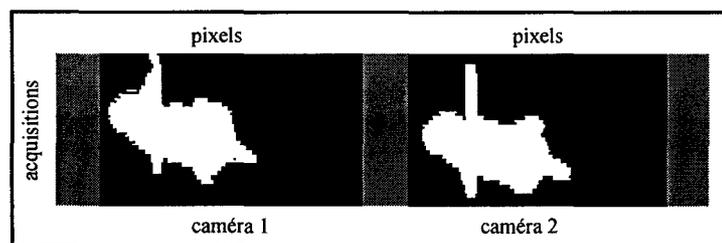


Figure 3.25 : exemple de forme

Nous appelons **silhouette**, la forme laissée par *un seul piéton* passé au travers des deux plans de surveillance (Figure 3.26). Une forme est en fait constituée de plusieurs silhouettes qui se chevauchent plus ou moins.

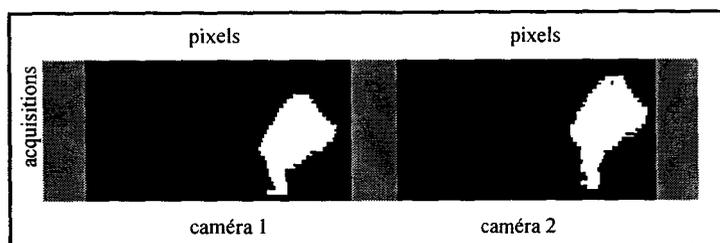


Figure 3.26 : exemple de silhouette

Nous appelons **noyau** le résultat des traitements réalisés sur une forme ou une silhouette. Ces traitements sont décrits dans les chapitres suivants. Une silhouette traitée fournit un noyau alors qu'une forme traitée fournit un ou plusieurs noyaux. Nous verrons dans le chapitre suivant que l'objectif est d'associer un noyau à un piéton.

3.4.2.2 Influence de la vitesse des piétons sur les images obtenues

La vitesse des piétons détectés va influencer les formes de ces piétons dans les images binaires.

Décalage temporel

Dans les deux images binaires présentées sur la figure 3.24, on peut noter le décalage temporel des deux formes correspondant au même piéton qui a d'abord traversé le plan surveillé par la caméra 1, puis celui associé à la caméra 2.

Dilatation temporelle

Même si la vitesse des piétons peut-être calculée pour chaque pixel affecté par le mouvement, on peut en avoir une idée par l'observation de la "longueur" des formes laissées par leur passage sous les caméras. La **longueur** d'une forme est le nombre de lignes d'acquisition pendant lesquels un plan de surveillance est occulté par la présence d'un piéton.

On constate sur la figure 3.27 que la longueur de la forme est inversement proportionnelle à la vitesse. Le piéton qui correspond à la forme 1 marche vite ou court sous les caméras, sa forme est courte car le plan de surveillance a été occulté pendant peu de périodes d'acquisitions. A l'inverse, le piéton lié à la forme 5 se déplace lentement conduisant ainsi à une forme plus longue dans l'image binaire.

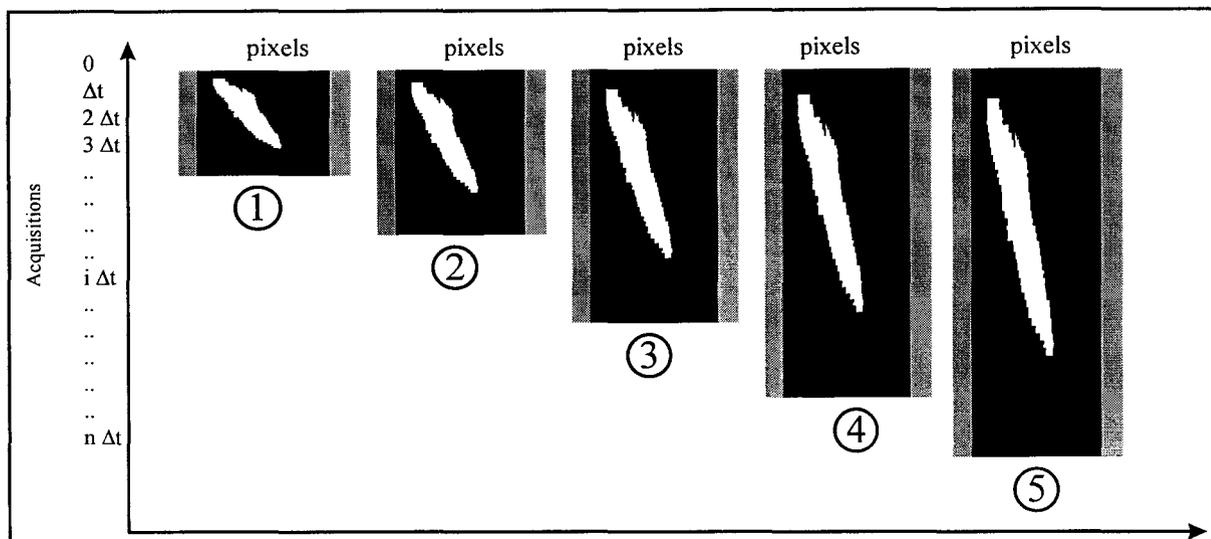


Figure 3.27 : description graphique des vitesses

3.5 LE CALCUL DE LA VITESSE

Grâce aux deux plans de surveillance, il est possible d'évaluer la vitesse des piétons ainsi que le sens de leur déplacement.

La vitesse d'un piéton qui passe sous les caméras peut être immédiatement obtenue à partir du temps qu'il a mis pour parcourir la distance D qui sépare les deux plans.

Connaissant D et le temps mis pour la parcourir, le résultat est immédiat. Ce raisonnement n'est valable que si l'on considère qu'un piéton génère une forme binaire qui ne se déforme pas pendant son passage à travers les plans de surveillance.

Mais, non seulement les piétons présentent des formes géométriques irrégulières, mais ces formes se déforment au cours de leur détection. Cela conduit à observer non pas une vitesse globale homogène, mais plusieurs vitesses différentes pour un même piéton. Lorsque les piétons traversent les plans de surveillance, ils occupent une partie de la largeur du site de mesure. Leur vitesse est évaluée pour chaque pixel affecté par le mouvement de ces piétons.

3.5.1 Calcul de la vitesse des piétons au niveau des pixels affectés par le mouvement

La manière dont la vitesse des piétons est calculée est expliquée sur les figures 3.28.a et 3.28.b.

Sur la figure 3.28.a, représentant une forme correspondant à un piéton dans le diagramme spatio-temporel, on remarque que la forme est présente entre les deux plans de surveillance pendant un certain temps. Le piéton traverse d'abord le plan observé par la caméra 1, puis celui observé par la caméra 2.

En réalité, chaque ligne binaire correspondant à l'une ou l'autre des caméras est représentée par une série de valeurs 1, si le plan de surveillance correspondant n'est pas occulté.

Par contre, quand un piéton traverse les plans de surveillance, les pixels correspondant à la largeur qu'il occupe sont occultés.

Soit la fonction $\Delta N(k)$ définie comme étant le nombre d'acquisitions séparant une transition de la valeur 1 à la valeur 0 pour le pixel k de la caméra 1 ($N1_k$) et la transition de la valeur 1 à la valeur 0 pour le pixel de même rang pour la caméra 2 ($N2_k$) (Figure 3.28.a)

$$\Delta N(k) = (N2_k - N1_k)$$

Pour chaque couple de pixels homologues, l'information vitesse est dérivée de la fonction $\Delta N(k)$, de la distance D entre les deux plans de surveillance et de la période d'acquisition ΔT . Cette vitesse est calculée pour chaque pixel sur les deux images binaires des deux caméras. On obtient ainsi le "signal vitesse" $S(k)$ qui est défini par :

$$S(k) = \frac{D}{\Delta T \cdot \Delta N(k)}, k = 1, \dots, N$$

Dans cette expression, le numérateur est exprimé en mètres, le dénominateur en secondes, la vitesse est donc obtenue en mètres par seconde.

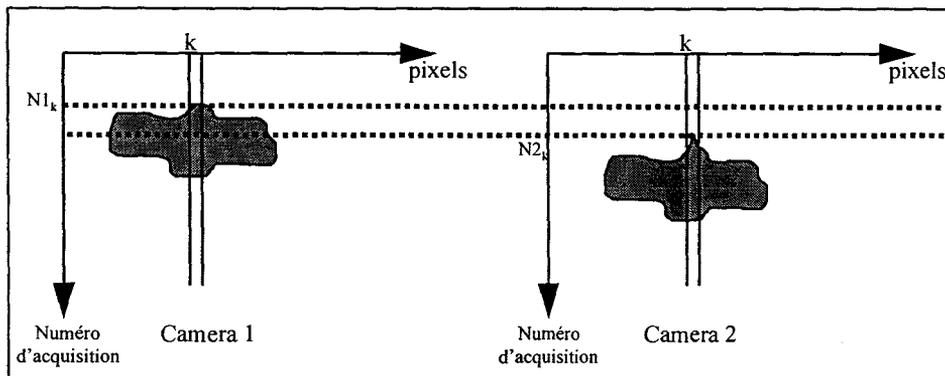


Figure 3.28.a : calcul de la vitesse

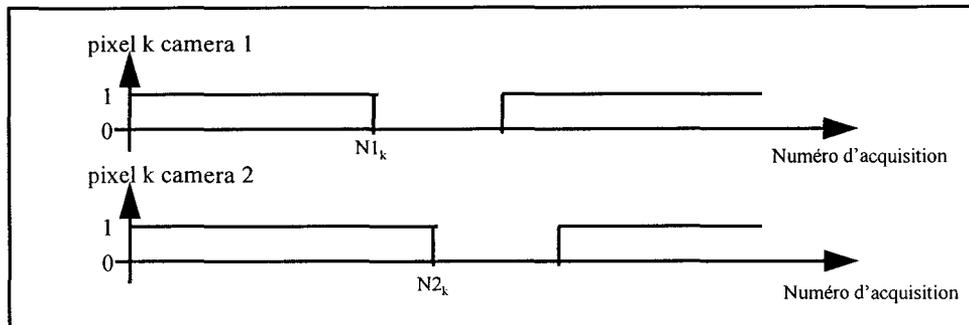


Figure 3.28.b : calcul de la vitesse

Dans la formule précédente, D et ΔT étant fixes, il est très facile à partir du temps mis par le piéton pour traverser les deux plans, exprimé sous la forme d'un multiple de ΔT , de calculer sa vitesse en mètres par seconde. Dans tout le reste de ce mémoire, la vitesse sera considérée en mètres par seconde ou par souci de commodité exprimée en nombre d'acquisitions.

3.5.2 Vitesses obtenues dans le cas d'un piéton

Dans le calcul des vitesses, pour chaque pixel affecté par le mouvement, nous faisons deux hypothèses :

- la vitesse d'entrée au pixel i est égale à la vitesse de sortie au même pixel. Seule, cette vitesse d'entrée est dorénavant prise en compte et appelée simplement vitesse,
- la vitesse est constante pendant le temps d'occultation d'un pixel affecté par le mouvement.

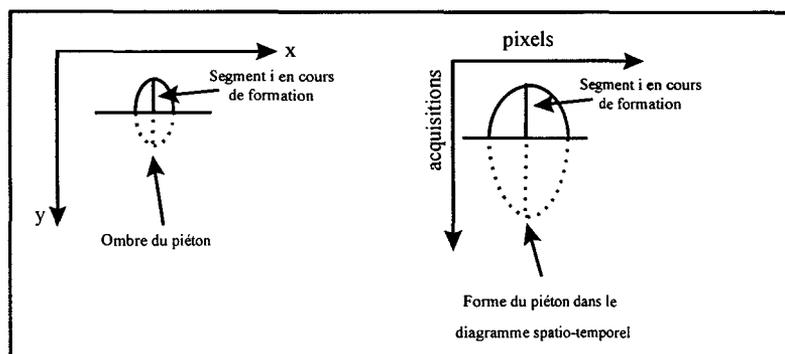


Figure 3.29 : association entre les segments d'une ombre d'un piéton et ceux de sa forme

Aux segments définis en 3.1.3.2, on peut faire correspondre des segments dans le diagramme spatio-temporel (Figure 3.29).

Ayant considéré la vitesse comme constante pendant tout le temps d'occultation d'un pixel affecté par le mouvement, il est logique d'affecter cette vitesse à tous les pixels du segment qui lui correspond dans la représentation spatio-temporelle (Figure 3.30).

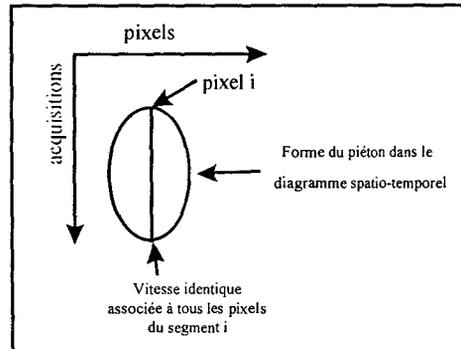


Figure 3.30 : Vitesses associées aux pixels intérieurs à une forme

Nous avons donc défini la vitesse à deux niveaux :

- sur les contours des formes, une série de vitesses calculées aux pixels de transitions,
- pour les points intérieurs aux formes, grâce à l'hypothèse de mouvement uniforme, on affecte pour chaque pixel de chaque segment, la vitesse calculée aux pixels de transition.

3.6 CONCLUSION

Nous avons vu, dans ce chapitre, que grâce au dispositif développé, il est possible de disposer de trois grandeurs bien différenciées qui vont être utilisées pour procéder au comptage. Il s'agit de :

l'information spatiale qui correspond à l'axe défini par l'orientation des barrettes CCD. Cette information spatiale est obtenue avec une résolution proche des N pixels, disponibles sur la barrette CCD, pour la largeur du passage surveillé. C'est sur cette dimension que nous allons évaluer la largeur des formes générées par les piétons qui passent sous les caméras. Nous verrons, dans le chapitre VII, dédié aux résultats de comptage, que l'expérimentation que nous avons menée pour tester le système global s'est déroulée dans un site où 4 personnes côte à côte pouvaient passer sous le système en même temps. L'information spatiale est donc

très importante, notamment pour les traitements décrits dans le chapitre V. Si la résolution des caméras linéaires utilisées est très importante, elle présente l'inconvénient de ralentir les traitements. Notre but, décrit dans l'introduction de cette thèse, est de concevoir et de développer un système fonctionnant en temps réel.

Ceci nous pousse à minimiser les temps de traitements autant que faire se peut. C'est pourquoi, nous verrons plus loin et notamment dans les chapitres V et VI, consacrés aux algorithmes de comptage, que nous avons diminué cette résolution pour certains traitements,

l'information temporelle qui est liée à la fréquence d'acquisition des lignes. Cette deuxième dimension, qui caractérise les formes obtenues par le capteur, est en fait une dimension spatio-temporelle. La période d'acquisition ΔT et la distance D entre les deux plans de surveillance ont été fixées de façon à répondre aux exigences suivantes :

- si la distance entre les deux plans de surveillance est trop élevée, la déformation des piétons entre les passages sous les deux caméras peut être importante et conduire ainsi à deux formes très différentes,
- si le nombre d'acquisitions par unité de temps n'est pas assez élevé, nous n'aurons pas suffisamment de lignes d'acquisitions quand un piéton traverse le système et par conséquent pas assez d'information pour évaluer sa vitesse.

Une période ΔT de 10 millisecondes, qui correspond à une fréquence de 100 acquisitions par seconde et une distance D de 10 centimètres est un bon compromis. En effet, une distance de 10 centimètres entre les plans évite les fortes déformations et une période de 10 millisecondes permet d'avoir une bonne information vitesse. Ce réglage permet de capter tous les piétons, même s'ils circulent à vitesse élevée. En effet, pour qu'un piéton traverse les deux plans en moins d'une acquisition, il faudrait qu'il coure à plus de 10 mètres/seconde, c'est à dire à plus de 36 km/h.

L'information vitesse qui est obtenue grâce aux deux plans de surveillance. Avec la méthode de calcul des vitesses précédemment expliquée, nous avons défini la vitesse des piétons au niveau de chaque pixel affecté par le mouvement. Nous obtenons donc pour chaque piéton une série de vitesses.

Lors de leur passage sous les caméras, les piétons génèrent des valeurs de vitesses erratiques liées aux mouvements de leurs membres pendant leur passage sous les caméras. Ces vitesses, telles que déterminées précédemment, sont très difficiles à utiliser pour

caractériser un mouvement global pour chaque piéton. Un filtrage de ces vitesses s'avère nécessaire. Il est exposé dans le chapitre suivant dans lequel sont également présentés un certain nombre de prétraitements.

A tous les pixels intérieurs aux formes, on a affecté une vitesse qui est celle du pixel de transition qui leur correspond sur les segments. Cette affectation de vitesse aux pixels intérieurs va être utilisée dans les chapitres V et VI.

Chapitre IV

Le prétraitement des données

Chapitre IV

LE PRETRAITEMENT DES DONNEES

4.1 LE FILTRAGE DES VITESSES

Le signal vitesse, décrit dans le chapitre précédent, contient pour chaque piéton une série de vitesses correspondant aux pixels affectés par le mouvement. Ces vitesses, très erratiques, ne donnent pas une information globale sur le déplacement du piéton. Les situations suivantes représentent des cas où les vitesses obtenues sont très difficiles à utiliser :

- pour un même piéton qui passe sous les caméras, certaines vitesses sont très difficiles à interpréter car elles sont dans le sens opposé au sens de déplacement du piéton. Ces vitesses sont générées par le balancement des bras, d'avant en arrière, du piéton lors de son passage sous les caméras,
- le balancement des membres d'un piéton fait apparaître des vitesses beaucoup plus élevées que la vitesse réelle du déplacement du piéton,
- les piétons qui sont immobiles sous les plans de surveillance conduisent à des vitesses nulles.

Ces vitesses ne sont pas exploitables directement. Il est nécessaire d'effectuer un filtrage, qui élimine les vitesses non représentatives du déplacement réel du piéton afin de déterminer une vitesse associée au mouvement global du piéton.

4.1.1 Représentation du signal vitesse

Un exemple de représentation du signal vitesse $S(k)$ défini au § 3.5 est fourni sur la figure 4.1. L'abscisse représente le rang des pixels affectés par le mouvement d'un piéton. Les vitesses correspondantes sont représentées en ordonnée. Les vitesses sont représentées de manière algébrique, correspondant aux deux sens de déplacement sous les caméras.

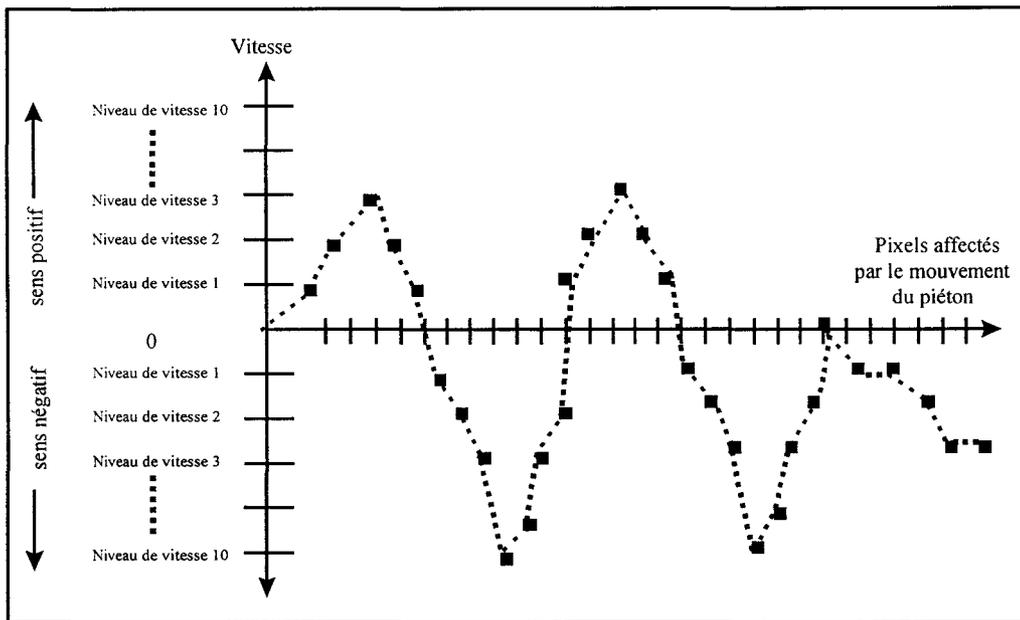


Figure 4.1 : signal vitesse brut

Quand un piéton traverse les plans de surveillance, on obtient autant de valeurs de vitesses qu'il y a de pixels affectés par le mouvement du piéton. Pour le filtrage de ces informations, il n'est pas nécessaire de quantifier très finement la vitesse. C'est pourquoi, nous procédons à une première discrétisation arbitraire des modules des vitesses brutes sur 10 niveaux. Dans un premier temps, ces 10 niveaux de vitesses sont suffisants. Sur la figure 4.1, les valeurs des vitesses sur l'axe des ordonnées sont discrétisées sur ces 10 niveaux dont les valeurs sont fournies dans le tableau 4.1.

Niveau de vitesse	Vitesse exprimée en nombre d'acquisitions	Vitesse exprimée en mètres par seconde
Niveau 1	1 à 2	10 à 5
Niveau 2	3 à 4	3,33 à 2,5
Niveau 3	5 à 6	2 à 1,67
Niveau 4	7 à 8	1,43 à 1,25
Niveau 5	9 à 10	1,11 à 1
Niveau 6	11 à 12	0,91 à 0,83
Niveau 7	13 à 14	0,77 à 0,71
Niveau 8	15 à 16	0,67 à 0,63
Niveau 9	17 à 18	0,59 à 0,56
Niveau 10	19 et plus	0,53 et moins

Tableau 4.1 : discrétisation préliminaire des vitesses

Le filtrage du signal vitesse est réalisé sur le signal discrétisé.

4.1.2 Choix du type de filtrage

Nous avons cherché dans la littérature un filtre qui permet, à partir de ces vitesses erratiques, d'obtenir une vitesse globale représentative du mouvement de chaque piéton. Il s'agit d'éliminer les vitesses qui ne sont pas représentatives du déplacement réel du piéton. Ces vitesses non représentatives correspondent à des "pics" et "vallées" dans le signal vitesse (Figure 4.1).

Il semble que l'approche morphologique soit bien adaptée au filtrage de ce type de signal. Nous rappelons les principes de base du filtrage morphologique.

4.1.3 Filtres morphologiques

La description de ce type de filtres a été largement développée et commentée dans d'autres écrits, aussi dans le cadre de cette thèse, seuls quelques rappels sont fournis [SER 82].

La morphologie mathématique consiste à extraire de l'information à partir du signal d'origine en choisissant tout d'abord un élément structurant qui est en fait un autre signal mais dont la forme est plus simple que celle du premier.

L'élément structurant agit sur le signal d'origine pour le transformer en une version plus simple, en éliminant les détails non significatifs pour ne conserver que les variations importantes du signal.

Filtrer morphologiquement une fonction consiste à procéder à une séquence d'opérations dont la "dilatation" et "l'érosion" représentent les opérations de base [SER 86].

Soit la fonction discrète $s_k : S(k) \rightarrow Z$, un signal 1-D et soit $b_k : B_k \rightarrow Z$, l'élément structurant centré sur le point k , où $S \subset Z$ et $B \subset Z$ sont les domaines de s et b respectivement.

L'érosion, notée $s \ominus b$, et la dilatation, notée $s \oplus b$, peuvent être exprimées sous la forme :

$$(s \ominus b)(k) = \min \left\{ s(k+x) - b_{k+x}(x) ; x \in B_{k+x} \text{ et } k+x \in S \right\}$$

$$(s \oplus b)(k) = \max \left\{ s(k-x) + b_{k-x}(x) ; x \in B_{k-x} \text{ et } k-x \in S \right\}$$

où b_{k+x} représente la translation de b_k par x , avec $x \in Z$.

Ces opérations sont basées sur l'addition et la soustraction de Minkowski [MIN 03]. En pratique, l'érosion et la dilatation sont souvent combinées. Elles conduisent ainsi à deux autres opérations morphologiques fondamentales appelées ouverture et fermeture. L'ouverture, notée $S \circ B$, est définie par une érosion suivie d'une dilatation et la fermeture, notée $S \bullet B$, est définie par une dilatation suivie d'une érosion.

$$(s \circ b)(k) \equiv ((s \ominus b) \oplus b)(k)$$

$$(s \bullet b)(k) \equiv ((s \oplus b) \ominus b)(k)$$

Ces deux opérations, l'ouverture puis la fermeture, sont appliquées au signal vitesse, en utilisant un élément structurant unidimensionnel de taille F .

4.1.3.1 Résultats obtenus avec le filtre morphologique

Les opérations d'ouverture et de fermeture ont été appliquées au signal de vitesse brut avec une fenêtre F de petite taille : de l'ordre de 5 à 10 pixels. Les résultats n'ont pas été satisfaisants en terme de vitesse globale de déplacement du piéton. En effet, sur le signal filtré, demeurent encore des composantes positives et négatives dans la vitesse du piéton.

Afin d'augmenter l'efficacité du filtrage pour obtenir une vitesse globale pour le piéton, d'autres opérations d'ouverture et de fermeture ont été effectuées avec des tailles de fenêtres plus importantes. Cependant, compte tenu du site où nos données ont été relevées, la taille maximale de fenêtre que l'on peut envisager est de 25 pixels. Cette taille correspond, dans ce site, approximativement à la largeur moyenne de la forme d'un piéton. Si la taille de la fenêtre est supérieure à 25 pixels, le filtrage peut porter sur les vitesses de plusieurs piétons en même

temps. Ceci peut conduire à affecter la même vitesse à des piétons se déplaçant à des vitesses différentes.

Sur la figure 4.2 sont représentés deux signaux de vitesse : le premier est un signal brut, constitué par les vitesses associées aux pixels sur la largeur du piéton, et le deuxième est le résultat du signal brut filtré par une ouverture suivie d'une fermeture avec une fenêtre de 25 pixels.

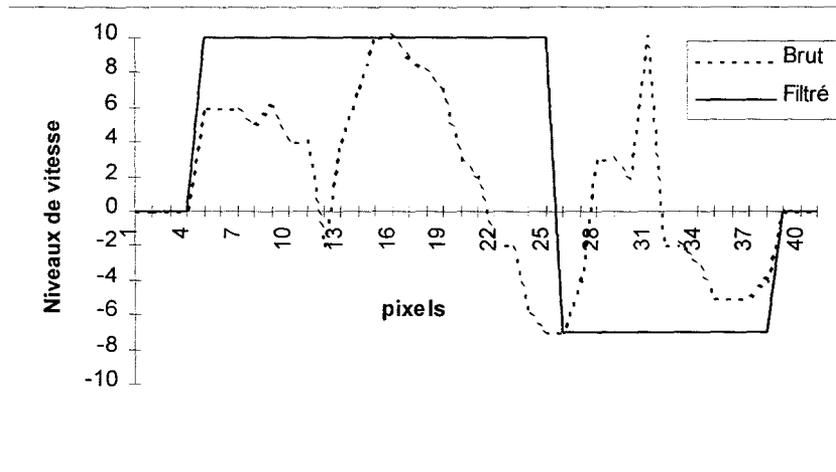


Figure 4.2 : résultat du filtrage morphologique

Nous pouvons remarquer, sur la figure 4.2, que les opérations morphologiques utilisées ont modifié le signal brut par l'élimination des pics et des vallées et la disparition des trous. Mais les variations du signal vitesse étant importantes et très fréquentes, les opérations morphologiques n'ont pu filtrer correctement les vitesses pour conduire à une vitesse globale pour le piéton. On observe sur le signal filtré, que la vitesse du piéton présente encore une composante positive entre les pixels de rang 4 et 25 et une composante négative du pixel de rang 26 à celui de rang 39.

Le problème rencontré dans l'utilisation des filtres morphologiques est que les valeurs des vitesses à filtrer dans la fenêtre de F pixels correspondant à la taille de l'élément structurant peuvent être positives et négatives. Les minima correspondant à l'érosion sont toujours négatifs et les maxima pour la dilatation sont toujours positifs.

Ces filtres sont adaptés pour des valeurs dont le domaine de validité est uniquement \mathbb{R}^+ ou \mathbb{R}^- . Ils sont bien adaptés pour le filtrage d'images en niveaux de gris.

Pour tenir compte des valeurs positives et négatives des vitesses, nous avons développé un filtre, appelé "**filtre conditionnel**", qui utilise les opérations d'ouverture et de fermeture en introduisant une condition sur le sens des vitesses.

4.1.4 Filtre conditionnel

Le signal à traiter et la taille de l'élément structurant sont les mêmes que ceux utilisés pour les filtres morphologiques. Les opérations de fermeture et d'ouverture sont appliquées au signal brut en introduisant une condition sur les calculs des maxima pour l'ouverture et des minima pour la fermeture.

4.1.4.1 Calcul des minima conditionnels

La condition que nous introduisons sur le calcul des minima dépend du sens des vitesses dans la fenêtre de F pixels correspondant à la taille de l'élément structurant. Deux cas peuvent se présenter (Figure 4.3) :

- si les vitesses de la fenêtre sont toutes de sens positif ou toutes de sens négatif, alors la **valeur filtrée** est égale au minimum des valeurs des vitesses dans cette fenêtre,
- si les vitesses calculées dans la fenêtre présentent les deux sens, deux cas peuvent se produire :
 - si plus de la moitié des vitesses sont de sens négatif, la **valeur filtrée** est égale au minimum des modules de vitesses dans la fenêtre,
 - sinon la **valeur filtrée** est égale à la valeur médiane des modules de vitesses de sens positif dans la fenêtre.

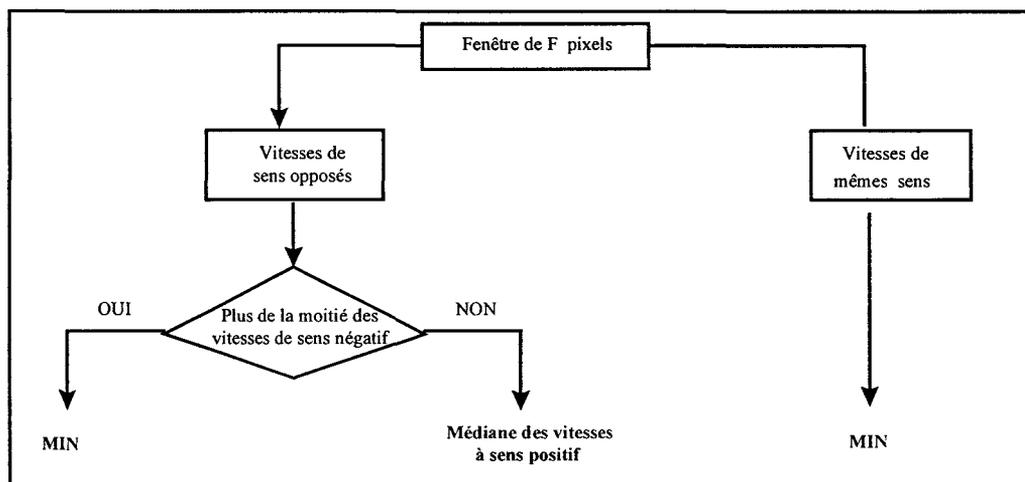


Figure 4.3 : calcul des minima conditionnels

4.1.4.2 Calcul des maxima conditionnels

La condition sur le calcul des maxima dépend, comme pour le calcul des minima, du sens des vitesses dans la fenêtre de calcul (Figure 4.4). Deux cas également se présentent :

- si les vitesses dans la fenêtre sont toutes de même sens, alors la **valeur filtrée** est égale au maximum des valeurs des vitesses dans la fenêtre,
- si les vitesses dans la fenêtre présentent les deux sens, deux cas peuvent se produire :
 - si plus de la moitié des vitesses sont de sens positif, la **valeur filtrée** est égale au maximum des modules de vitesses dans la fenêtre,
 - sinon, la **valeur filtrée** est égale à la médiane des modules des vitesses de sens négatif dans la fenêtre.

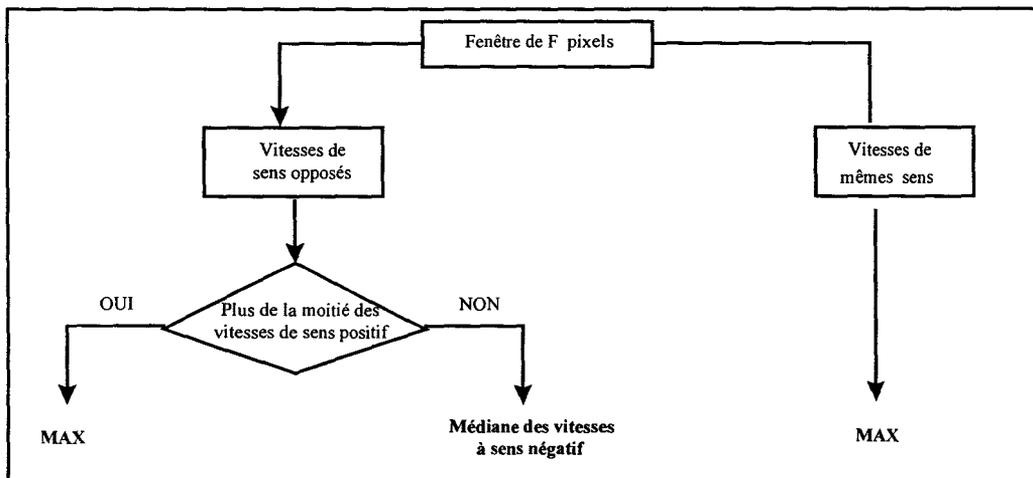


Figure 4.4 : calcul des maxima conditionnels

Finalement, que ce soit pour le calcul des minima ou des maxima, lorsque les vitesses sont toutes de même sens, le filtre conditionnel est exactement le même que le filtre morphologique testé auparavant. Par contre, lorsque ces vitesses présentent les deux sens sur la largeur de la fenêtre, l'introduction de cette condition sur la majorité des sens modifie notablement les résultats.

4.1.4.3 Résultats obtenus avec le filtre conditionnel

Sur la figure 4.5 sont représentés deux signaux de vitesse. L'un représente le signal brut, le même que celui utilisé pour le test du filtre morphologique; l'autre le signal filtré en appliquant le filtre conditionnel.

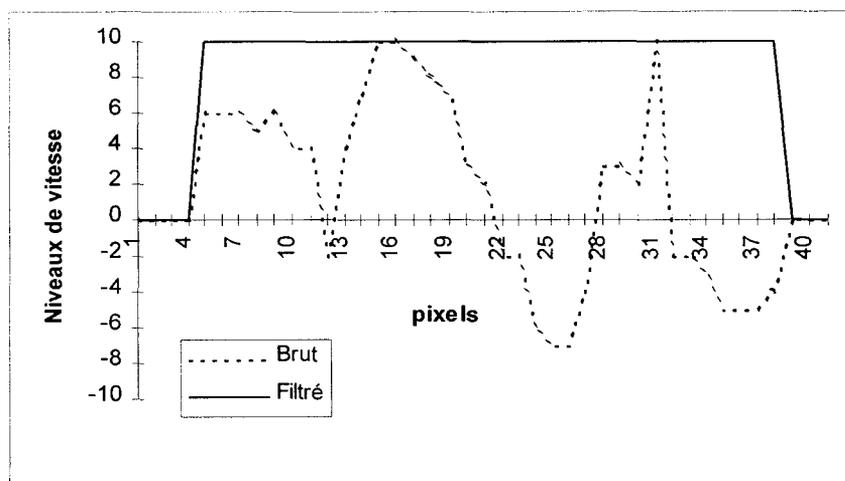


Figure 4.5 : résultat du filtre conditionnel

On peut noter maintenant que le signal d'origine est correctement filtré, selon le but recherché, car il conduit à une vitesse globale représentative du mouvement d'ensemble du piéton.

4.2 CLASSIFICATION FINALE DES VITESSES

Une fois toutes les vitesses brutes calculées, puis filtrées à l'aide du filtre conditionnel, leurs modules sont assignés à 5 classes, comme le montre le tableau 4.2.

Cette classification, n'est pas, contrairement à la discrétisation sur 10 niveaux, arbitraire. Ici toutes les vitesses calculées et filtrées sont prises en compte pour être classées par un algorithme de classification non hiérarchique du type "Nuées dynamiques" ou "Centres mobiles" [SAP 90].

Sachant que l'on désire classer un certain nombre de valeurs de vitesses, en un nombre c de classes, la méthode des centres mobiles permettant d'obtenir ces c classes est la suivante :

- définir au hasard, ou selon des considérations extérieures à la méthode, c noyaux (en général c vitesses parmi les valeurs de vitesses à classer),
- associer à chaque noyau l'ensemble des valeurs de vitesses les plus proches de la valeur de vitesse de ce noyau, selon une distance euclidienne. On obtient alors une partition en c classes,
- déterminer le barycentre de chaque classe, qui devient le nouveau noyau,
- itérer le processus, jusqu'à stabilisation du résultat de la classification.

On peut aussi utiliser un indice pour juger de la qualité d'une classification, tel que l'inertie inter-classes. L'inertie totale des valeurs de vitesses à classer se décompose en deux types d'inerties complémentaires :

- l'inertie intra-classe qui est la somme des distances de toutes les valeurs de vitesses assignées à une classe, au barycentre de cette classe,
- l'inertie inter-classes qui est la somme des distances de toutes les valeurs de vitesses des barycentres des différentes classes, au barycentre général.

On essaie en général de maximiser l'inertie inter-classes et de minimiser l'inertie intra-classes de façon à obtenir des classes aussi différenciées que possible entre elles, et très homogènes sur un plan interne. Nous avons effectué des classifications en 2, 3, 4, 5, 6, 7 classes en observant à chaque fois l'évolution des valeurs des inerties.

Les meilleures inerties ont été trouvées pour la classification en 5 classes. Il faut noter que les valeurs de vitesses négatives et positives ont été classées séparément et que la même classification en 5 classes a été retenue. Sur le tableau 4.2 figure les résultats de cette classification en 5 classes.

En considérant les deux sens pour les vitesses et les vitesses nulles, cela fait 11 classes.

Cette classification finale est utilisée dans tous les traitements de comptage développés dans les chapitres V et VI où chaque piéton possède une vitesse globale qui appartient à une des classes de vitesse ci-dessous. Dans certains cas, cependant, nous verrons que lorsque le signal vitesse du piéton est très erratique, deux classes de vitesse différentes peuvent lui être affectées. Dans ce cas, on constate expérimentalement que ces deux classes sont adjacentes.

Classe de vitesse	Vitesse exprimée en nombre d'acquisitions	Vitesse exprimée en mètres par seconde
Classe 1	1 à 4	10 à 2,5
Classe 2	5 à 8	2 à 1,25
Classe 3	9 à 12	1,11 à 0,83
Classe 4	13 à 16	0,77 à 0,63
Classe 5	17 et plus	0,59 et moins

Tableau 4.2 : classification finale des vitesses

Le prétraitement suivant, qui est aussi un préalable aux algorithmes de comptage présentés dans les chapitres V et VI, consiste à extraire l'intersection de deux images binaires.

4.3 SELECTION DE L'INTERSECTION DES FORMES

Nous avons vu, dans les chapitres précédents, que les images binaires obtenues grâce au dispositif de comptage contiennent des formes ou des silhouettes correspondant aux passages de piétons sous les caméras.

Pendant les tranches horaires creuses, les piétons passent sous les caméras de manière isolée et les traces qu'ils laissent sont en général des silhouettes.

Lorsque la densité de piétons devient élevée, pendant des tranches horaires de pointe par exemple, les traces laissées par les piétons sont presque toutes des formes car les piétons passent de manière groupée sous les caméras et leurs silhouettes se chevauchent.

Le but de cette recherche est de développer un système temps réel de comptage de piétons, dont la précision soit satisfaisante, quelle que soit la densité de personnes. Il s'agit d'associer le bon nombre de piétons aux traces enregistrées par les caméras.

Pour cela, nous avons développé des algorithmes, présentés dans les chapitres V et VI, qui consistent à traiter les formes pour mettre en évidence les silhouettes qui les composent.

Chaque piéton qui traverse le double plan de surveillance laisse une trace dans les séquences acquises par chaque caméra. Pour les traitements qui sont développés dans les chapitres V et VI, nous utilisons une image qui résulte de l'intersection des deux traces laissées par les piétons. Dans certains cas simples, cette image intersection permet, à elle seule, d'individualiser les silhouettes qui constituent les formes.

La figure 4.6 est divisée en trois parties : à gauche une forme correspondant à la trace associée à la caméra 1, à droite celle associée à la caméra 2, et au milieu la réunion et l'intersection des deux formes précédentes obtenues par superposition. La réunion est représentée en gris alors que l'intersection correspond à la partie blanche. Nous pouvons constater que l'intersection des deux formes initiales met en évidence deux silhouettes correspondant probablement à deux piétons.

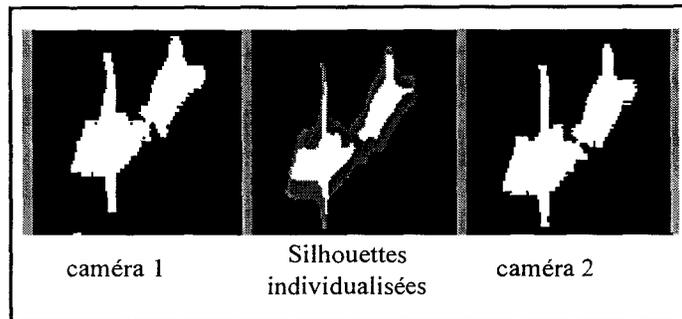


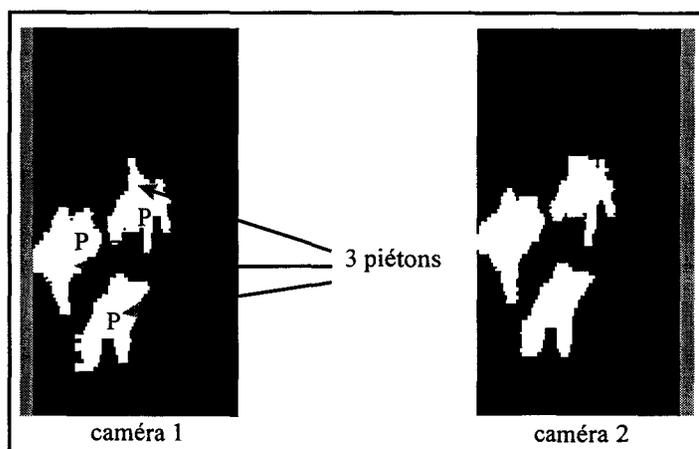
Figure 4.6 : prise en compte de l'intersection

Les algorithmes décrits dans les chapitres suivants seront testés sur des séquences d'images sélectionnées comme représentatives des principales situations rencontrées dans une station de métro en exploitation. Ces séquences, que nous retrouverons tout au long de ce mémoire, ont été relevées sur un site réel et présentent plusieurs niveaux de difficulté dans le traitement.

4.4 SEQUENCES D'IMAGES UTILISEES POUR LE TEST DES ALGORITHMES

4.4.1 Séquence No 1

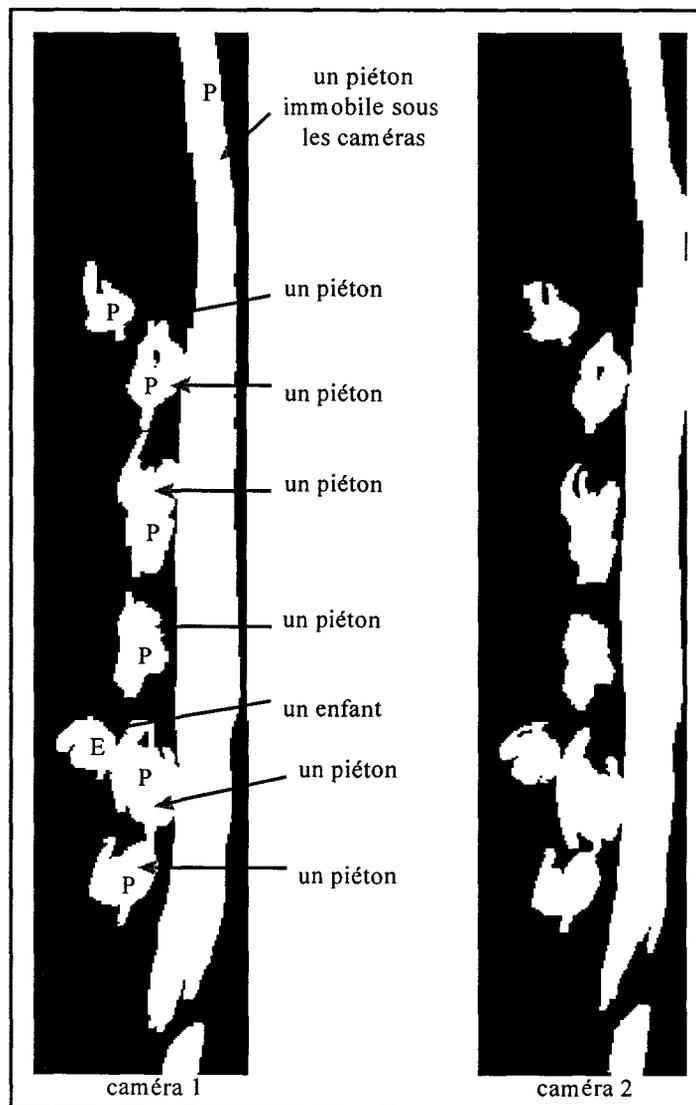
La première séquence, reproduite ci-dessous, représente le cas simple de piétons nettement séparés passant sous les caméras. Les formes correspondant aux piétons sont représentées dans les deux diagrammes spatio-temporels associés aux deux caméras. Dans cette séquence, les traces laissées par les piétons après leur passage sont des silhouettes. Les lettres sur les formes renseignent sur le type d'usagers ou d'objets traversant les plans de surveillance.



Séquence 1 : piétons isolés (P : piétons)

4.4.2 Séquence No 2

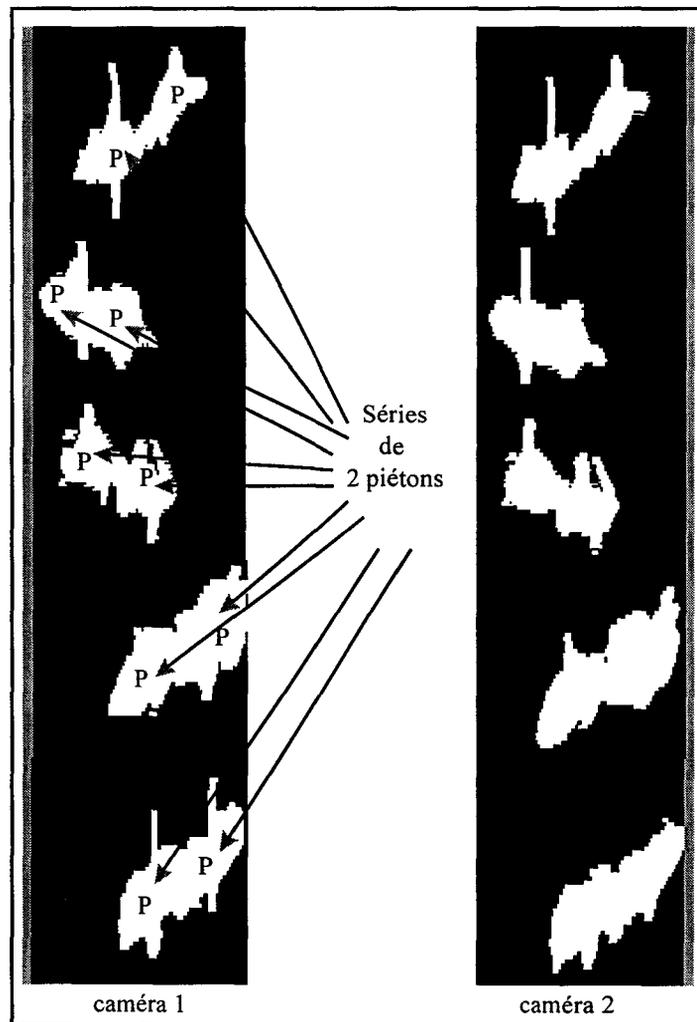
La deuxième séquence présente une première difficulté dans le traitement des formes : un piéton reste immobile sous les caméras, générant ainsi une forme très allongée, apparaissant ici sur les parties droites des deux diagrammes spatio-temporels. Pendant l'immobilité de ce piéton, d'autres traversent les plans de surveillance : ils sont représentés par les autres formes. On remarque une certaine connexité entre les silhouettes des piétons.



Séquence 2 : piéton immobile (P : piéton, E : enfant)

4.4.3 Séquence No 3

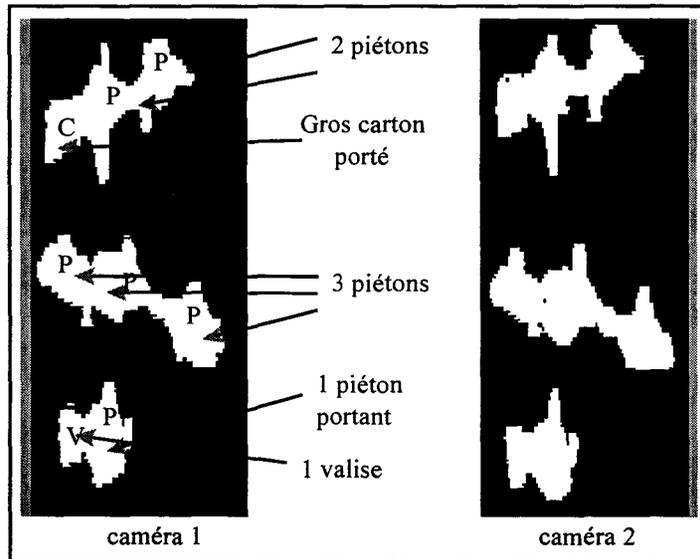
La séquence 3 est une sélection où cinq couples de deux piétons passent sous les caméras de manière très rapprochée en se donnant la main ou serrés l'une contre l'autre, conduisant ainsi à des silhouettes qui se chevauchent.



Séquence 3 : série de deux silhouettes connexes (P : piéton)

4.4.4 Séquence No 4

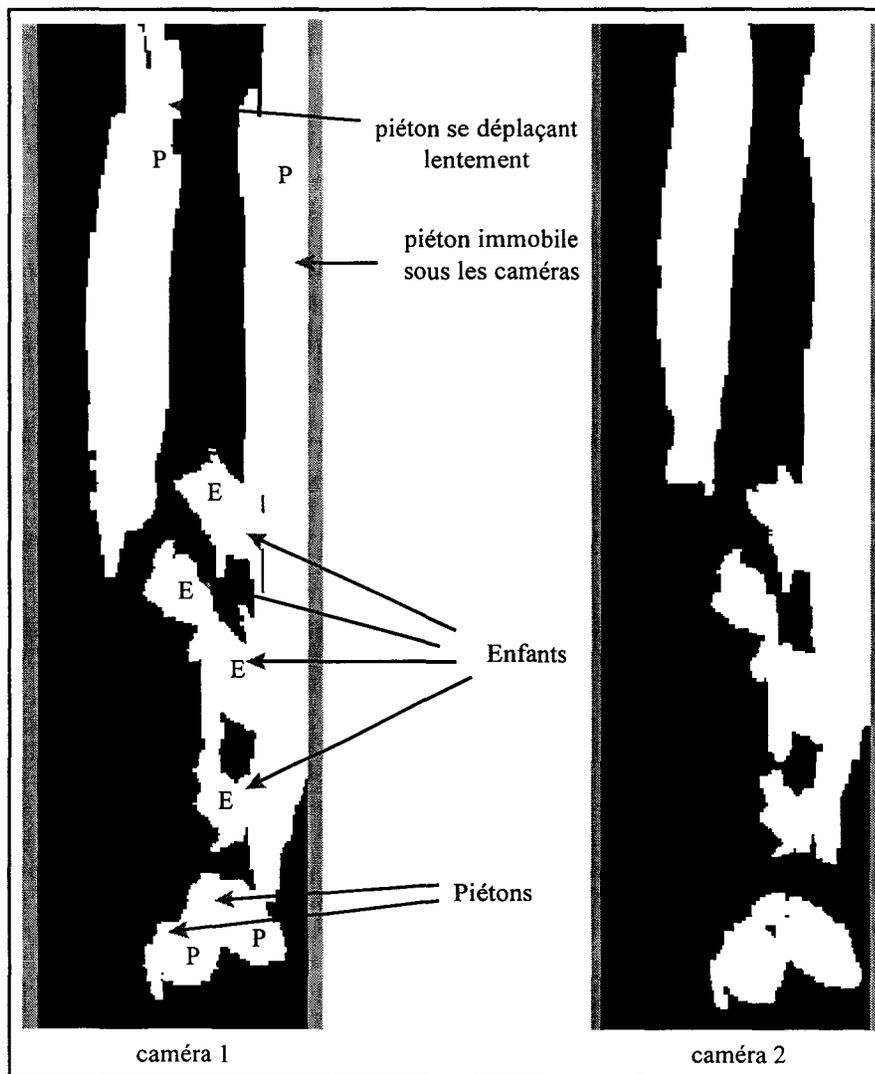
La séquence 4 illustre une configuration où les piétons ont généré des silhouettes connexes et/ou transportent des objets tels que des valises ou des cartons.



Séquence 4 : piétons et objets (P : piéton, V : valise, C : carton)

4.4.5 Séquence No 5

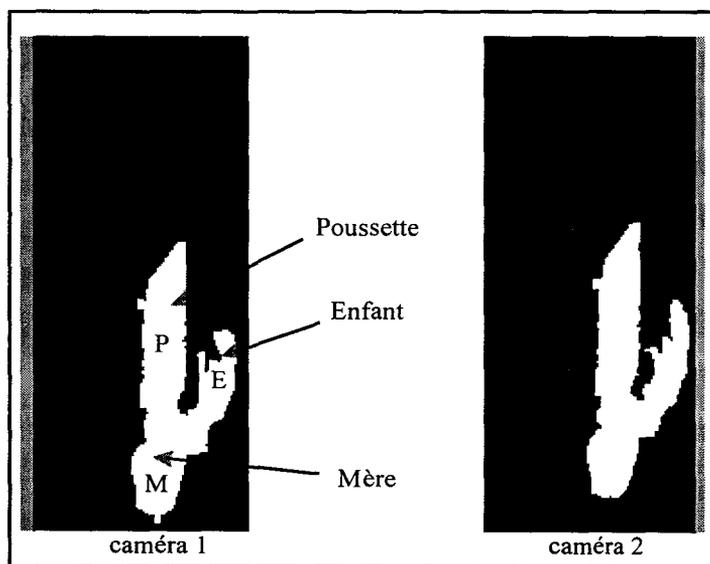
La séquence 5 présente la même difficulté que la séquence 2, résultant d'un piéton immobile sous les caméras avec d'autres piétons passant à proximité, mais avec un niveau supérieur de connexité entre les silhouettes. Sur les diagrammes spatio-temporels on note, à proximité de la forme du piéton immobile, un ensemble de silhouettes connexes de largeur réduite. Ces silhouettes correspondent à des enfants.



Séquence 5 : groupe d'enfants (P: piéton, E: enfant)

4.4.6 Séquence No 6

La dernière séquence présentée, très atypique, représente le passage sous le dispositif d'une personne poussant un landau tout en donnant la main à un enfant.



Séquence 6 : piétons et objets particuliers (*M* : mère, *P* : poussette, *E* : enfant)

4.5 CONCLUSION

Nous avons vu, dans ce chapitre, qu'un certain nombre de prétraitements sur les données de base sont indispensables avant de procéder au comptage des piétons proprement dit.

Le filtrage des vitesses est un élément très important de la procédure, car tous les traitements qui suivent se basent sur les résultats de ce filtrage. Le but de ce filtrage est d'obtenir une vitesse unique représentative du mouvement d'un piéton. Un filtrage morphologique classique a été utilisé. Celui-ci conduit à de mauvais résultats, notamment lorsque le signal filtré présente des composantes de vitesse négatives et positives.

Nous avons donc développé un nouveau filtre, appelé filtre conditionnel, qui tient compte du signe du signal de vitesse brut. Ce filtre donne des résultats satisfaisants dans la majorité des signaux de vitesse rencontrés. Cependant, des cas existent où après filtrage conditionnel, la classe de vitesse affectée à un piéton n'est pas unique. On constate malgré tout que ces classes sont adjacentes.

La classification finale des vitesses est aussi un élément important, car il intervient directement dans les procédures de comptage développées dans le chapitre suivant.

Un prétraitement également important consiste à ne considérer que les séquences d'images résultat de l'intersection des images binaires captées par les deux caméras. Dans certains cas simples, l'image intersection individualise à elle seule des silhouettes qui se chevauchent.

En fin de chapitre, nous avons présenté six séquences d'images sur lesquelles les algorithmes de comptage seront évalués. Ces séquences représentent des niveaux de complexité différents en termes de connexité des silhouettes.

Nous proposons maintenant d'aborder dans les chapitres V et VI les algorithmes de comptage.

Chapitre V

Les algorithmes de comptage

Chapitre V

LES ALGORITHMES DE COMPTAGE

5.1 INTRODUCTION

Les 3 informations prises en compte pour accomplir le comptage sont :

- l'information spatiale extraite de chaque image ligne. Elle concerne essentiellement le profil des niveaux de gris dans le sens de la largeur du site de mesures contrôlé par les caméras,
- l'information temporelle correspond à l'information qui peut être extraite de la séquence des images lignes acquises. Elle concerne essentiellement le diagramme spatio-temporel obtenu en empilant les images lignes,
- l'information de vitesse, étroitement liée à l'information temporelle, permet d'estimer la vitesse des piétons lors de leur passage sous les caméras.

Les données que nous sommes amenés à traiter, et qui correspondent à des passages de piétons sous les caméras, peuvent être classées en deux catégories.

Lorsque les piétons passent sous les caméras isolément, ce qui correspond à des périodes de faible fréquentation, les traces recueillies sont des silhouettes (Figure 5.1).

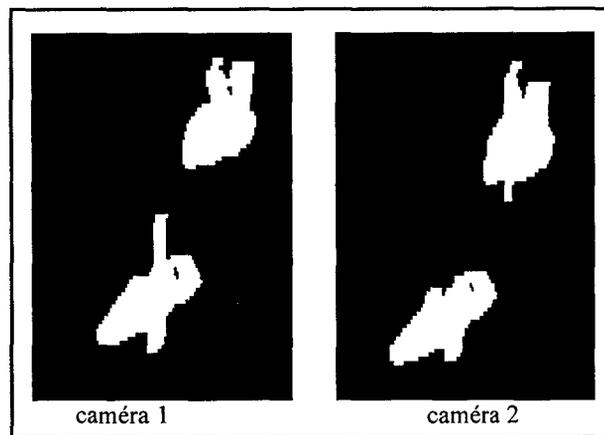


Figure 5.1 : piétons isolés

Dans ce cas, le comptage des piétons est relativement aisé. Un algorithme, travaillant sur les pixels constituant la silhouette et vérifiant leur connexité, permettrait dans l'image binaire de détecter le contour de chaque silhouette. Dans la figure 5.1, en associant un piéton à chaque silhouette, nous compterions 2 piétons.

En revanche, pendant une période où le flot de piétons est dense, les piétons passent de manière groupée sous les caméras. Un exemple de formes obtenues dans ce cas est donné par la figure 5.2.

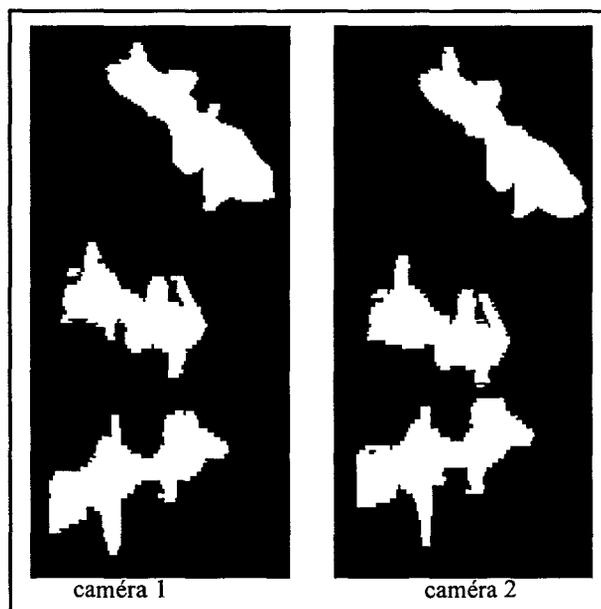


Figure 5.2 : piétons groupés

Sur la figure 5.2, les silhouettes des piétons sont maintenant connexes et il est difficile de voir immédiatement si les formes obtenues correspondent à un ou plusieurs piétons.

Dans ce cas, l'algorithme de comptage par l'analyse de la connexité des pixels, évoqué ci-dessus, fournirait de très mauvais résultats puisqu'il associerait un piéton à chaque forme, quelle que soit la configuration de la forme. Dans cet exemple précis, l'algorithme d'analyse de connexité compterait 3 piétons au lieu de 6 en réalité. Il serait donc extrêmement hasardeux de procéder directement au comptage.

A cela il faut ajouter que les piétons qui passent sous les caméras ne sont pas toujours seuls. Ils peuvent être accompagnés d'enfants dont les silhouettes sont moins larges, certains portent des sacs ou des valises, poussent des landaus, portent des parapluies s'il fait un temps pluvieux et que le système est installé à l'extérieur. Finalement, les formes obtenues peuvent correspondre à une multitude de possibilités.

5.1.1 Choix de la séparation des formes

Nous savons, grâce à des observations manuelles effectuées sur des images vidéo, que toutes les formes présentes sur la figure 5.2 correspondent à plusieurs piétons dont les silhouettes correspondantes sont connexes. Pour mettre en évidence ces silhouettes nous avons développé des méthodes qui consistent à travailler sur les formes obtenues pour individualiser les silhouettes qui les composent. C'est un travail de "séparation" qui consiste à travailler sur la géométrie des formes.

Les algorithmes développés et décrits dans le cadre de ce chapitre s'inspirent de méthodes basées sur la morphologie mathématique. Les principes de base de la morphologie mathématique binaire ont été repris, élargis et adaptés à notre problématique. Trois algorithmes ont ainsi été développés, utilisant des éléments structurants de géométries différentes.

Nous savons que, lorsque les piétons passent de manière groupée sous les caméras, leurs silhouettes correspondantes sont connexes. Ce que nous recherchons, en nous inspirant de la morphologie mathématique, ce sont les "zones centrales" ou "parties centrales" des formes, appelées "noyaux". Après traitement, le résultat peut conduire à une mise en évidence de un ou plusieurs noyaux au sein d'une forme.

5.2 PROCEDURE DE SEPARATION DES FORMES

Nous avons dit, en introduction de ce chapitre, que nous cherchons à traiter les formes pour mettre en évidence les silhouettes qui les composent. Le résultat attendu du traitement des formes est un ensemble de noyaux dont chacun correspond à un piéton.

Avant de mettre en œuvre les algorithmes basés sur la morphologie en vue de la séparation des formes, nous avons développé un algorithme qui permet, dans certaines situations, de séparer directement les formes sans utiliser les méthodes qui s'inspirent de la morphologie.

5.2.1 Procédure de séparation par la vitesse

C'est l'information de vitesse qui va permettre de procéder à la séparation directe des formes. Nous rappelons qu'avant cette procédure, au chapitre III, on a affecté une vitesse à tous les pixels dans les formes, qu'ils soient des pixels de transition ou des pixels intérieurs. Toutes ces vitesses, contenues dans une base de données, ont été calculées, filtrées et assignées à 5 classes, 11 classes si l'on considère les deux sens des vitesses et la vitesse nulle.

La procédure de séparation est la suivante. Pour chaque ligne d'acquisition, les vitesses associées à deux pixels adjacents sont comparées deux à deux (Figure 5.3) et la séparation, c'est à dire la "coupure" de la forme est effectuée dans les cas suivants :

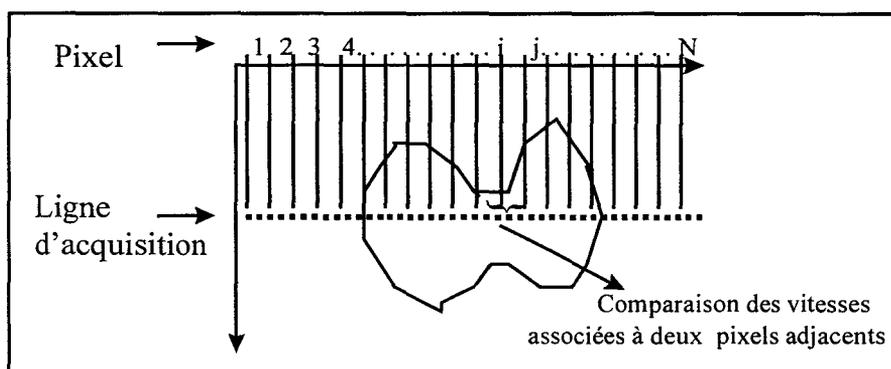


Figure 5.3 : principe de séparation

- la comparaison des vitesses associées à deux pixels adjacents révèle qu'elles ont des sens opposés, c'est à dire qu'elles correspondent à des déplacements de sens contraires,

- la comparaison des vitesses associées à deux pixels adjacents révèle que la vitesse associée à l'un des deux pixels est nulle, correspondant probablement à une personne arrêtée sous les caméras,
- la comparaison des vitesses associées à deux pixels adjacents révèle qu'ils ont des vitesses de même sens, mais appartiennent à des classes non adjacentes. En effet, après filtrage et classification finale, la vitesse d'un piéton appartient soit à une classe unique ou à deux classes adjacentes (Cf. § 4.2).

Si le test entre les pixels adjacents conduit à un diagnostic de séparation, la forme analysée est "coupée", c'est à dire que le premier pixel (on se déplace de gauche à droite) des deux pixels adjacents testés est mis à 1 sur toute la longueur de la forme analysée.

Les figures 5.4, 5.5 et 5.6 illustrent de manière détaillée les trois cas de séparation directe que nous venons de présenter.

La figure 5.4.a présente le cas d'une forme constituée de deux silhouettes : la silhouette très longue correspond à un piéton arrêté sous les caméras, la plus courte à un autre piéton frôlant le premier. Le calcul des vitesses associées aux pixels de la forme révèle que certaines sont nulles. Celles qui sont nulles correspondent à des pixels de la silhouette du piéton statique. La séparation de la forme se fait alors sur le premier pixel associé à une vitesse nulle (Figure 5.4.b). Après séparation, la forme d'origine est divisée en deux silhouettes (Figure 5.4.c).

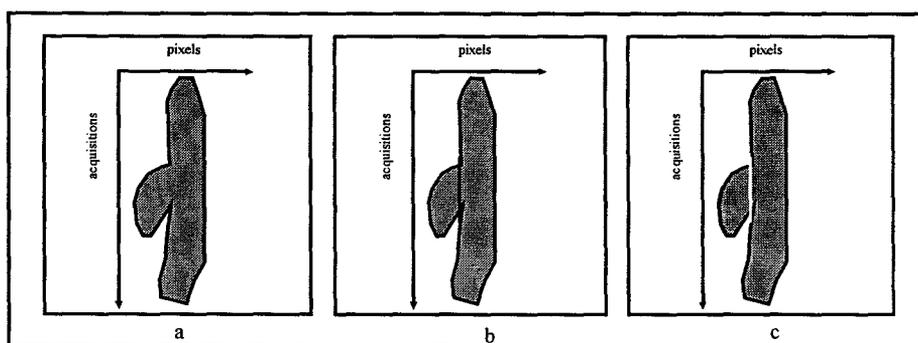


Figure 5.4 : premier scénario de séparation directe

La figure 5.5 présente une forme avec deux silhouettes connexes, correspondant à deux piétons se déplaçant dans le même sens (Figure 5.5.a). Le calcul des vitesses associées aux couples de pixels adjacents révèle que ces vitesses appartiennent à deux classes non

adjacentes et ne peuvent donc pas correspondre à un seul piéton. La forme est alors coupée au premier des deux pixels (Figure 5.5.b) pour donner deux silhouettes (Figure 5.5.c).

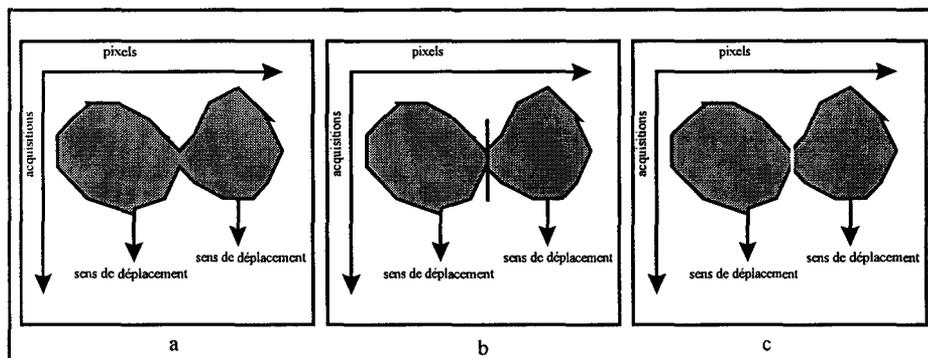


Figure 5.5 : deuxième scénario de séparation directe

La figure 5.6 présente une forme avec deux silhouettes connexes et correspondant à des piétons se déplaçant en sens inverse (Figure 5.6.a). Le calcul des vitesses associées aux couples de pixels adjacents révèle que ces vitesses ont des sens opposés. La forme est alors coupée au premier des deux pixels (Figure 5.6.b) pour donner deux silhouettes (Figure 5.6.c).

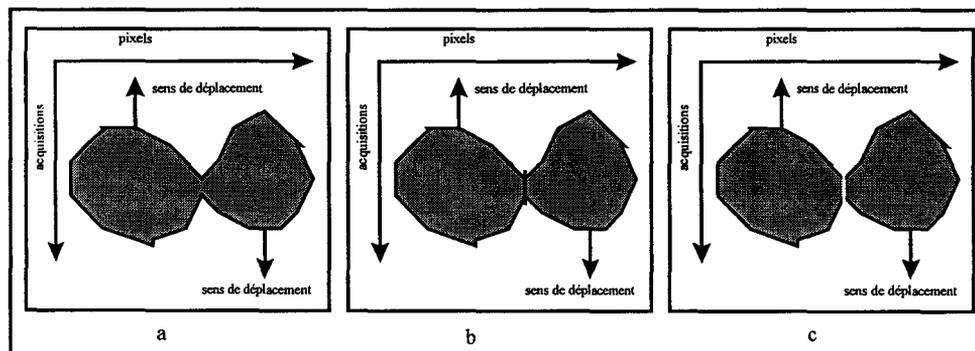


Figure 5.6 : troisième scénario de séparation directe

La figure 5.7, divisée en trois parties, présente une séquence réelle où l'on voit une forme constituée apparemment de plus d'une silhouette. La partie très longue de la forme correspond probablement à la silhouette d'un piéton immobile sous les caméras alors que les autres silhouettes correspondent à des piétons passant à proximité de ce piéton immobile. Les figures 5.7.a et 5.7.c représentent les formes obtenues par les deux caméras. La figure 5.7.b représente l'intersection des deux formes captées par les deux caméras. Nous pouvons remarquer qu'une séparation directe des formes a pu être opérée.

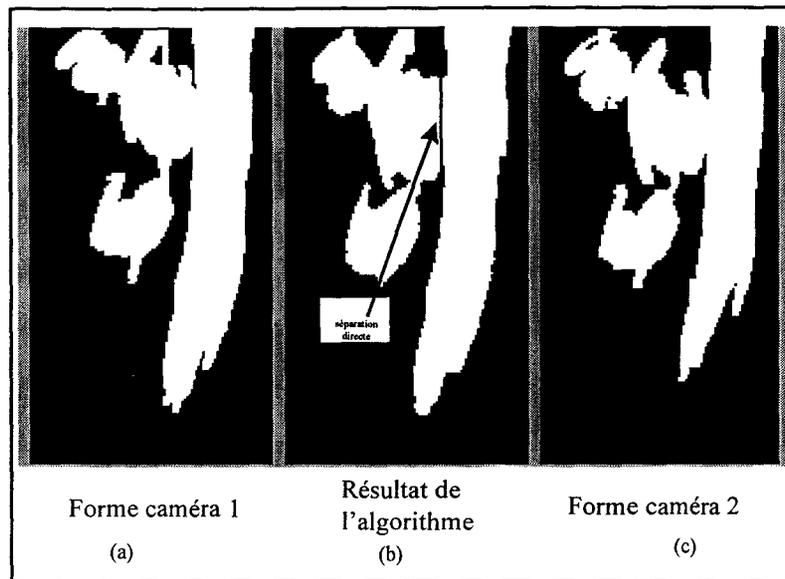


Figure 5.7 : exemple réel de séparation directe

Nous voyons que cette procédure permet une séparation directe des formes. Bien évidemment, elle ne permet pas de couvrir tous les cas de figure, une grande partie des formes n'étant pas séparable directement par analyse de la vitesse. C'est pourquoi nous avons développé des algorithmes supplémentaires que nous présentons au § 5.4 pour compléter l'algorithme de séparation directe. Ceux-ci s'inspirent des principes de base de la morphologie mathématique dont un bref rappel est fourni au § 5.3.

5.3 PRINCIPES DE BASE DE LA MORPHOLOGIE MATHÉMATIQUE

L'idée de base de la morphologie mathématique est de comparer les objets que l'on veut analyser à un autre objet de forme connue, appelé **élément structurant**.

La morphologie mathématique utilise des opérations ensemblistes pour transformer l'image. Ces opérations peuvent être classées en deux catégories : les transformations ensemblistes classiques et les transformations en tout ou rien utilisant un élément structurant.

Dans les paragraphes suivants, nous présentons uniquement les transformations de base en tout ou rien.

5.3.1 Transformations en tout ou rien par un élément structurant

Considérons un objet uni-, bi- ou tridimensionnel et appelons X l'ensemble des points de l'espace constituant cet objet. La structure de X est donnée par l'ensemble de toutes les

relations entre ces différents points. Les transformations en tout ou rien constituent un moyen systématique pour étudier ces relations.

Pour faire une transformation en tout ou rien, nous devons choisir un élément B de géométrie connue, appelé élément structurant. Cet élément structurant est ensuite déplacé de façon à ce que son origine passe par toutes les positions de l'espace. Pour chaque position, on pose une question relative à l'union, à l'intersection ou à l'inclusion de B avec ou dans X . La réponse sera positive ou négative, d'où le nom de **transformation en tout ou rien**. L'ensemble des points correspondants à des réponses positives forme un nouvel ensemble qui constitue ce que l'on appelle l'**image transformée**.

Chaque forme et chaque taille d'élément structurant donne une information relative à la géométrie et à la topologie de l'ensemble étudié. Chaque élément structurant différent donne un résultat différent. Le choix de l'élément dépend du problème à résoudre.

5.3.2 Transformations morphologiques élémentaires

Pour définir les différentes transformations, on va faire intervenir l'élément structurant noté B ou B_x pour indiquer qu'il est centré sur x .

5.3.2.1 Transformation par érosion

Pour définir l'opération d'érosion, nous allons nous situer dans un espace R^2 partiellement occupé par un ensemble X . Prenons un élément structurant B représentant une figure géométrique simple, par exemple une ellipse. Cet élément B_x est repéré par son centre qui est placé en x dans l'espace R^2 (Figure 5.8).

Il est ensuite déplacé de telle sorte que son centre occupe successivement toutes les positions x de l'espace.

Pour chaque position, on pose la question suivante : est-ce que B_x est complètement inclus dans X ?, c'est à dire en notation ensembliste :

$$B_x \subset X ?$$

L'ensemble des positions x correspondant à une réponse positive forme un nouvel ensemble Y appelé érodé de X par B_x . Cet ensemble satisfait l'équation :

$$Y = \{ x : B_x \subset X \}$$

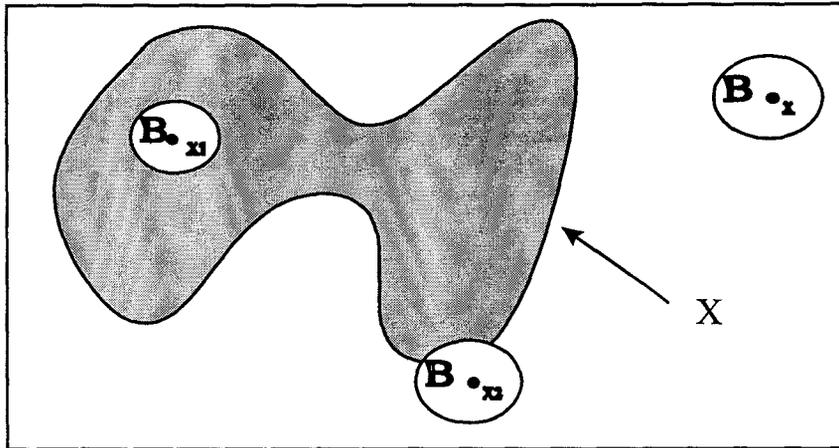


Figure 5.8 : principe de la transformation par érosion avec un élément structurant B_x centré en x : B_{x1} est inclus dans X , tandis que B_{x2} ne l'est pas.

L'érosion est habituellement représentée par le symbole \ominus .

La figure 5.9 présente un exemple d'érosion, par un élément structurant carré, d'un ensemble X .

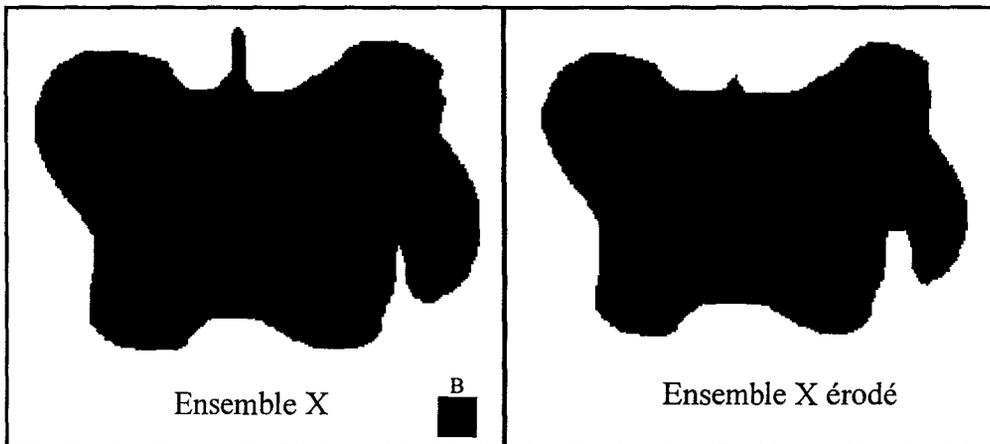


Figure 5.9 : érosion d'un ensemble X par un élément structurant carré

5.3.2.2 Transformation par dilatation

L'opération de dilatation se définit d'une manière analogue. En prenant le même élément structurant B_x , on pose, pour chaque point de R^2 , la question : B_x touche-t-il l'ensemble X ?, c'est à dire :

$$B_x \cap X \neq \emptyset?$$

L'ensemble des points de \mathbb{R}^2 correspondant aux réponses positives forme un nouvel ensemble Y dont la frontière est le lieu des centres géométriques de B_x lorsque B_x touche X . L'ensemble Y satisfait à l'équation :

$$Y = \{x : B_x \cap X \neq \emptyset\}$$

La dilatation est habituellement représentée par le symbole \oplus .

La figure 5.10 présente la dilatation de l'ensemble X de la figure 5.9, par l'élément structurant carré utilisé pour l'érosion de la figure 5.9.

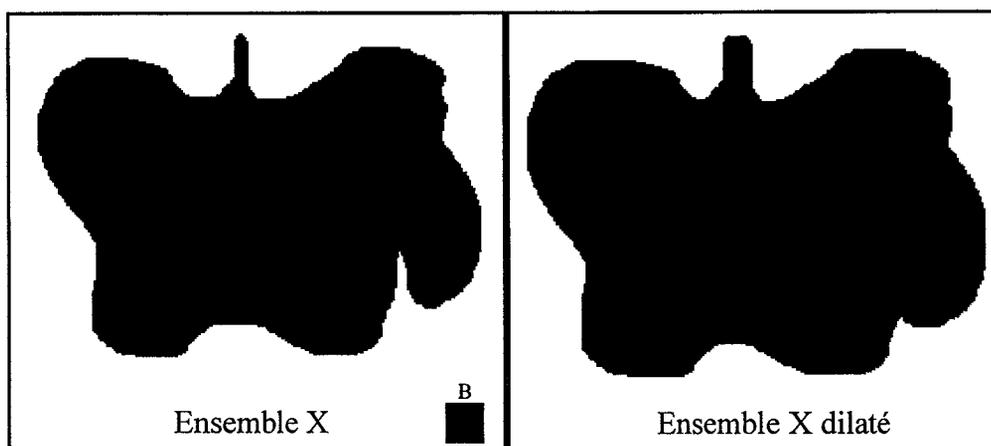


Figure 5.10 : dilatation d'un ensemble X par un élément structurant carré

5.3.2.3 Ouverture et fermeture morphologiques

Ouverture

Puisque l'érosion et la dilatation sont des transformations que l'on peut itérer, il est possible d'effectuer, sur un ensemble X , une érosion par B puis de dilater l'ensemble érodé par le même élément structurant. Cet ensemble est dénommé ouvert de X par B (noté $X \text{ ov } B$).

$$X \text{ ov } B = (X \ominus \tilde{B}) \oplus B$$

avec \tilde{B} symétrique de l'élément structurant B par rapport au centre.

La figure 5.11 présente l'ouverture de l'ensemble X de la figure 5.9 par l'élément structurant carré utilisé précédemment.

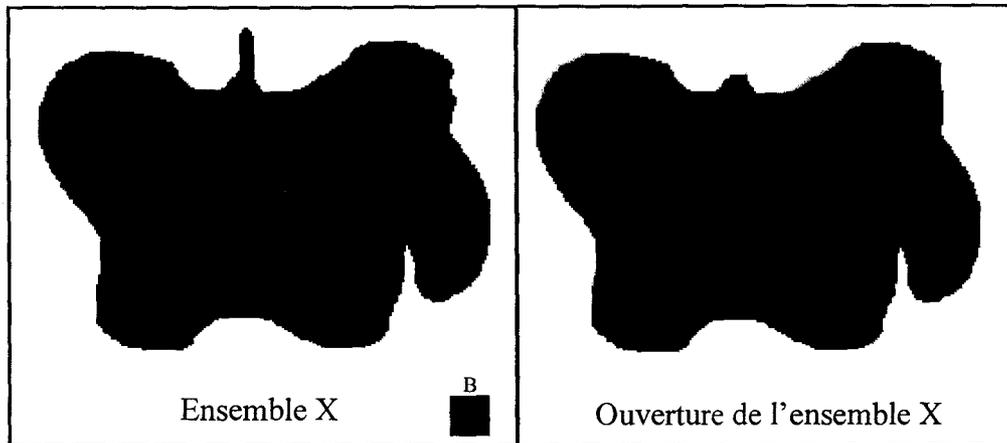


Figure 5.11 : opération d'ouverture sur un ensemble X avec un élément structurant carré

Fermeture

L'opération de fermeture est une dilatation suivie d'une érosion. C'est l'opération duale de l'opération d'ouverture.

$$(X \text{ ov } B)^c = X^c \text{ fer } B$$

$$X^c \text{ ov } B = (X \text{ fer } B)^c$$

où X^c est le complémentaire de X .

5.4 TRAITEMENTS INSPIRÉS DE LA MORPHOLOGIE MATHÉMATIQUE AVEC UN ÉLÉMENT STRUCTURANT DÉFORMABLE DYNAMIQUE

Habituellement, les opérations morphologiques telles que l'érosion et la dilatation sont utilisées avec un même élément structurant appliqué sur toute l'image à traiter. A chaque passe, le même élément structurant est utilisé. La géométrie de ce dernier, même si elle est invariante dans l'espace, peut-être variée : on trouve ainsi dans la littérature des éléments structurants ronds, rectangulaires, etc. Les opérations telles que la dilatation et l'érosion avec un élément structurant de géométrie unique, et invariant dans l'espace, ne peuvent être efficaces que sur des formes fixes.

La problématique du comptage est très différente dans la mesure où les formes que nous devons traiter sont variables aussi bien dans l'espace que dans le temps. Nous avons vu dans le

chapitre III que la variété des piétons et des objets détectés est telle que l'on n'a jamais deux fois la même silhouette. De plus, on a vu aussi que la taille de la silhouette à analyser peut varier dans les deux dimensions :

- en largeur car les personnes détectées n'ont pas le même gabarit : un enfant génère une silhouette moins large que celle d'un adulte,
- en longueur, en fonction de la vitesse de déplacement des personnes. Plus la vitesse est lente et plus la longueur de la silhouette est importante.

Ainsi, nous serons amenés à observer des personnes correspondant à des gabarits plus ou moins importants (personnes fortes ou non, chargées de bagages ou non), allant plus ou moins vite, et cela conduit à une multitude de silhouettes différentes.

Pour ces raisons, il n'est pas possible d'utiliser une opération morphologique avec un élément structurant unique. Il est donc nécessaire d'adapter l'élément structurant à la forme sur laquelle il est appliqué.

C'est pourquoi nous traitons les formes, en vue de leur séparation, avec une série d'éléments structurants différents pour la même passe. Nous utilisons pour cela des éléments structurants dont la taille est adaptée à la vitesse des piétons observés. Il en résulte que les axiomes de la morphologie mathématique ne sont plus vérifiés. C'est pourquoi, dans tout ce mémoire, on ne parle pas de méthodes morphologiques mais de méthodes *inspirées* de la morphologie mathématique.

5.4.1 Élément structurant rectangulaire dynamique

Le premier élément structurant que nous avons utilisé est rectangulaire [KHO 96/1]. Ses deux dimensions sont nommées :

- largeur, dans la dimension spatiale,
- longueur, dans la dimension temporelle.

Cet élément est illustré sur la figure 5.12.

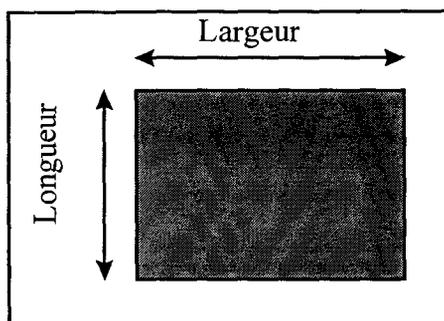


Figure 5.12 : dimensions de l'élément structurant

5.4.1.1 Utilisation de la vitesse

Les vitesses observées servent à dimensionner l'élément structurant. Ces vitesses sont, rappelons-le, filtrées par le filtre conditionnel présenté au chapitre IV, puis assignées à 5 classes.

Dans une forme donnée, à chaque pixel correspond une vitesse calculée, filtrée et associée à une classe (Chapitre IV). A chaque classe de vitesse est associée un élément structurant adapté, que nous décrirons plus loin. Par conséquent, à chaque pixel est associé un élément structurant adapté.

La procédure est alors la suivante : chaque pixel d'une forme dans l'image binaire est considéré comme le centre de l'élément structurant qui lui est appliqué. Si cet élément est complètement inclus, au sens ensembliste, dans la forme traitée, le pixel concerné est **marqué**. L'ensemble des pixels marqués forme une nouvelle image binaire (Figure 5.13).

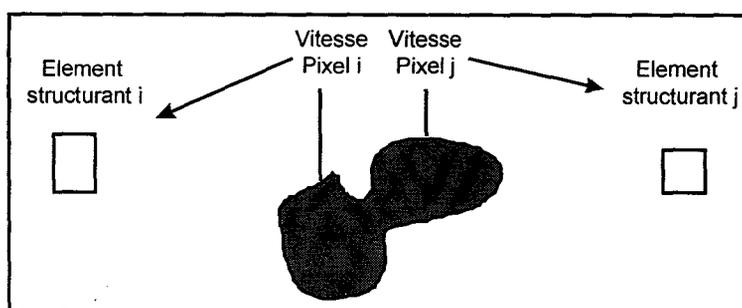


Figure 5.13 : traitement avec une série d'éléments structurants

5.4.1.2 Largeur de l'élément structurant

Une analyse statistique sur toutes les largeurs des silhouettes contenues dans notre base de données a été effectuée afin de déterminer la largeur moyenne.

La largeur de l'élément structurant est alors définie comme étant un sous-multiple de cette largeur moyenne calculée. Pour les données que nous avons traitées, la largeur de l'élément structurant a été fixée à 1/8 de la largeur moyenne. Cette dimension est très importante comme nous le verrons plus loin dans ce chapitre.

5.4.1.3 Longueur de l'élément structurant

L'élément structurant qui correspond à un pixel donné dépend de la classe de vitesse associée à ce pixel. La longueur de l'élément structurant est contrôlée par cette vitesse comme indiqué dans la figure 5.14 [KHO 96/2].

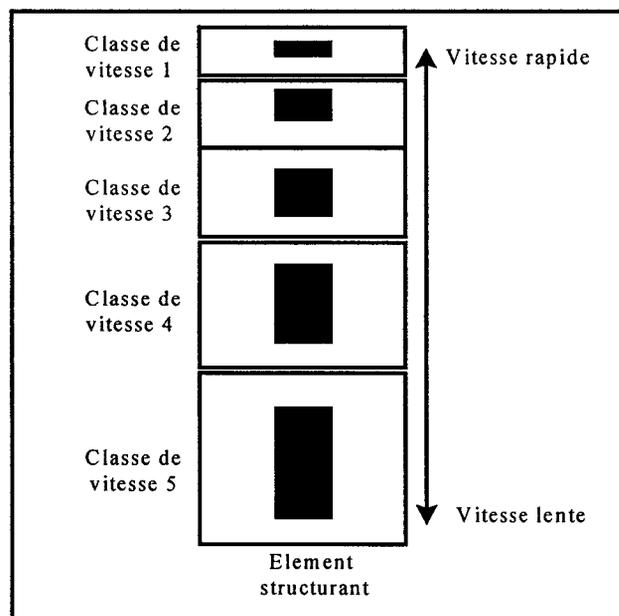


Figure 5.14 : correspondance vitesse-élément structurant

Les vitesses correspondant à la classe 1 concernent des piétons qui marchent vite ou courent sous les caméras. L'élément structurant correspondant est court puisque les silhouettes associées à ces piétons sont elles-mêmes courtes.

A l'inverse, les vitesses de la classe 5 correspondent à des piétons qui se déplacent lentement, conduisant à de longues silhouettes. Un élément structurant plus long est alors nécessaire pour tenir compte de cette vitesse lente.

5.4.1.4 Définition de la largeur de l'élément structurant

La largeur de l'élément structurant est un paramètre fixe, très important dans l'opération d'individualisation des silhouettes qui constituent les formes. Il a été dit au § 5.4.1.2 que ce

paramètre est défini en fonction de la largeur moyenne des silhouettes. Sa détermination est souvent délicate car un mauvais choix peut conduire à des erreurs de comptage.

Largeur d'élément structurant trop réduite

Si la largeur de l'élément structurant est trop petite, la forme traitée ne sera pas séparée en silhouettes (Figure 5.15).

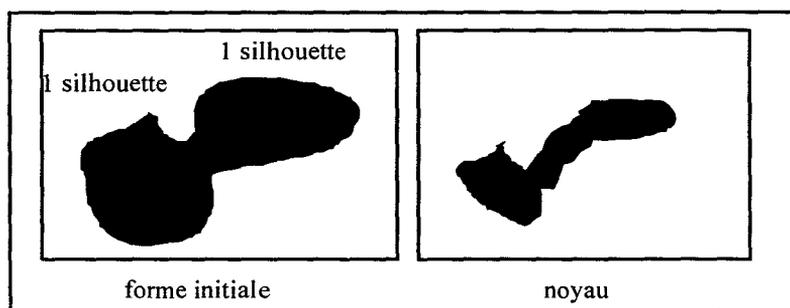


Figure 5.15 : éléments structurants trop étroits

Sur la figure 5.15, on constate que des éléments structurants insuffisamment larges ne permettent pas d'individualiser les deux silhouettes constituant la forme initiale. Après traitement, un seul noyau subsiste. Un sous-comptage sera observé dans ce cas, car le comptage final associe un noyau à un piéton.

Largeur d'élément structurant trop élevée

Si l'élément structurant est trop large, le phénomène inverse se produit (Figure 5.16). La nouvelle image créée ne contient plus de noyau. Dans ce cas, un sous-comptage encore plus important est observé.

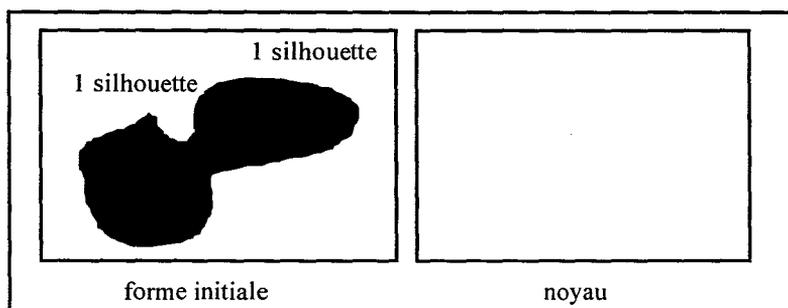


Figure 5.16 : éléments structurants trop larges

Largeur d'élément structurant conseillée

Il n'y a pas de règle générale quant à la définition de la "bonne" largeur des éléments structurants, les formes obtenues étant très différentes d'un site de mesures à l'autre. Nous avons choisi d'utiliser un sous-multiple de la largeur moyenne des silhouettes contenues dans la base de données. Le choix du sous-multiple le mieux adapté, 1/2 ou 1/3 ou 1/4 etc, est effectué manuellement, par approches successives, dans une phase de réglage du dispositif.

La figure 5.17 fournit le résultat d'un traitement avec une largeur d'élément définie correctement.

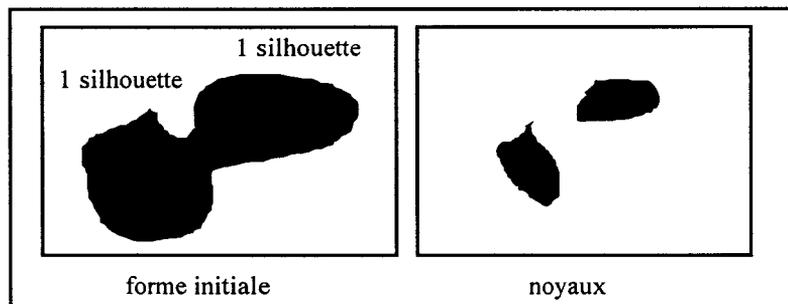


Figure 5.17 : éléments structurants de largeur correcte

5.4.1.5 Résultats

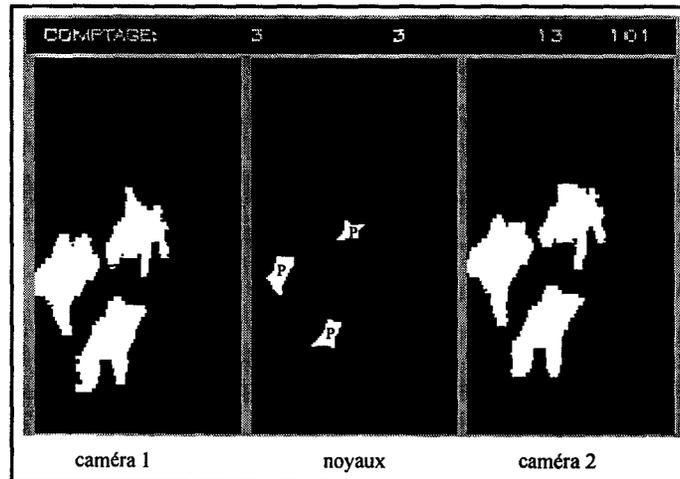
Les paragraphes qui suivent fournissent les résultats de traitement des séquences présentées dans le chapitre IV. Ces résultats doivent être consultés en même temps que ces séquences. Nous rappelons que chaque traitement porte sur l'intersection des images binaires issues des deux caméras.

Dans chaque séquence résultat, les formes gauche et droite correspondent aux données d'origine recueillies par les deux caméras. La partie centrale contient les noyaux, qui sont les résultats du traitement par l'algorithme à éléments structurants rectangulaires. Dans le comptage final un noyau est associé à un piéton.

A chaque résultat de séquence, des lettres sont placées sur les noyaux pour aider à leur interprétation.

Séquence 1

La séquence 1 correspond à des piétons isolés. Les noyaux montrent que les structures géométriques des silhouettes d'origine ont été modifiées par l'application de l'élément structurant. En effet, l'élément structurant est rectangulaire alors que les silhouettes ont des géométries plutôt rondes. Le problème de séparation des formes ne se pose pas pour cette séquence car les piétons sont isolés dans la séquence d'origine.



Séquence 1 : piétons isolés (P : piéton)

Séquence 2

Dans cette séquence, un grand nombre de formes ont été séparées, soit directement par la vitesse, soit par l'application de l'algorithme avec un élément structurant rectangulaire. Un noyau correspond encore à deux personnes, à savoir un enfant et un adulte.



Séquence 2 : piéton immobile (P : piéton, E : enfant)

Séquence 3

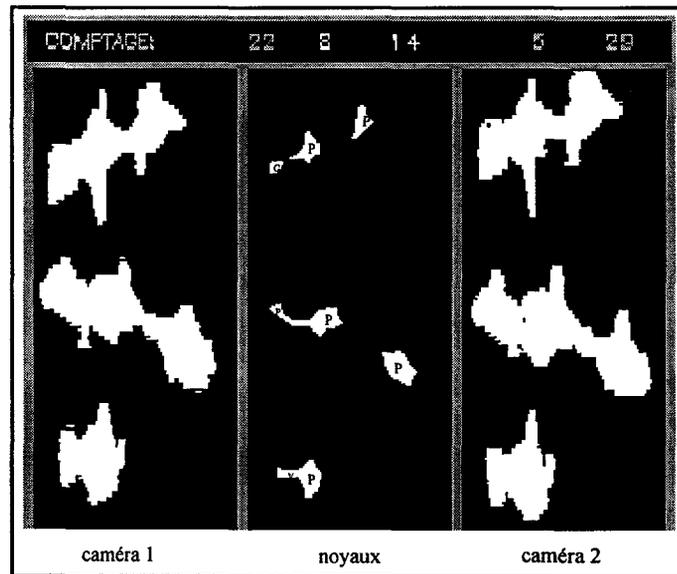
Dans la séquence 3, toutes les formes correspondent à deux silhouettes connexes. Certaines de ces formes n'ont pas été divisées en silhouettes. Ceci peut être dû à un mauvais paramétrage de la largeur de l'élément structurant utilisé.



Séquence 3 : série de deux silhouettes connexes (P : piéton)

Séquence 4

Le traitement de cette séquence montre que le carton, aussi large que la personne qui le porte, a été assimilé à une silhouette. La forme correspondant au triplet de piétons a été séparée en deux et le piéton portant une valise a été correctement identifié puisque l'on obtient un seul noyau.



Séquence 4 : piétons et objets (P : piéton, C : carton, V : valise)

Séquence 5

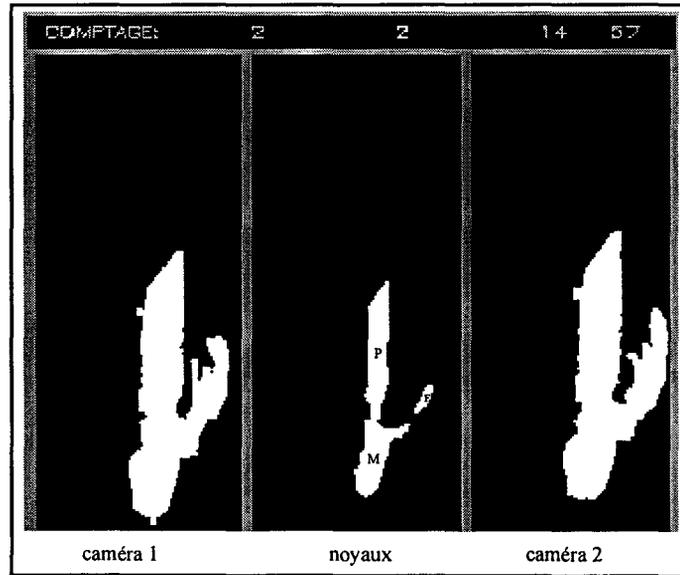
Tous les piétons liés aux formes de cette séquence ont été identifiés soit directement par analyse de la vitesse, soit en utilisant l'algorithme avec un élément structurant rectangulaire. Il faut noter néanmoins que la longue forme de droite (partie centrale) a été divisée en deux dans sa partie inférieure. Nous verrons plus loin que le comptage final tient compte de ce phénomène.



Séquence 5 : groupe d'enfants (P : piéton, E : enfant)

Séquence 6

Dans cette séquence plutôt atypique, il faut remarquer que la forme initiale a été divisée en deux : d'un côté le noyau correspondant à la mère poussant le landeau et de l'autre l'enfant.



Séquence 6 : piétons et objets particuliers (M : mère, E : enfant, P : poussette)

5.4.2 Élément structurant ellipsoïdal dynamique

Bien qu'ayant obtenu des résultats satisfaisants en terme d'identification de silhouettes avec l'élément structurant rectangulaire, nous avons pu remarquer ses limites dans certaines configurations.

En effet, cet élément structurant, une fois sa largeur fixée, ne varie qu'en longueur en fonction de la vitesse. Sa géométrie reste par contre identique, c'est à dire celle d'un rectangle plus ou moins long.

C'est pourquoi, nous avons envisagé de procéder aux mêmes traitements, mais avec une toute autre géométrie d'élément structurant. Il s'agit en fait d'une série d'éléments structurants qui vont changer non seulement de taille mais aussi de forme en fonction de la vitesse des piétons.

L'élément structurant ellipsoïdal [KHO 96/4], comme le rectangulaire, est un élément à deux dimensions, largeur et longueur, comme le montre la figure 5.18.

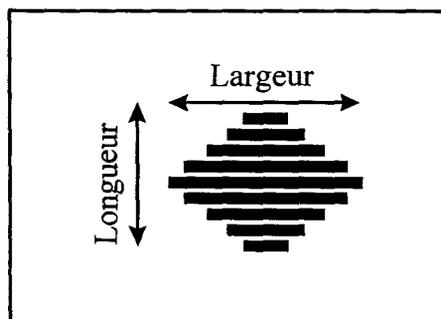


Figure 5.18 : dimensions de l'élément structurant

Utilisation de la vitesse

Dans ce cas de figure, la vitesse définit non seulement la longueur de l'élément structurant, mais aussi sa forme. Ainsi en fonction des vitesses mesurées, on n'utilise pas un élément structurant unique mais une série d'éléments structurants de longueurs différentes.

5.4.2.1 Largeur et longueur de l'élément structurant

Largeur de l'élément structurant

La partie la plus large de l'élément structurant ellipsoïdal est calculée comme la largeur de l'élément structurant rectangulaire, c'est à dire $1/8$ de la largeur moyenne calculée.

Longueur de l'élément structurant

Comme précédemment, les vitesses sont calculées, filtrées et associées à des classes. Les 5 classes de vitesses utilisées pour définir la longueur des éléments structurants ellipsoïdaux sont les mêmes que celles utilisées pour l'élément structurant rectangulaire.

L'élément structurant qui correspond à un pixel donné dépend de la classe de vitesse associée à ce pixel. La longueur de l'élément structurant est contrôlée par cette vitesse.

Les vitesses correspondant à la classe 1 concernent les piétons qui marchent vite ou courent sous les caméras. L'élément structurant correspondant est court puisque les formes associées à ces piétons sont elles-mêmes courtes.

A l'inverse, les vitesses de la classe 5 correspondent aux piétons qui se déplacent lentement, conduisant ainsi à des formes longues. Un élément structurant plus long est alors nécessaire pour séparer d'éventuelles silhouettes connexes.

5.4.2.2 Application de l'élément structurant ellipsoïdal

La correspondance entre la taille et la forme de l'élément structurant est illustrée sur la figure 5.19.

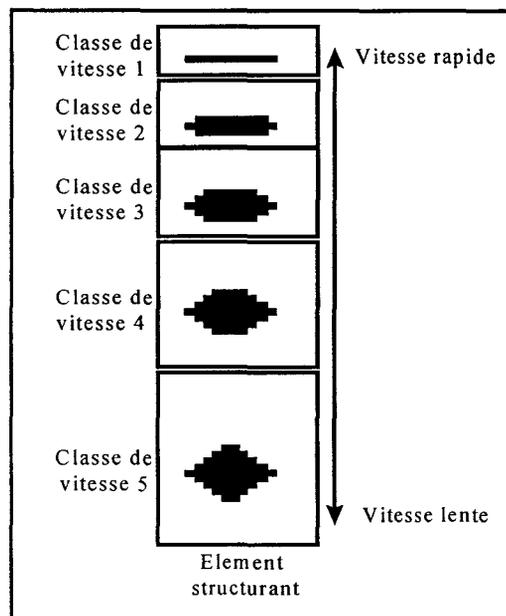


Figure 5.19 : correspondance vitesse-élément structurant

Par la suite, et comme pour le traitement avec l'élément structurant rectangulaire, la procédure est la suivante : chaque pixel d'une forme dans l'image binaire est considéré comme le centre de l'élément structurant qui lui correspond.

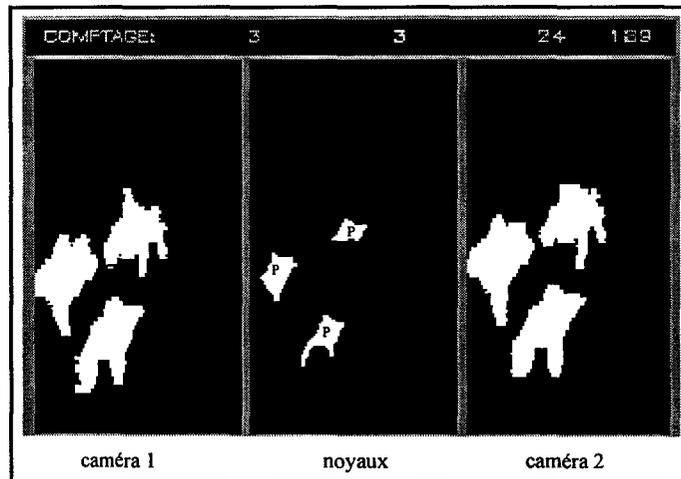
Si cet élément est complètement inclus, au sens ensembliste, dans la forme traitée, le pixel concerné est marqué. L'ensemble des pixels marqués forme une nouvelle image binaire.

On peut noter aussi, sur la figure 5.19, que pour chaque classe de vitesse, l'élément structurant correspondant est différent, non seulement en longueur, mais aussi en apparence géométrique.

5.4.2.4 Résultats obtenus

Séquence 1

Dans cette séquence, on remarque que les noyaux sont plus petits que ceux résultant du traitement par l'élément rectangulaire, qui avait tendance à modifier la géométrie des formes initiales. L'élément structurant ellipsoïdal a tendance à les "épouser", à les lisser. C'est pourquoi, dans les résultats du traitement de cette séquence, la géométrie des noyaux s'apparente plus à celle des formes d'origine.



Séquence 1 : piétons isolés (P : piéton)

Séquence 2

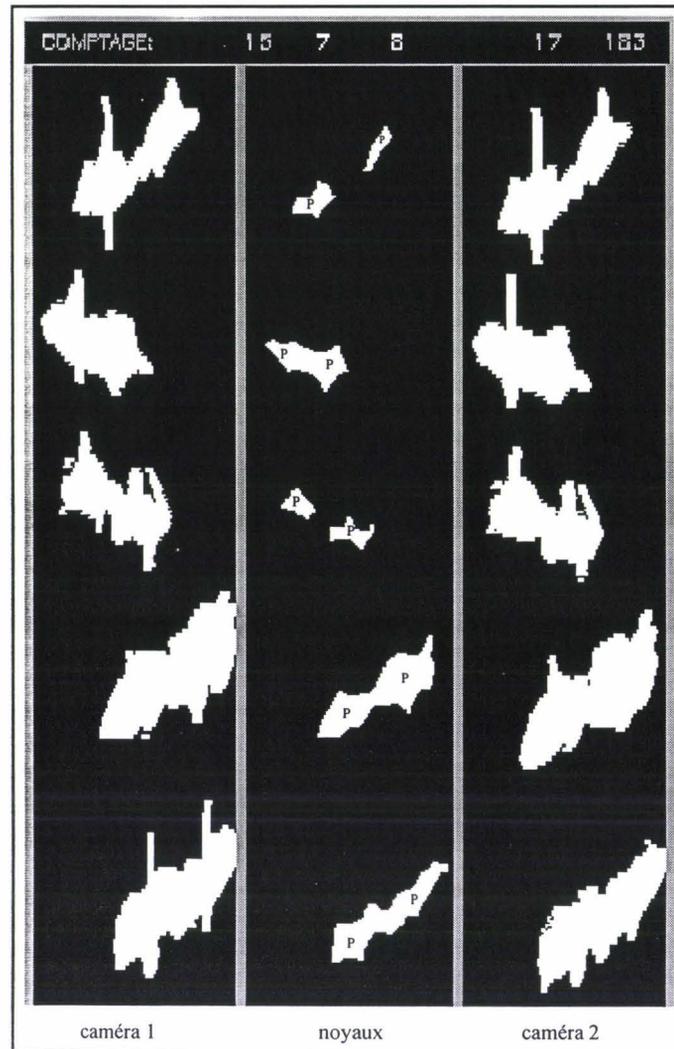
En terme de séparation de formes, les résultats pour cette séquence sont identiques à ceux observés avec l'élément rectangulaire : seule la silhouette liée à l'enfant n'a pas été séparée de celle qui lui est adjacente et qui correspond à un piéton adulte.



Séquence 2 : piéton immobile (P : piéton, E : enfant)

Séquence 3

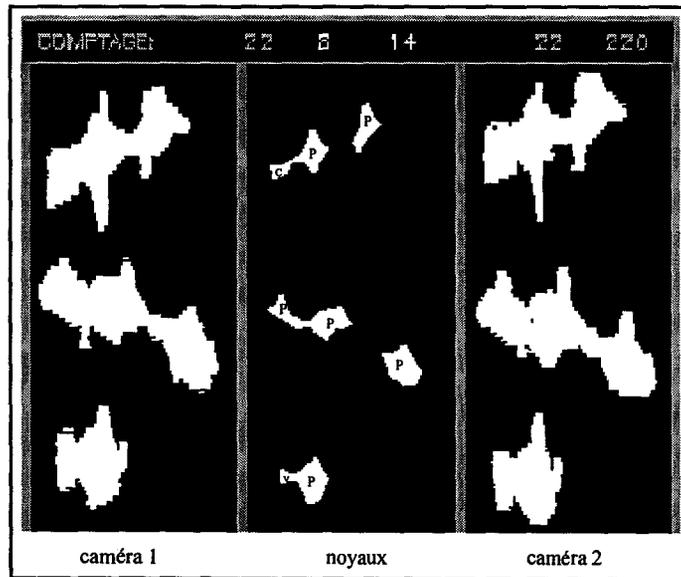
Toutes les silhouettes qui ont été individualisées par l'élément rectangulaire le sont aussi par l'élément ellipsoïdal. Des formes contenant des silhouettes connexes n'ont pas été séparées.



Séquence 3 : Série de deux silhouettes connexes (P : piéton)

Séquence 4

Dans le traitement de cette séquence, la forme relative aux deux piétons et au carton a été divisée en deux, alors que l'élément rectangulaire séparait le carton de son propriétaire conduisant ainsi à trois noyaux. Dans les autres cas, les résultats sont identiques.



Séquence 4 : piétons et objets (P : piéton, V : valise, C : carton)

Séquence 5

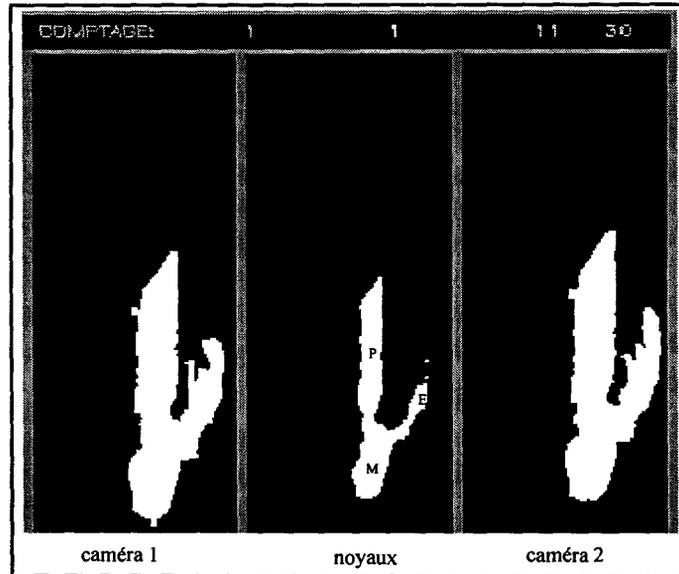
Dans cette séquence, on observe à nouveau une différence avec l'algorithme précédent. La forme très longue de droite n'est plus divisée en deux dans sa partie inférieure. Les autres silhouettes sont correctement identifiées.



Séquence 5 : groupe d'enfants (P : piéton, E : enfant)

Séquence 6

La capacité de séparation de l'élément ellipsoïdal étant moins forte que celle de l'élément rectangulaire, les silhouettes correspondant à la mère et à l'enfant n'ont pas été séparées.



Séquence 6 : piétons et objets particuliers (M : mère, P : poussette, E : enfant

On peut résumer à l'issue du traitement de ces séquences que l'élément structurant ellipsoïdal modifie moins la géométrie des formes d'origine que l'élément structurant rectangulaire. Dans certains cas, sur la séquence 6 par exemple, sa capacité à séparer les formes en silhouettes est moins forte que celle du rectangulaire.

5.4.3 Élément structurant directionnel dynamique

La troisième méthode que nous avons développée, en nous inspirant de la morphologie mathématique, utilise aussi une série d'éléments structurants qui dépendent des vitesses observées. Elle permet de traiter de manière plus performante les formes qui correspondent à des piétons se déplaçant en biais par rapport aux plans de surveillance.

En fonction de la localisation du système de comptage, on observe très souvent des piétons traversant les plans de surveillance en biais et non perpendiculairement aux plans de surveillance (Figure 5.20).

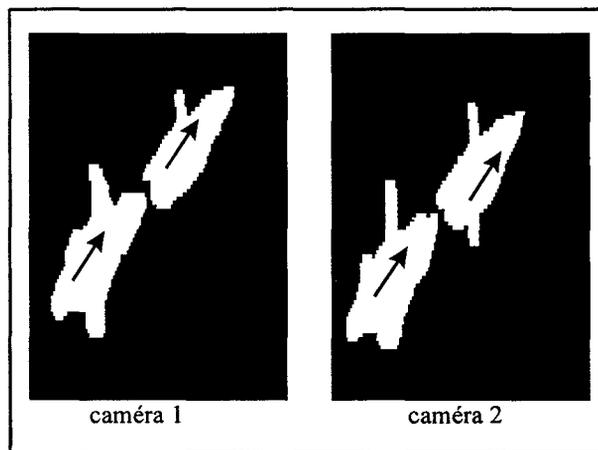


Figure 5.20 : déplacements de biais

Les éléments structurants rectangulaires et ellipsoïdaux ne permettent pas de traiter correctement ces formes car leurs tailles ne varient qu'en longueur. C'est pourquoi, nous avons développé une méthode qui prend en compte la direction du mouvement des piétons [KHO 96/3].

5.4.3.1 Classification des trajectoires des piétons

Toutes les directions de mouvement possibles des piétons par rapport aux plans de surveillance ont été réparties en 6 grandes classes de direction, séparées chacune de 30 degrés (Figure 5.21).

Sur la figure 5.21, la direction numérotée 1 représente un déplacement perpendiculaire aux plans de surveillance dans le sens positif, le sens négatif de déplacement étant représenté par la flèche en pointillé de direction opposée. La direction numérotée 6 correspond à un déplacement parallèle aux plans de surveillance dans le sens positif, le sens négatif de

déplacement étant représenté par la flèche en pointillé de direction opposée. Les autres directions de mouvement, 2, 3, 4, 5 sont intermédiaires.

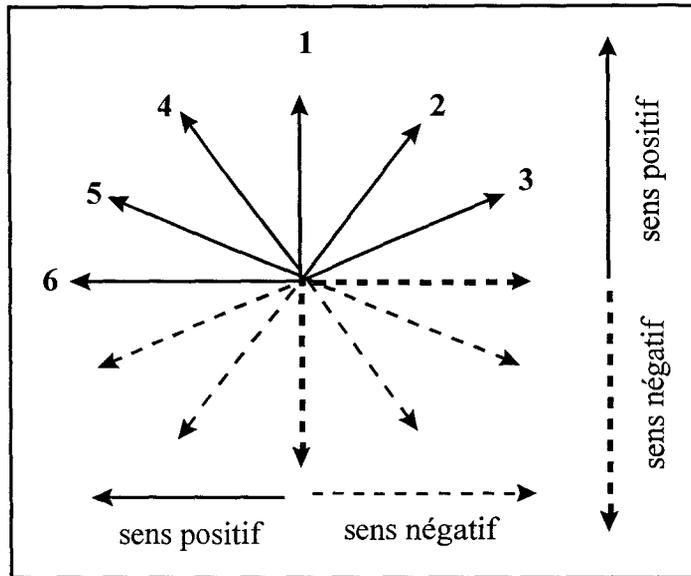


Figure 5.21 : classification des directions de mouvement des piétons

5.4.3.2 La série d'éléments structurants utilisée

Le système de comptage que nous voulons mettre en œuvre est un système temps réel, dont le temps de traitement doit être le plus court possible. Le traitement est réalisé de manière séquentielle, c'est à dire que chaque image ligne acquise est immédiatement traitée pour éviter toute opération de stockage sur disque.

Nous ne connaissons pas à l'avance, c'est à dire pendant l'acquisition des images lignes, la direction du mouvement des piétons détectés. C'est pourquoi, pour tenir compte de cette direction, on imagine pour chaque pixel tous les scénarios possibles et vraisemblables.

Application de l'élément structurant le plus adapté aux formes

La procédure de traitement développée avec l'élément structurant directionnel est la même que celle utilisée avec les éléments structurants rectangulaires et ellipsoïdaux. Chaque pixel d'une forme est considéré comme le centre de l'élément structurant qui lui correspond. Si cet élément est complètement inclus, au sens ensembliste, dans la forme traitée, le pixel concerné est marqué. L'ensemble des pixels marqués forme une nouvelle image binaire.

Largeur de l'élément structurant

La procédure de définition de la largeur de l'élément structurant directionnel est la même que celle utilisée pour l'élément rectangulaire, c'est à dire 1/8 de la largeur moyenne calculée.

Longueur de l'élément structurant

L'élément structurant qui correspond à un pixel donné dépend de la classe de vitesse associée à ce pixel. La longueur de l'élément structurant est contrôlée par cette vitesse. Cette procédure est identique à celle utilisée pour l'élément structurant rectangulaire.

Direction de l'élément structurant

Il reste à choisir la direction de l'élément structurant. Nous avons vu précédemment que les directions de mouvement des piétons ont été réparties en 6 classes. Pour chaque classe de direction, le traitement est réalisé avec un un élément structurant différent (Figure 5.22).

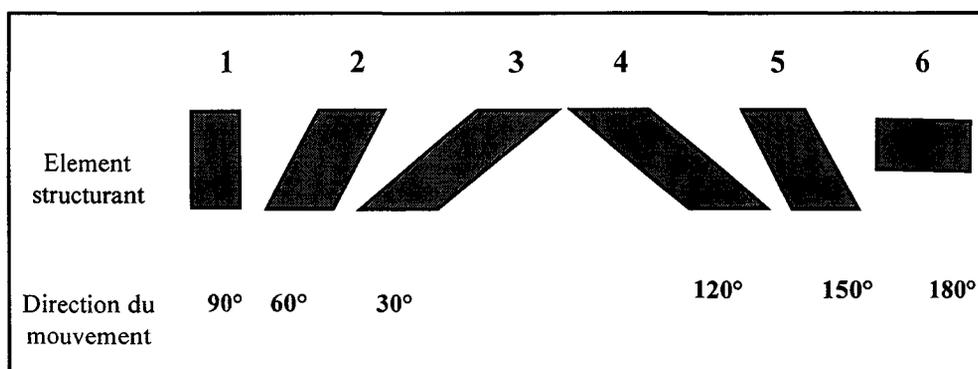


Figure 5.22 : correspondance direction de mouvement - élément structurant

Nous avons dit précédemment qu'à chaque pixel correspond un élément structurant défini par sa largeur, sa longueur et une direction encore inconnue.

Pour chaque pixel, les éléments directionnels sont testés tour à tour dans l'ordre croissant du numéro 1 au numéro 6. On choisit celui qui convient, c'est à dire qui est complètement inclus au sens ensembliste dans la forme analysée. Bien évidemment, dans beaucoup de cas, pour un pixel donné, plusieurs éléments directionnels conviennent. Dans ce cas, on choisit arbitrairement le premier élément testé qui convient afin de réduire le temps de traitement. Ainsi, les éléments suivants ne sont pas testés.

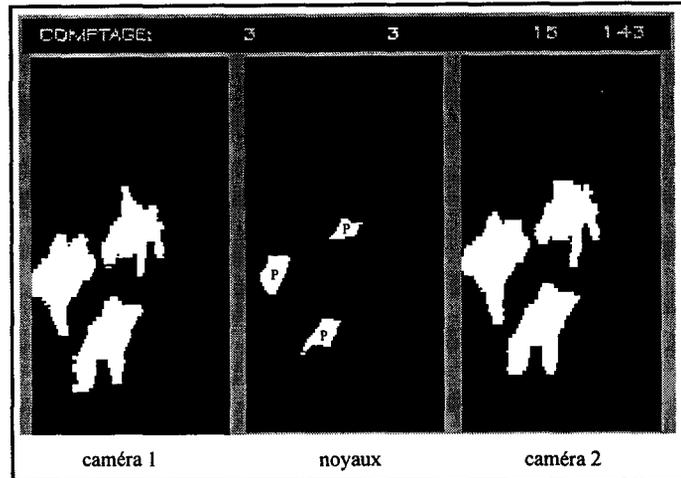
Si aucun des 6 éléments ne convient, le pixel n'est pas marqué et ne fait pas partie de la nouvelle image binaire.

Ainsi, chaque forme analysée ne sera pas traitée par un seul élément mais par une série d'éléments qui diffèrent en longueur et en orientation, leur largeur ayant été fixée une fois pour toutes au départ.

5.4.3.3 Résultats obtenus

Séquence 1

Avec l'élément directionnel, la géométrie des formes initiales est encore mieux respectée qu'avec l'élément rectangulaire.



Séquence 1 : piétons isolés (P : piéton)

Séquence 2

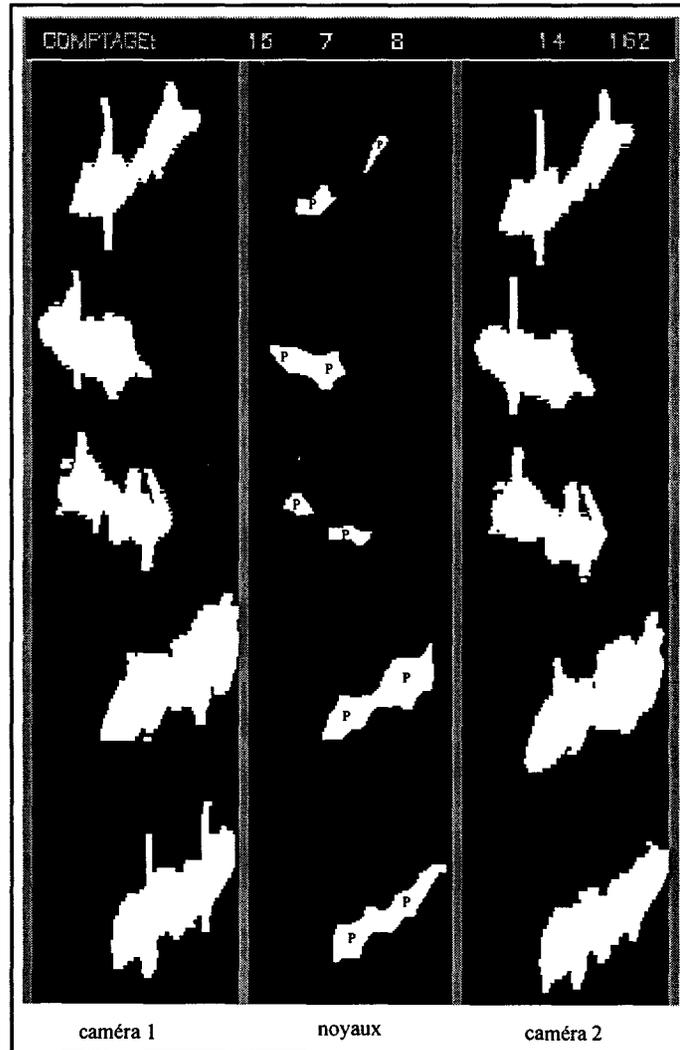
Pour cette séquence, l'algorithme utilisant l'élément directionnel donne le même résultat en termes de comptage que les deux précédents. Seule la forme relative à l'adulte et l'enfant n'a pas été divisée en deux noyaux.



Séquence 2 : piéton immobile (P : piéton, E : enfant)

Séquence 3

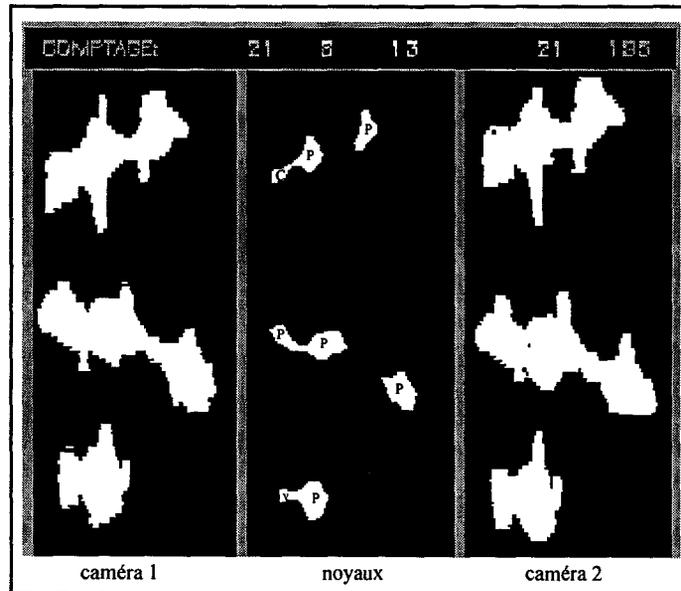
Ici, en dehors du respect de la géométrie des formes initiales, les résultats en termes de séparation sont les mêmes que pour les algorithmes précédents.



Séquence 3 : série de deux silhouettes connexes (P : piéton)

Séquence 4

Les résultats sont les mêmes que pour l'algorithme avec élément ellipsoïdal.



Séquence 4 : piétons et objets (P : piéton, V : valise, C : carton)

Séquence 5

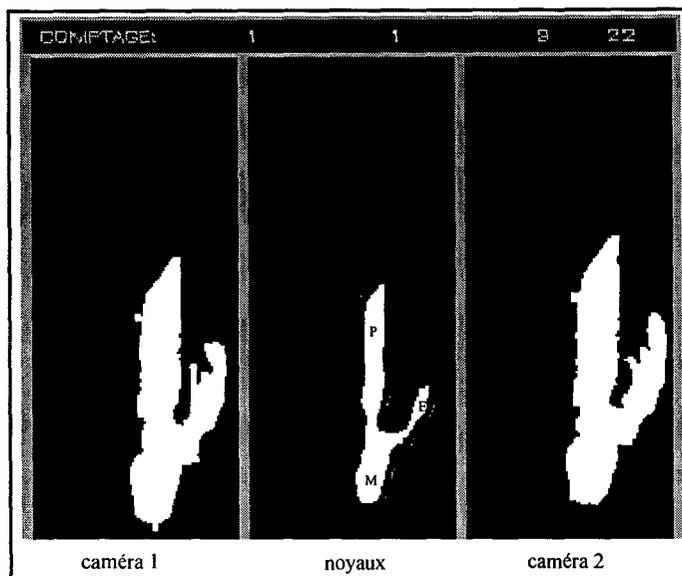
Les résultats sont les mêmes que pour l'algorithme avec élément ellipsoïdal.



Séquence 5 : groupe d'enfants (P : piéton, E : enfant)

Séquence 6

Les résultats sont les mêmes que pour l'algorithme avec l'élément structurant ellipsoïdal.



Séquence 6 : piétons et objets particuliers (M : mère, P : poussette, E : enfant)

A l'issue du traitement des séquences par l'élément structurant directionnel, la synthèse est la suivante :

- l'élément structurant directionnel respecte bien la géométrie des formes d'origine mais sépare moins bien les formes dans certains cas,
- l'élément structurant ellipsoïdal respecte mieux la géométrie des formes que le rectangulaire mais moins bien que le directionnel. Sa capacité à séparer les formes est meilleure que la capacité du directionnel et à peu près la même que celle du rectangulaire,
- l'élément structurant rectangulaire modifie fortement la géométrie des formes d'origine.

Nous verrons dans le chapitre VII, qui fournit les résultats du comptage, que les précisions de comptage des 3 algorithmes sont voisines.

5.5 SUIVI DES FORMES

Nous avons vu, dans les paragraphes précédents, trois algorithmes de comptage avec des éléments structurants rectangulaires, ellipsoïdaux et directionnels.

Une fois les éléments structurants appliqués, on obtient une nouvelle image binaire contenant les noyaux. On considère que chaque noyau dans cette nouvelle image binaire correspond à un piéton.

Un simple comptage de ces noyaux, en associant un piéton à chaque noyau, conduirait à de fortes erreurs si l'on ne tenait pas compte de certains phénomènes indésirables. C'est pourquoi deux procédures supplémentaires sont utilisées pour que le comptage final soit correct.

5.5.1 La procédure de suivi des formes

La procédure de suivi des formes a été développée pour éviter des erreurs de comptage après l'extraction des noyaux. Pour ce faire, elle utilise les formes ou silhouettes dont ces noyaux ont été extraits. Cette procédure est nécessaire quel que soit le type d'élément structurant utilisé.

La figure 5.23 est organisée en 3 parties : à gauche apparaît une forme provenant de la camera 1, à droite celle correspondant à la caméra 2 et au milieu sont représentés les noyaux correspondant au traitement de la forme par l'élément structurant rectangulaire.

On sait que cette forme traitée est une silhouette alors que le dénombrement des noyaux indiquerait deux piétons pour cette silhouette. Nous aurions donc un sur-comptage.

Ce cas se produit lorsque les piétons traversent les plans de surveillance lentement et de biais. Il peut se produire aussi lorsqu'un piéton est arrêté sous les caméras et ne bouge que ses membres.

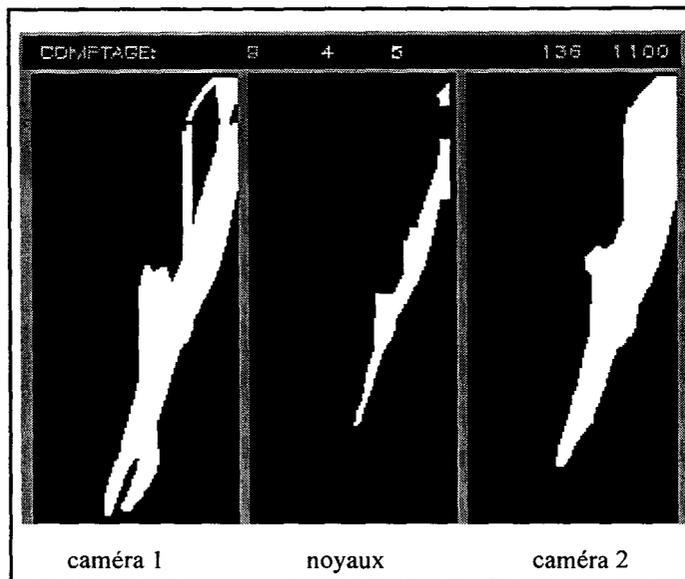


Figure 5.23 : silhouette divisée en deux noyaux

La procédure de suivi des formes, décrite ci-dessous, permet de palier cet inconvénient.

La figure 5.24 est identique à la figure 5.23, sauf pour la partie centrale qui représente les noyaux, en blanc, et l'intersection des deux silhouettes des parties gauche et droite, en gris. Nous rappelons que les traitements avec les éléments structurants se font toujours sur cette intersection, obtenue rappelons-le par superposition des deux images binaires captées par les deux caméras.

Nous appelons **zone de connexité** la zone occupée par la silhouette grise de la partie centrale.

Pour le comptage final, on ne comptera qu'un seul piéton, si on montre que les deux noyaux appartiennent à la même zone de connexité ou à la même silhouette. C'est sur cette idée qu'est basée la procédure de suivi présentée ici. Le suivi se fait sur la zone de connexité.

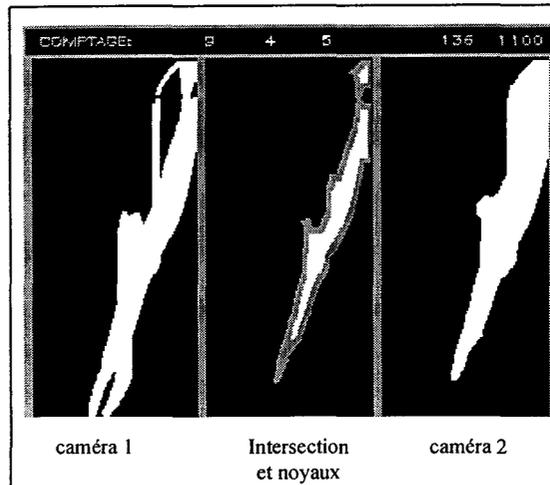


Figure 5.24 : noyaux et zone de connexité

Le suivi des formes

La procédure de suivi des formes est expliquée à l'aide de scénarios.

5.5.1.1 Scénario 1

C'est le scénario le plus classique dans le cas d'un piéton qui reste sous les caméras en ne bougeant que ses membres, ce qui conduit à observer, après traitement par l'un quelconque des éléments structurants, plusieurs noyaux pour ce piéton. Les noyaux figurent en gris foncé sur la figure 5.25.

La procédure de suivi n'est mise en œuvre que si la longueur de la forme ou de la silhouette analysée, en nombres de lignes d'acquisition, dépasse un certain seuil. Ce seuil a été fixé à 100 lignes d'acquisition. Pour une période d'échantillonnage de 10 ms, cela veut dire que le piéton reste au moins une seconde sous les caméras. Au delà d'une seconde, cela signifie que le piéton traverse les 10 cm qui séparent les deux plans de surveillance à moins de 0.1 m/s, soit à moins de 0.36 km/h. A ce niveau de vitesse, nous considérons que le piéton est presque arrêté. C'est pourquoi, nous appliquons l'analyse de suivi sur sa silhouette.

Une fois la silhouette traitée, le premier noyau, la forme ou silhouette étant parcourue sur sa longueur, est comptabilisé (Figure 5.25). La procédure de suivi, notée "analyse de connexité", est mise en œuvre si le numéro de la ligne i , qui correspond au numéro de la dernière ligne du noyau, est supérieur à 100. La numérotation des lignes commence à l'origine de la silhouette initiale, origine par rapport au temps.

On effectue ensuite une analyse de connexité qui consiste à suivre la silhouette dans son ensemble à partir de la ligne i , et à vérifier si de nouveaux noyaux existent à l'intérieur. Ces derniers ne sont pas comptés s'ils sont situés dans la zone de connexité (en pointillé sur la figure 5.25).

Dans le cas de la figure 5.25, un seul noyau sera comptabilisé car les deux suivants se trouvent dans la zone de connexité, et correspondent donc au même piéton.

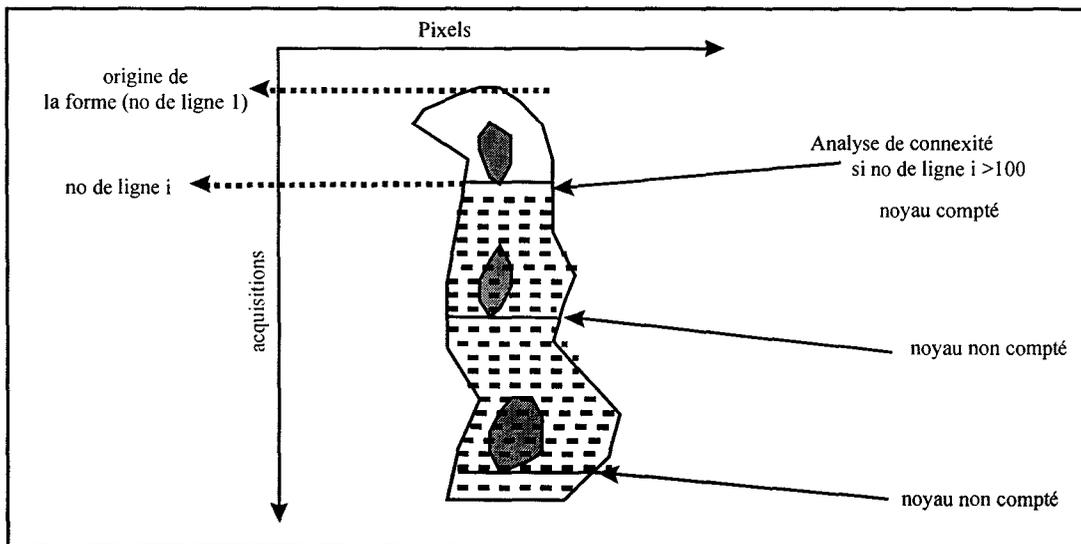


Figure 5.25 : scénario 1

5.5.1.2 Scénario 2

Le scénario 2 est légèrement différent du premier : un piéton, statique sous les caméras, bouge ses membres. Pendant ce temps, un autre piéton passe suffisamment près de lui pour que leurs deux silhouettes soient connexes. Cette situation est illustrée sur la figure 5.26.

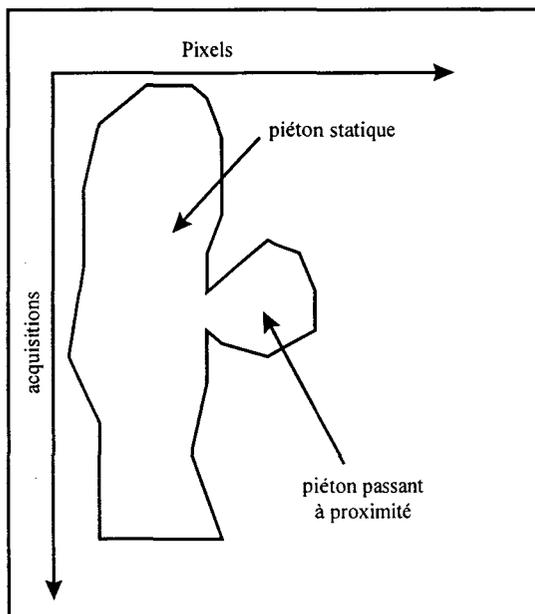


Figure 5.26 : scénario 2

La procédure de séparation directe par la vitesse, développé au §5.2 , permet de diviser la forme de la figure 5.25 en deux silhouettes (Figure 5.27).

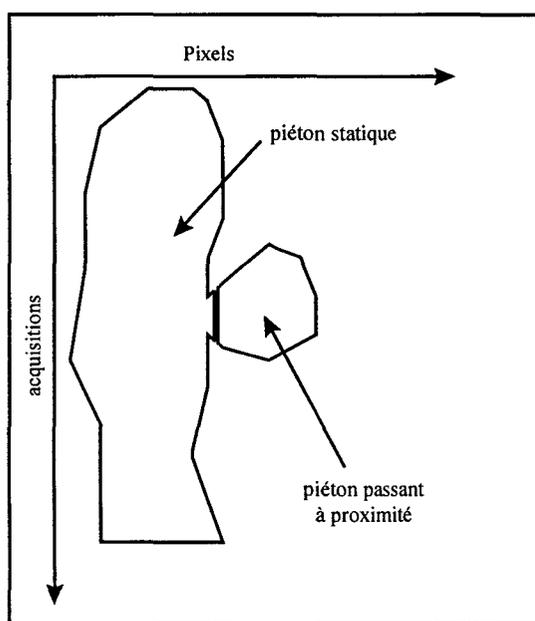


Figure 5.27 : scénario 2 avec silhouettes séparées

Si, après un traitement avec un des éléments structurants, plusieurs noyaux correspondent à une même silhouette, on est ramené au scénario 1 que l'on sait traiter.

Dans tous les cas de figure, grâce à la procédure de séparation directe par la vitesse, il est possible de se ramener à la situation du scénario 1. Cette procédure est très importante car les deux scénarios précédemment décrits se produisent très souvent dans la pratique.

5.5.2 La procédure de filtrage des formes

C'est une procédure qui intervient en tout dernier lieu pour compter les noyaux, une fois toutes les procédures mises en œuvre : séparation directe par la vitesse, traitement par éléments structurants et analyse de connexité.

Parmi les noyaux trouvés, certains ne sont pas comptés si leur surface en pixels est inférieure à un seuil. Ce seuil a été fixé à 30 pixels. Ceci permet d'éliminer les noyaux trop petits pour correspondre à un piéton.

Ce cas de figure, assez fréquent, est illustré par la figure 5.28. Une silhouette, traitée par l'élément structurant ellipsoïdal, est séparée en deux noyaux dont l'un n'est pas comptabilisé en raison de sa taille insuffisante.

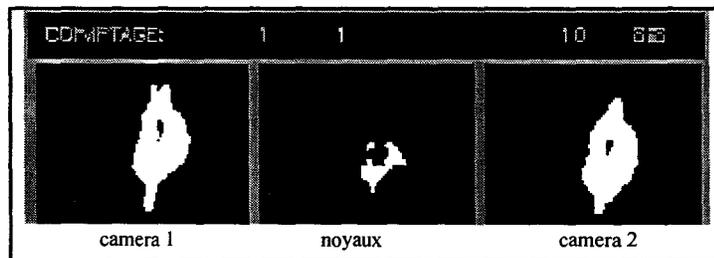


Figure 5.28 : silhouette divisée en deux noyaux

5.6 CONCLUSION

Dans ce chapitre, nous avons présenté trois algorithmes de comptage de piétons, tous inspirés des concepts de la morphologie mathématique, qui ont été élargis et enrichis. La vocation de ces algorithmes est d'associer le bon nombre de piétons aux formes extraites des images lignes saisies par les caméras. Nous avons vu que l'algorithme utilisant une série d'éléments structurants rectangulaires tend à modifier les structures géométriques des formes d'origine, alors que ceux utilisant les éléments ellipsoïdaux et directionnels les préservent mieux.

Les trois algorithmes présentés dans ce chapitre ne suffisent pas à résoudre le problème du comptage des piétons. Des procédures supplémentaires, notamment celle de la séparation directe par la vitesse, celle du suivi des formes et du filtrage des noyaux pour le comptage final, sont un complément indispensable.

Les résultats présentés montrent que la plupart des piétons correspondant à des silhouettes connexes, présentées dans les différentes séquences, ont été correctement identifiés.

Les séquences que nous avons présentées comme illustrations sont des "morceaux choisis" surtout destinés à mettre en évidence les difficultés rencontrées.

Nous verrons, dans le chapitre VII qui fournit les résultats globaux et analyse les performances, que ces trois algorithmes ont donné des résultats tout à fait prometteurs.

Cependant, dans certains cas, les silhouettes sont difficiles à individualiser. Nous avons développé un algorithme supplémentaire de comptage qui permet de résoudre certains de ces cas difficiles.

Le chapitre suivant est consacré à la description de ce nouvel algorithme, appelé ASSESA, acronyme de **A**lgorithme de **S**egmentation par **S**ous-**E**chantillonnage et **S**euillage **A**daptatif.

Chapitre VI

L'algorithme ASSESA

Chapitre VI

L'ALGORITHME ASSESA

6.1 INTRODUCTION

Nous avons vu dans le chapitre précédent que les algorithmes qui s'inspirent de la morphologie mathématique rencontrent, dans certains cas, des difficultés pour l'individualisation des silhouettes incluses dans les formes. Parmi les trois éléments structurants utilisés dans le chapitre V, le rectangulaire est celui qui permet d'obtenir la meilleure individualisation des silhouettes, et nous avons remarqué que celui-ci tend à modifier fortement les formes d'origine.

Pour le comptage des piétons, tel que nous l'abordons, il n'est pas nécessaire de maintenir la géométrie des formes d'origine. L'important est d'associer le bon nombre de piétons aux formes traitées. Une fois le comptage des piétons réalisé, il n'est pas nécessaire, pour nous, de revenir aux formes d'origine. En utilisant l'algorithme ASSESA, on recherche les "zones principales" ou "centrales" des formes en éliminant progressivement les détails tels que les membres des piétons, les objets portés, etc...

Le principe de l'algorithme est de mettre en évidence, dans les formes, une ou plusieurs structures linéaires globales perceptibles à basse résolution.

C'est pourquoi, en utilisant cet algorithme, nous allons volontairement exagérer la modification géométriques des formes d'origine pour ne garder que les zones principales [KHO 96/5].

Dans de nombreuses applications de vision artificielle, disposer d'une image binaire constitue le but final des traitements dits de "bas niveau". Les procédures qui suivent concernent l'analyse de l'image binaire par des traitements dits de " haut niveau".

Dans notre application, binariser les images lignes d'origine constitue une première segmentation. Lorsque les piétons passent sous les caméras isolément, lors de périodes de faible fréquentation par exemple, le résultat de la première segmentation est un ensemble de régions appelées silhouettes. Le comptage des piétons est, dans ce cas, facile (Cf. §5.1).

En revanche, pendant une période où le flot de piétons est dense, les piétons passant de manière groupée sous les caméras, le résultat de cette segmentation est une image binaire constituée par un ensemble de régions appelées formes. Dans ce cas, avant le comptage final des piétons, une individualisation des silhouettes contenues dans les formes est nécessaire. Pour cela, nous opérons une transformation de cette image binaire par un algorithme que nous avons appelé ASSESA pour "Algorithme de Segmentation par Sous-Echantillonnage et Seuillage Adaptatif". Cet algorithme est organisé en deux étapes. Chaque étape se compose d'une procédure de sous-échantillonnage d'une image binaire suivie d'une procédure de seuillage.

6.2 PREMIERE ETAPE

6.2.1 Principe

L'image binaire initiale est sous-échantillonnée pour obtenir une nouvelle image de résolution inférieure. Pour cela, on découpe l'image initiale en rectangles $R1$ de taille fixe, $N1$ dans le sens de la largeur et $M1$ dans celui de la longueur (Figure 6.1).

Sur la figure 6.1, on note que les rectangles $R1$ de l'image initiale sont tous adjacents. L'image étant de taille $[N,L]$, N pixels et L lignes d'acquisitions, $N1$ est un diviseur de N et $M1$ un diviseur de L .

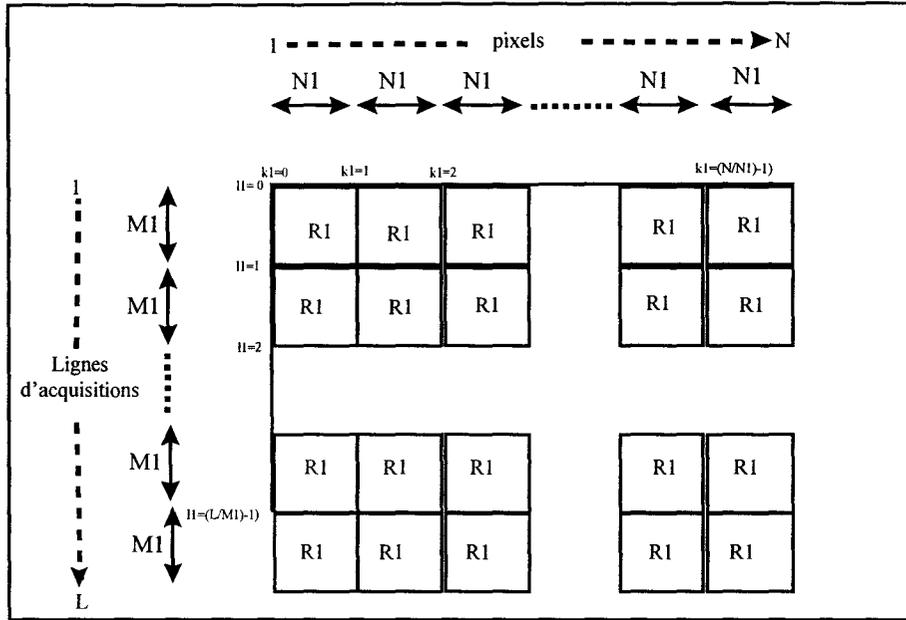


Figure 6.1 : sous-échantillonnage de l'espace correspondant à l'image binaire initiale

Dans chaque rectangle de l'image binaire initiale, appelée **image1**, le nombre de pixels de valeur 0 est comptabilisé ainsi :

$$NB_PIXEL_0[R1(k1,l1)] = NB_PIXEL_R1 - \sum_{i=1}^{N1} \sum_{j=1}^{M1} image1(k1 * N1 + i, l1 * M1 + j)$$

où

$NB_PIXEL_0[R1(k1,l1)]$ est le nombre de pixels à 0 dans le rectangle R1 dont la position dans l'image binaire (**image1**) est $(k1,l1)$ (Figure 6.1).

L'image binaire d'origine étant de taille (N, L) , N étant le nombre de pixels des barrettes CCD des caméras linéaires et L le nombre d'images lignes acquises, alors $k1$ varie de 0 à $(N/N1)-1$ et $l1$ de 0 à $(L/M1)-1$.

NB_PIXEL_R1 : nombre de pixels dans le rectangle R1 ($N1 \times M1$).

$Image1(i,j)$: valeur du pixel de coordonnées (i,j) dans **image1** (0 ou 1).

A cette valeur $NB_PIXEL_0[R1(k1,l1)]$ correspond ensuite un niveau de gris codé sur 8 bits, c'est à dire de 0 à 255, tel que :

$$NdG[R1(k1,l1)] = \frac{NB_PIXEL_0[R1(k1,l1)]}{NB_PIXEL_R1} \cdot 255$$

où $NdG[R1(k1,l1)]$ est le niveau de gris affecté au rectangle R1 dont la position dans **image1** est $(k1,l1)$.

Un rectangle R1 de image1 dont tous les pixels sont à 0 aura un niveau de gris 255 car pour ce rectangle NB_PIXEL_0 vaudra $N1 \times M1$. A l'inverse, un rectangle R1 dont tous les pixels sont à 1 aura un niveau de gris 0 car NB_PIXEL_0 vaudra 0. On voit donc que cette formule consiste à affecter de manière linéaire les niveaux de gris dans les différents rectangles en fonction des pixels à 0 qu'ils contiennent.

Cette opération d'affectation des niveaux de gris est effectuée pour chaque rectangle de l'image binaire (image1). Nous obtenons une image en niveaux de gris de résolution inférieure à celle de l'image binaire initiale.

Cette image en niveaux de gris est ensuite binarisée en utilisant une procédure de seuillage adaptatif. Cette procédure utilise les vitesses associées aux pixels dans les différents rectangles construits.

Pour chaque rectangle R1 considéré dans l'image binaire, on calcule la vitesse moyenne associée aux pixels à 0 qui le composent. Cette vitesse moyenne est associée à une classe de vitesse, dans la partition en 5 classes définie au chapitre IV.

Ensuite est établie une relation entre la classe de vitesse de chaque rectangle et un niveau de seuillage pour ce rectangle. Soit CCI_Ref le centre de la classe de vitesse des piétons se déplaçant le plus lentement, c'est à dire la classe 5. Par commodité, les classes de vitesse sont ici considérées en nombre d'acquisitions.

$$Seuil_Cl(i) = \frac{CCI(i)}{CCI_Ref} .255$$

où

$Seuil_Cl(i)$ est le seuil de binarisation correspondant à la classe de vitesse i ($i=1, \dots, 5$),

$CCI(i)$ est le centre de la classe de vitesse i , exprimée en nombre d'acquisitions.

On voit que, dans cette relation, CCI_Ref étant fixé, que plus la classe de vitesse correspond à une vitesse rapide, c'est à dire $CCI(i)$ faible et plus le niveau de seuillage est bas.

Le tableau 6.1 fournit le niveau de seuillage appliqué pour chaque classe de vitesse. Le centre de classe de vitesse 5 a été fixé à 20.

Pour les rectangles associés à la classe de vitesse nulle, on applique le niveau de seuillage maximum, c'est à dire 255.

Ainsi, on binarise chaque rectangle dans l'image en niveaux de gris avec le seuil correspondant. On obtient alors une nouvelle image binaire, appelée **image2**, de taille $(N/N1), (L/M1)$.

Classe de vitesse	Vitesse exprimée en nombre d'acquisitions	CCI(i)	CCI_Ref	Seuil_CI(i)
Classe 1	1 à 4	2,5	20	32
Classe 2	5 à 8	6,5	20	83
Classe 3	9 à 12	10,5	20	134
Classe 4	13 à 16	14,5	20	185
Classe 5	17 et plus	20	20	255

Tableau 6.1 : affectation des niveaux de binarisation en fonction des classes de vitesse

6.2.2 Application aux images

Sur la figure 6.2, nous présentons un exemple de ce traitement pour une image binaire.

Cette figure est divisée en trois parties :

- à gauche, une image binaire, appelée image1 dans l'algorithme, sur laquelle figure une forme. Cette forme correspond à deux silhouettes connexes,
- au milieu, une image en niveaux de gris résultant de l'affectation des niveaux de gris aux rectangles de dimensions $[N1, M1]$ de l'image1,
- à droite, le résultat du seuillage adaptatif appliqué sur l'image en niveaux de gris. C'est une image binaire, appelée image2.

On remarque que les silhouettes connexes constituant la forme dans image1 le sont encore dans image2, mais elles commencent à se détacher. Une deuxième procédure, analogue à celle développée dans la première étape de l'algorithme ASSESA, est utilisée pour individualiser les deux silhouettes de image2. Cette deuxième procédure est présentée ci-dessous.

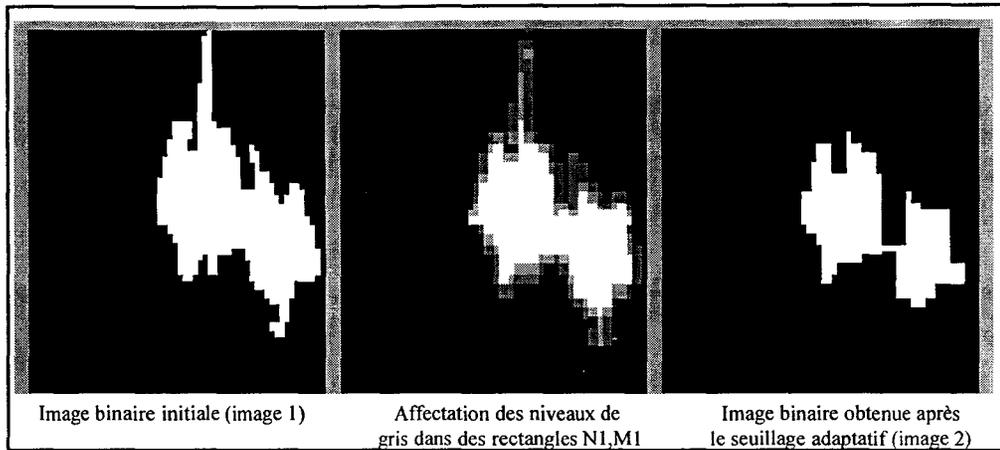


Figure 6.2 : exemple de traitement d'une image binaire en utilisant la première étape de l'algorithmme ASSESA

6.3 DEUXIEME ETAPE

6.3.1 Principe

L'image binaire appelée image2, résultat du seuillage adaptatif, est sous-échantillonnée pour obtenir une nouvelle image de résolution inférieure. Pour cela, on découpe cette image en rectangles R2 de taille fixe, N2 dans le sens de largeur et M2 dans celui de la longueur (Figure 6.3).

Sur la figure 6.3, on note que les rectangles R2 de image2 sont tous adjacents. L'image étant de taille $[N/N1, L/M1]$, $(N/N1)$ pixels et $(L/M1)$ lignes d'acquisitions, N2 est un diviseur de $(N/N1)$ et M2 un diviseur de $(L/M1)$.

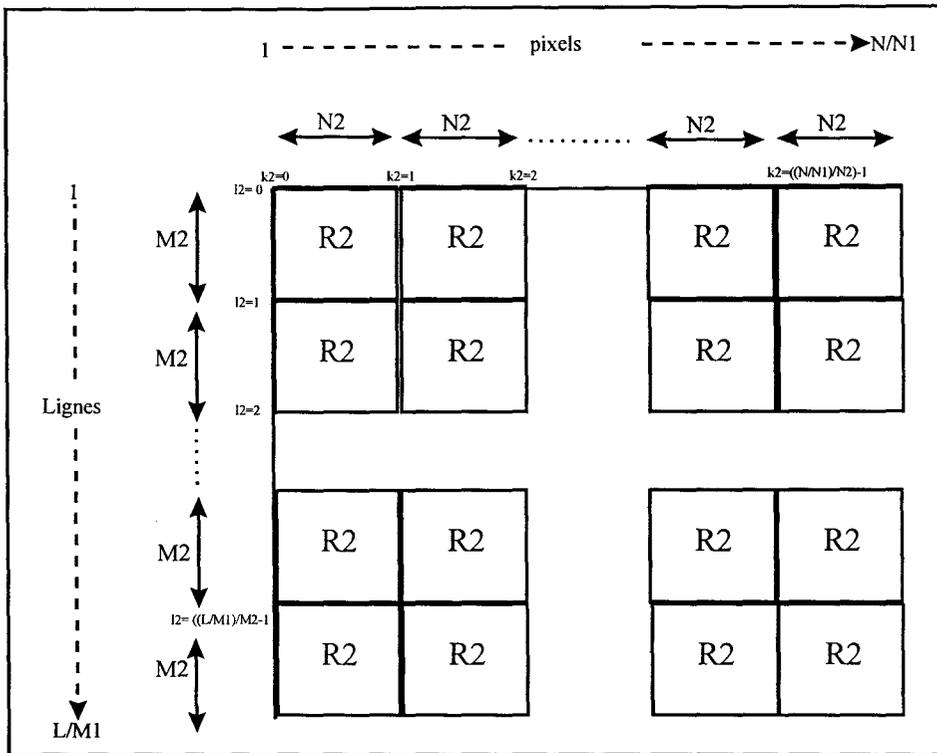


Figure 6.3 : sous-échantillonnage en rectangles R2 de image2

Dans chaque rectangle de l'image binaire initiale, appelée **image2**, le nombre de pixels de valeur 0 est comptabilisé ainsi :

$$NB_PIXEL_0[R2(k2,l2)] = NB_PIXEL_R2 - \sum_{i=1}^{N1} \sum_{j=1}^{M1} image2(k2 * N2 + i, l2 * M2 + j)$$

où $NB_PIXEL_0[R2(k2,l2)]$ est le nombre de pixels à 0 dans le rectangle R2 dont la position dans image2 est $(k2,l2)$.

NB_PIXEL_R2 est le nombre de pixels dans le rectangle R2 ($N2 \times M2$).

Après le seuillage adaptatif, image2 étant de taille $[N/N1, L/M1]$, $k2$ varie de 0 à $(N/N1)/N2 - 1$ et $l2$ de 0 à $(L/M1)/M2 - 1$.

$image2(i,j)$ est la valeur du pixel de coordonnées (i,j) dans image2 (0 ou 1).

Comme dans la première étape de l'algorithme ASSESA, à la valeur $NB_PIXEL_0[R2(k2,l2)]$ correspond un niveau de gris codé sur 8 bits, tel que :

$$NdG[R2(k2,l2)] = \frac{NB_PIXEL_0[R2(k2,l2)]}{NB_PIXEL_R2} \cdot 255$$

où $NdG[R2(k2,l2)]$ est le niveau de gris affecté au rectangle R2 dont la position dans image2 est $(k2,l2)$.

Cette opération d'affectation des niveaux de gris est effectuée pour chaque rectangle de l'image2. Nous obtenons à nouveau une image en niveaux de gris de résolution inférieure à celle de image2.

Cette image en niveaux de gris est ensuite binarisée. Un seuil de binarisation unique lui est appliqué. Tous les pixels de cette image dont le niveau de gris est inférieur au seuil sont mis à 1, les autres à 0.

6.3.2 Applications aux images

La figure 6.4 est l'illustration du traitement de image2 de la figure 6.2 selon la procédure décrite dans la deuxième étape de l'algorithme ASSESA. Cette figure est aussi divisée en 3 parties :

- l'image2 est présentée à gauche,
- au milieu, on observe une image en niveaux de gris résultant de l'affectation des niveaux de gris dans des rectangles $[N2, M2]$ dans image2,
- le résultat du seuillage global appliqué sur l'image en niveaux de gris est donné à droite sous la forme d'une nouvelle image binaire.

On remarque que les silhouettes qui commençaient à se détacher dans l'image2 sont maintenant individualisées car deux noyaux résultent du traitement complet de l'algorithme ASSESA.

Le décompte final des piétons consiste à associer un piéton à un noyau.

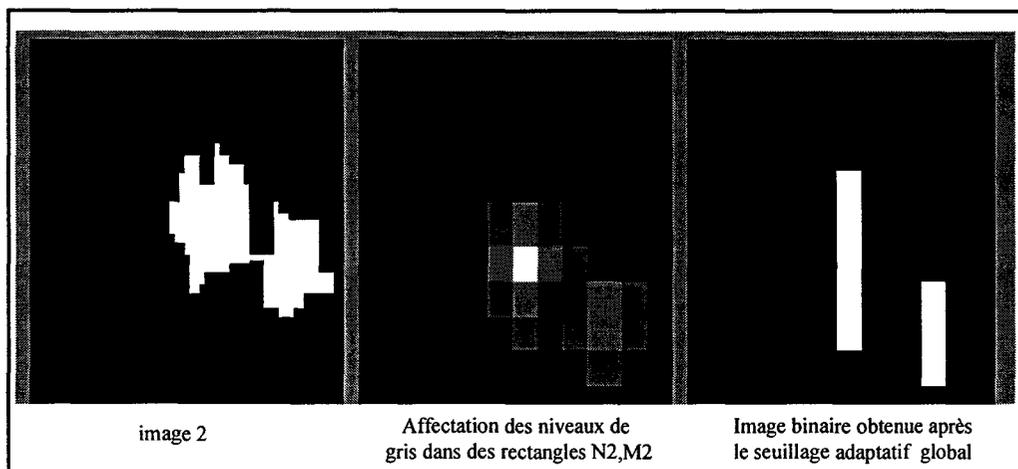


Figure 6.4 : exemple de traitement d'une image binaire en utilisant la deuxième étape de l'algorithme ASSESA

6.4 EFFET PROGRESSIF DE L'ALGORITHME

Dans les procédures de sous-échantillonnage et seuillage, on recherche, en plusieurs étapes, les "zones principales" ou noyaux des formes correspondant aux piétons.

Une première version de l'algorithme ASSESA effectuait un seul sous-échantillonnage, avec des tailles de rectangles supérieures à celles de $[N1, M1]$, suivi d'un seuillage adaptatif. Les résultats ont révélé des performances, en termes de comptage, très moyennes. Ces performances moyennes étaient dûes soit à une individualisation insuffisante des silhouettes par non séparation des formes, soit à l'élimination de silhouettes de petite taille après le seuillage. Nous avons alors développé une nouvelle version de l'algorithme qui traite les formes en deux étapes.

Les figures 6.2 et 6.4 montrent bien l'aspect progressif du traitement par l'algorithme. On a vu, sur ces deux exemples, que la forme de départ n'a pas été divisée après le premier seuillage mais que sa géométrie a changé, mettant en évidence des parties de cette forme qui commençaient à se détacher. Par contre après le deuxième seuillage, la forme est divisée, conduisant à deux noyaux, c'est à dire deux piétons.

6.5 CHOIX DE LA TAILLE DES RECTANGLES $[N1, M1]$ ET $[N2, M2]$

Dans les deux procédures de sous-échantillonnage, le choix des paramètres est très important et dépend beaucoup du site de mesures où le dispositif de comptage est installé. En effet, les contraintes géométriques du site de mesures influent sur la taille des formes à analyser.

Néanmoins, dans le premier sous-échantillonnage, $N1$ et $M1$ sont choisis assez petits, 5 à 6 pixels pour $N1$ et 3 à 4 lignes d'acquisition pour $M1$, afin qu'ils ne séparent pas immédiatement les formes. Une séparation trop rapide des formes peut conduire à individualiser des objets tels que des valises ou des sacs portés par les piétons. Leur taille permet un premier effet filtrant sur les formes, pour supprimer les détails comme les membres des piétons et les objets qu'ils portent. Une fois ces détails supprimés, le deuxième sous-échantillonnage, avec des valeurs de paramètres $N2$ et $M2$ plus élevées, sert à mettre en évidence les noyaux des formes traitées qui correspondent aux silhouettes qui les composent. Dans le deuxième sous-échantillonnage de l'algorithme, $N2$ respectivement $M2$ a été pris comme le quadruple de $N1$ respectivement $M1$.

6.6 INTERET DU SEUILLAGE ADAPTATIF

L'exemple de la figure 6.5 permet de montrer l'utilité et l'intérêt du seuillage adaptatif présenté dans la première étape de l'algorithme ASSESA.

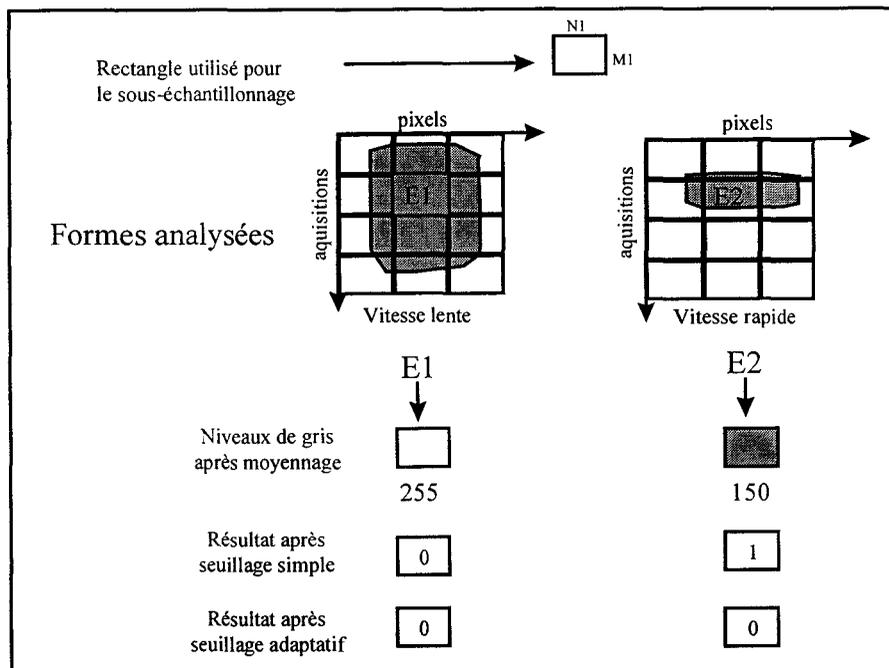


Figure 6.5 : exemple de seuillage adaptatif

Dans l'exemple de la figure 6.5, figurent deux images binaires contenant chacune une forme à traiter. La première forme, représentée dans le diagramme spatio-temporel de gauche, correspond à un piéton qui se déplace lentement (en classe 5 par exemple). La seconde, représentée dans le diagramme spatio-temporel de droite, correspond à un piéton qui se déplace rapidement (en classe 1 par exemple).

La procédure de sous-échantillonnage en rectangles $[N1, M1]$ de l'image binaire est appliquée et les niveaux de gris affectés. A l'intérieur de la première forme, on a repéré un rectangle par la lettre $E1$ alors qu'à l'intérieur de la deuxième, le rectangle repéré est appelé $E2$.

Après la procédure d'affectation des niveaux de gris sur les deux images binaires, on a affecté aux rectangles $E1$ et $E2$ leurs niveaux de gris respectifs en fonction du nombre de pixels à 0 situés à l'intérieur. On peut noter que l'élément $E1$ a un niveau de gris supérieur à celui de $E2$ car tous les pixels qui le composent sont à 0 contrairement à ceux de $E2$. Le

niveau de gris de E1 est donc égal au maximum, c'est à dire à 255 alors que l'on affecte à E2 un niveau de gris égal à 150.

Se pose maintenant le problème de la binarisation de ces rectangles en niveaux de gris. Ce que nous recherchons dans la procédure de binarisation, c'est à obtenir les zones principales des formes, c'est à dire leurs noyaux.

Si on binarise globalement toute l'image en niveaux de gris sans tenir compte des vitesses avec un seuil donné, le résultat après seuillage pourrait supprimer le rectangle E2 si son niveau de gris est sous la valeur du seuil.

On voit sur la figure 6.5 qu'avec un seuillage global, le rectangle E1 est mis à 0 car son niveau de gris est supérieur au seuil alors que E2 est mis à 1 car son niveau de gris est en dessous. On voit bien qu'un seuillage global élimine la forme correspondant à un déplacement rapide. Cela signifie que, dans le comptage final, on élimine un piéton.

La vitesse est une information très importante pour caractériser le passage des piétons dans les plans de surveillance. Nous utilisons ici cette information pour transformer le seuillage global en un seuillage qui dépend de la vitesse. Ce seuillage adaptatif, appliqué à des cas comme celui de la figure 6.5, permet de maintenir, pour le comptage final, les piétons qui se déplacent rapidement.

6.7 RESULTATS

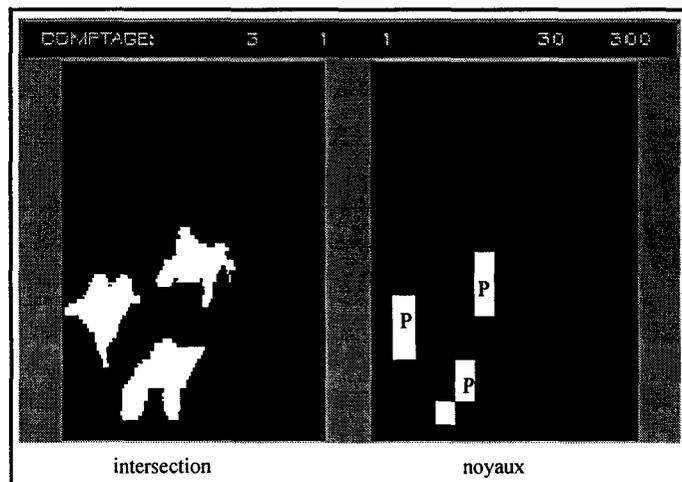
Les résultats du traitement des formes par l'algorithme ASSESA porte sur les séquences d'images présentées dans le chapitre IV. Ces résultats doivent être consultés en même temps que ces séquences.

Comme pour les traitements avec les algorithmes développés dans le chapitre V, on utilise pour le traitement par l'algorithme ASSESA, les images de l'intersection des formes d'origine obtenues avec les deux caméras.

Dans chaque séquence résultat, la partie gauche correspond à l'intersection des formes d'origine obtenues avec les deux caméras. La partie droite contient ce que l'on a appelé les noyaux qui sont les résultats du traitement par l'algorithme ASSESA. Dans le comptage final, un noyau est associé à un piéton.

Séquence 1

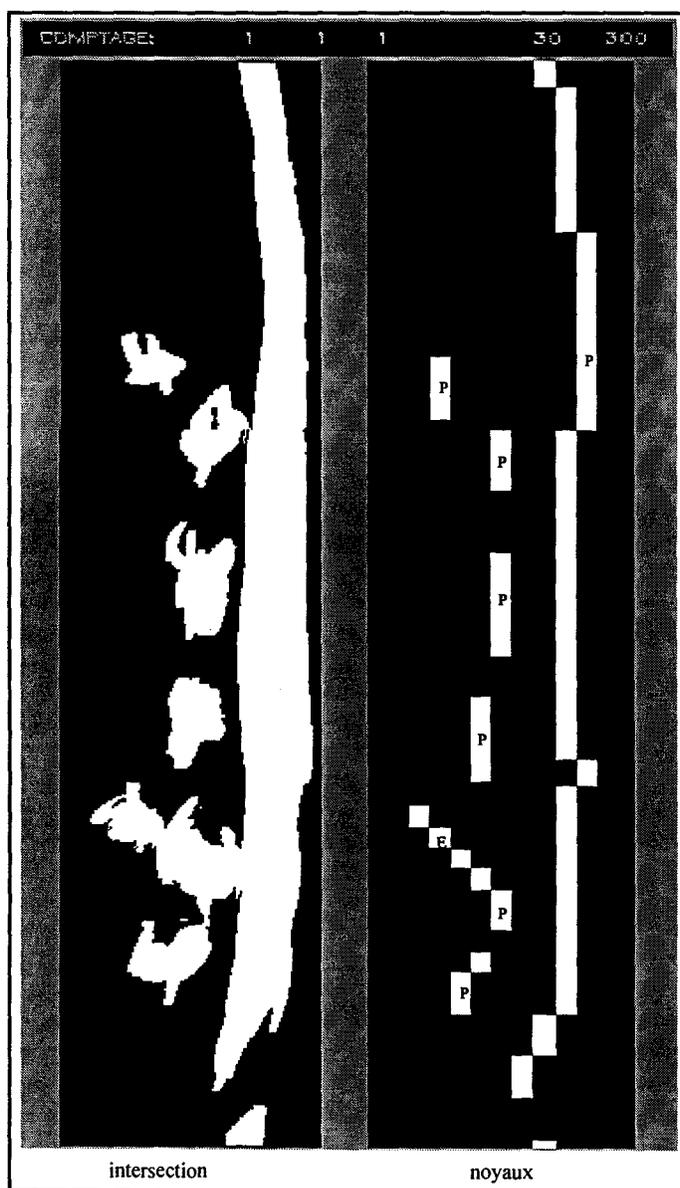
Cette séquence correspond à des piétons isolés. Les noyaux montrent que la géométrie des formes d'origine a été complètement modifiée par l'application des deux sous-échantillonnages des images binaires d'origine. Le problème de séparation des formes ne se pose pas pour cette séquence puisque les piétons sont isolés dans la séquence d'origine.



Séquence 1 : piétons isolés (P : piéton)

Séquence 2

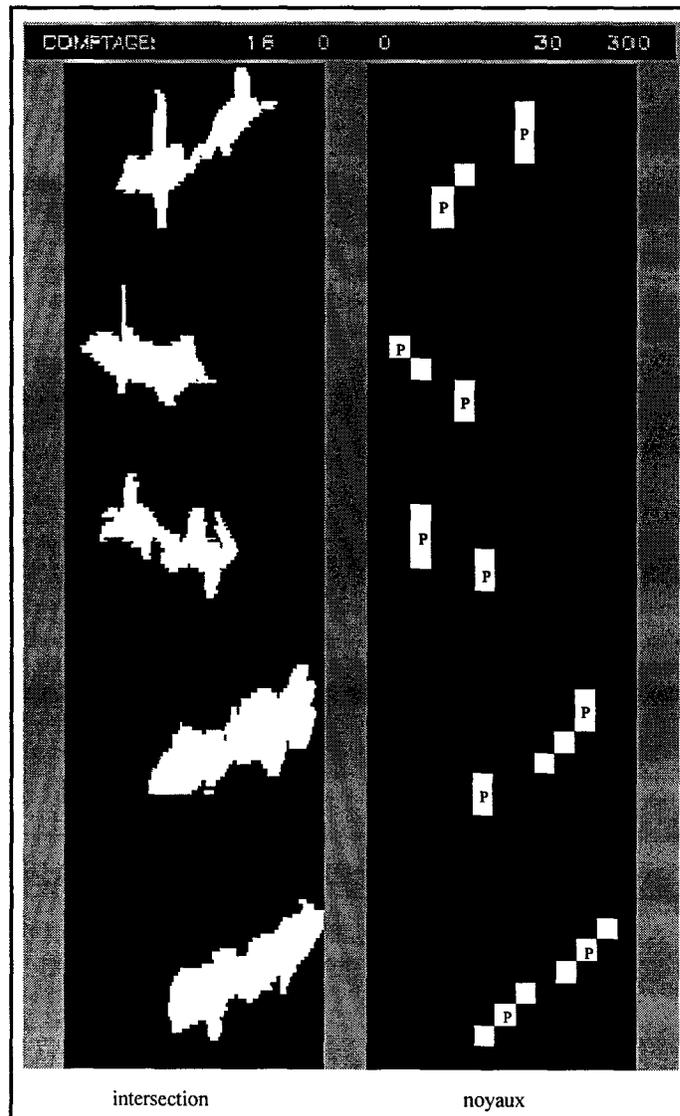
Tous les piétons relatifs aux formes de cette séquence ont été identifiés. La petite forme correspondant à l'enfant a été correctement individualisée par l'algorithme ASSESA, ce que n'ont pas réussi à faire les algorithmes inspirés de la morphologie mathématique.



Séquence 2 : piéton immobile (P : piéton, E : enfant)

Séquence 3

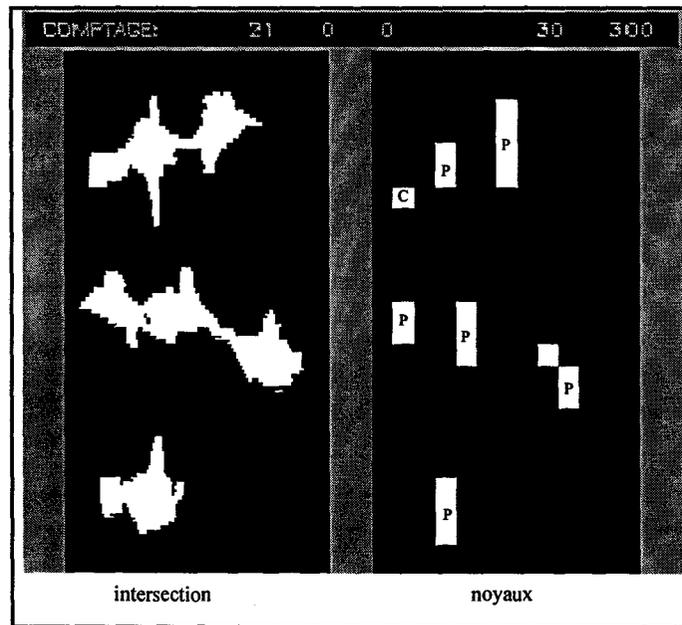
Cette séquence illustre plusieurs formes correspondant à deux silhouettes connexes. La séparation de ces formes a mis en évidence chaque piéton dans une forme isolée. Il faut noter que toutes les silhouettes connexes de la séquence ont été correctement individualisées alors que plusieurs d'entre elles restaient encore connexes, après l'application des algorithmes inspirés de la morphologie mathématique.



Séquence 3 : série de deux silhouettes connexes (P : piéton)

Séquence 4

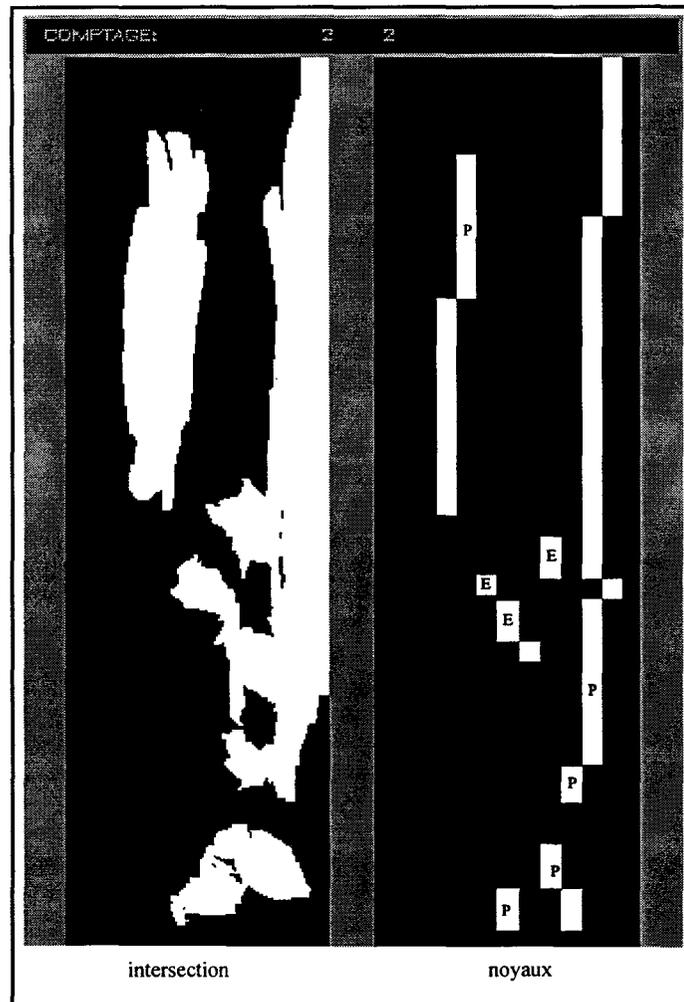
La séquence 4 comporte des piétons qui portent des objets. Avec l'algorithme ASSESA, le carton a été séparé du piéton qui le porte. Ce résultat a déjà été observé dans les résultats du traitement avec l'algorithme à élément structurant rectangulaire. Ceci est dû à la largeur du carton qui est proche de celle du piéton qui le porte. La valise, par contre, compte-tenu de sa largeur réduite, n'a pas été séparée. Le triplet de piétons, dont les silhouettes sont connexes, est correctement identifié.



Séquence 4 : piétons et objets (P : piéton, C : carton)

Séquence 5

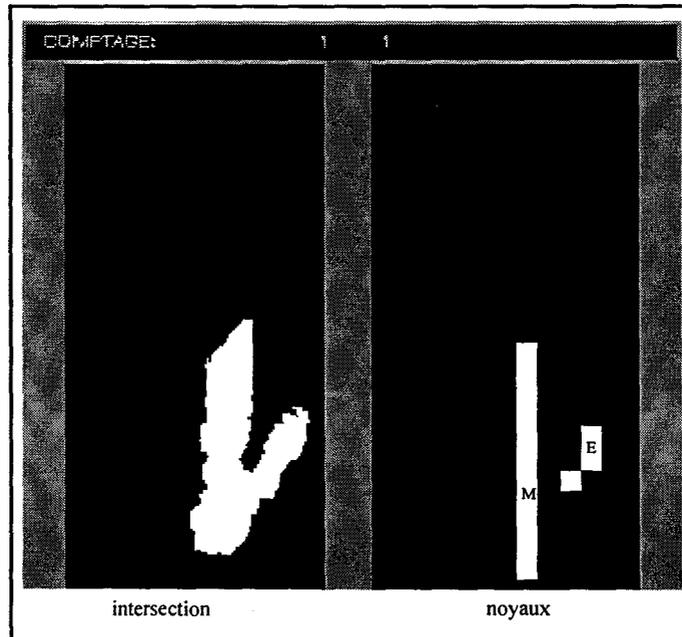
Tous les piétons de cette séquence dont les silhouettes sont connexes ont été correctement individualisés.



Séquence 5 : groupe d'enfants (P : piéton, E : enfant)

Séquence 6

Le résultat du traitement de cette séquence montre que la mère et l'enfant ont été individualisés.



Séquence 6 : piétons et objets particuliers

6.8 CONCLUSION

Dans ce chapitre, nous avons présenté un nouvel algorithme, que nous avons baptisé ASSESA pour Algorithme de Segmentation par Sous-Echantillonnage et Seuillage Adaptatif. Cet algorithme agit en plusieurs étapes sur les formes pour effectuer l'association piétons-formes.

Nous avons vu que cette approche avec deux résolutions par l'algorithme ASSESA est particulièrement appropriée au problème du comptage.

En terme de séparation de formes, comparé aux algorithmes inspirés de la morphologie mathématique, l'algorithme ASSESA est plus performant dans la mesure où il parvient à mieux individualiser les silhouettes connexes.

Nous avons vu aussi que le succès de cet algorithme réside dans sa capacité à rechercher les noyaux dans les formes.

Dans le chapitre suivant, sont présentés les résultats de comptage pour les algorithmes développées dans les chapitres V et VI.

Chapitre VII

Les résultats du comptage

Chapitre VII

LES RESULTATS DU COMPTAGE

7.1 CONTEXTE

Dans ce chapitre, on présente les résultats d'évaluation sur un site réel des algorithmes présentés dans les chapitres V et VI. Les résultats présentés incluent les performances :

- en termes de comptage pour les différents algorithmes présentés dans les chapitres V et VI et sur tous les échantillons d'images recueillies,
- en termes de temps de traitement.

7.1.1 L'environnement de l'expérimentation

L'enregistrement de séquences d'images linéaires s'est déroulé au Parc des Expositions de Villepinte, au nord de Paris, pendant le Salon SIPPA-bureautique les 11 et 12 janvier 1996. Le système de transport filoguidé SK (Soulé-Kermadec) assure la desserte interne du parc. Ce système de transport est exploité avec une vingtaine de véhicules sur une ligne de quelques centaines de mètres comportant uniquement 2 stations terminales.

Les véhicules de ce système peuvent contenir une quinzaine de personnes et sont mus par un système à câble avec un intervalle entre véhicules d'environ 30 secondes.

L'expérimentation dans ce site a duré deux journées.

Le système de comptage a été installé près de l'une des stations afin de détecter toutes les personnes entrant ou sortant de cette station.

Les piétons détectés traversaient un passage de 2,70 mètres de large si bien qu'à peu près 4 personnes pouvaient passer en même temps sous le système (Figure 7.1).

Les prises de vues ont été réalisées en extérieur avec des conditions climatiques très défavorables puisque la scène était très ensoleillée. De plus, compte tenu du contexte dans lequel les mesures ont été effectuées, à savoir un parc d'expositions un jour de salon, la majorité des piétons détectés portaient des objets, des valises, des cartons en ce qui concerne les exposants et des sacs remplis de documentation pour les visiteurs.

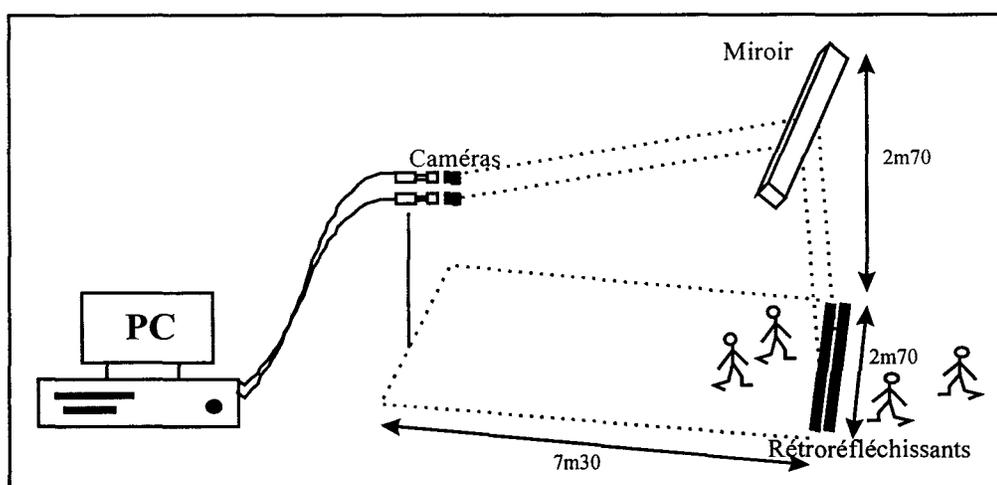


Figure 7.1: environnement de l'expérience

7.1.2 Le matériel utilisé

Nous avons utilisé 2 caméras linéaires du type IDC100, chacune étant connectée à une carte d'acquisition du type ISM197. Ces 2 cartes sont installées dans un PC à microprocesseur 486 DX II 66.

Les 2 caméras sont rendues actives par l'adjonction du système de LED infrarouge puissant permettant notamment de s'affranchir des éclairages ambiants souvent gênants. De plus, un filtre est inséré entre l'élément photosensible de la caméra et l'objectif.

La prise de vue est réalisée par l'intermédiaire d'un miroir placé au-dessus de la scène à observer et positionné à 45° afin de créer les 2 plans verticaux de détection souhaités tout en installant les caméras horizontalement à quelques mètres de la scène (Figure 7.1). L'intérêt

majeur d'une telle installation est de permettre l'implantation du système sans difficulté, quelle que soit la hauteur de plafond disponible.

Un revêtement adhésif rétroréfléchissant (3M) est collé au sol. Ce revêtement est composé de microprismes renvoyant la lumière émise par les LED dans un angle très fermé, de moins de 1 degré. Ceci oblige un positionnement des LED le plus proche possible de l'axe optique de la caméra.

Les 2 caméras sont montées sur un support mécanique permettant un réglage fin de la distance séparant les deux plans de surveillance et un réglage de leur parallélisme.

7.2 PERFORMANCES DES ALGORITHMES INSPIRES DE LA MORPHOLOGIE MATHÉMATIQUE

7.2.1 Les paramètres généraux

Une fois le système installé sur un site, un certain nombre de paramètres doivent être déterminés préalablement avant de procéder au comptage. En effet, ce sont les caractéristiques du lieu d'expérimentation comme la hauteur sous plafond, la distance des caméras au miroir, la distance focale choisie qui vont fixer les valeurs de ces paramètres qui vont déterminer les caractéristiques des formes obtenues. En effet, si les caméras sont fixées à un plafond bas, 2m50 par exemple, qui est un cas courant, les formes correspondant au passage des piétons détectés seront plutôt larges. A l'inverse si les caméras sont installées très haut, à 3 ou 4 mètres, les formes correspondant aux mêmes piétons seront plus petites.

Il y a ainsi un certain nombre de réglages de paramètres pour l'initialisation du système dans un site donné.

Le tableau 7.1 fournit la liste et un descriptif de ces paramètres. Parmi eux, citons les plus importants :

- NBPIX : sur les 1024 pixels des barrettes CCD des deux caméras, il est possible d'effectuer un sous-échantillonnage spatial pour n'en utiliser qu'une partie. Pour augmenter la vitesse de traitement, nous avons progressivement testé les algorithmes en diminuant à chaque fois le nombre de pixels utilisés. Nous avons pu diminuer le nombre de pixels traités jusqu'à 128 sans que les

résultats du comptage en soient affectés. Le choix des 128 pixels sur les 1024 est une simple sélection d'un pixel sur 8,

- La largeur des éléments structurants utilisés,
- les niveaux de vitesses, déterminées en nombre d'acquisitions et en mètres par seconde, qui dépendent du site d'expérimentation,
- la longueur des éléments structurants en fonction des niveaux de vitesses.

Morphologie				Élément structurant				
Nom du paramètre	Définition	Valeur	Unité	Rectangulaire	Ellipsoïdal	directionnel	Valeur	Unité
NBPIX	Nombre de pixels utilisés sur les 1024 de la barrette CCD	128	pixels	X	X	X		
MASQSIZE	Largeur de l'élément structurant utilisé	7	pixels	X	X	X		
VIT1	Classe de vitesse 1	[1,4]	acquisitions	X	X	X	[10 ; 2,5]	m/s
VIT2	Classe de vitesse 2	[5,8]	acquisitions	X	X	X	[2 ; 1,25]	m/s
VIT3	Classe de vitesse 3	[9,12]	acquisitions	X	X	X	[1,11 ; 0,83]	m/s
VIT4	Classe de vitesse 4	[13,16]	acquisitions	X	X	X	[0,76 ; 0,62]	m/s
VIT5	Classe de vitesse 5	[17 et +]	acquisitions	X	X	X	[0,58 ; 0]	m/s
PROF1	Correspondance entre la profondeur de l'élément structurant et la classe de vitesse 1	1	ligne	X	X	X		
PROF2	Correspondance entre la profondeur de l'élément structurant et la classe de vitesse 2	3	lignes	X	X	X		
PROF3	Correspondance entre la profondeur de l'élément structurant et la classe de vitesse 3	5	lignes	X	X	X		
PROF4	Correspondance entre la profondeur de l'élément structurant et la classe de vitesse 4	7	lignes	X	X	X		
PROF5	Correspondance entre la profondeur de l'élément structurant et la classe de vitesse 5	9	lignes	X	X	X		
CHAMP	Profondeur maximum de l'élément structurant	9	lignes	X	X	X		
DEGREE	Nombre de directions de mouvement envisagées	6				X		
SEGMIN	Nombre de pixels minimum pour un objet considéré	15	pixels	X	X	X		
LMOYEN	Nombre de lignes au delà duquel la procédure de suivi est effectuée	100	lignes	X	X	X		
LMINCOUNT	Nombre de lignes minimum pour un objet	2	lignes	X	X	X		
LMINPOINT	Nombre de pixels minimum pour une ligne d'un objet	3	pixels	X	X	X		

Tableau 7.1 : paramètres des algorithmes s'inspirant de la morphologie mathématique

7.2.2 Élément structurant rectangulaire dynamique

7.2.2.1 Les résultats globaux

Les résultats obtenus avec l'algorithme à élément structurant rectangulaire figurent dans le tableau 7.2. Parmi les 3941 piétons dénombrés avec un comptage manuel, plusieurs densités de piétons sont représentées. La base de donnée n'est pas représentative de toutes les conditions qu'il est possible de rencontrer dans la réalité, mais grâce à cette expérimentation un certain nombre de difficultés de traitement des formes ont été mises en évidence : piétons statiques sous les caméras, couples de piétons dont les silhouettes sont connexes, personnes portant des sacs, des valises, groupes de personnes.

Les algorithmes ont été testés sur un certain nombre d'échantillons de piétons représentant chacun une certaine densité de piétons. Par exemple, l'échantillon 1 représente un niveau de fréquentation faible, c'est à dire que les piétons traversaient isolément les plans de surveillance alors que l'échantillon 10 correspond plutôt à des piétons passant en groupes compacts sous les caméras.

Sur le tableau 7.2, nous avons porté, pour chaque échantillon, le résultat d'un comptage manuel effectué grâce à une caméra vidéo servant de témoin, le comptage automatique réalisé avec l'algorithme et l'erreur locale à l'échantillon. Soit EL cette erreur locale. Elle est calculée ainsi :

$$(EL)_{\text{éch}(i)} = ((\text{comptage automatique})_{\text{éch}(i)} - (\text{comptage manuel})_{\text{éch}(i)} * 100) / (\text{comptage automatique})_{\text{éch}(i)}$$

où éch(i) représente l'échantillon i.

Si l'on considère tous les échantillons pris ensemble, 3917 piétons ont été comptabilisés automatiquement par l'algorithme à élément structurant rectangulaire alors que 3941 sont en réalité passés sous les caméras, ce qui conduit à une erreur globale de -0.6%. Cette erreur globale est obtenue en comparant les deux totaux correspondant au comptage manuel et au comptage automatique.

Numéro d'échantillon	Comptage manuel	Comptage automatique	Erreur locale(%)
1	160	159	-0,63
2	277	271	-2,21
3	210	207	-1,45
4	202	197	-2,54
5	174	174	0,00
6	228	225	-1,33
7	405	406	0,25
8	452	449	-0,67
9	495	496	0,20
10	640	639	-0,16
11	137	138	0,72
12	153	154	0,65
13	212	212	0,00
14	196	190	-3,16
Somme	3941	3917	
Erreur globale (%)	-0,61		

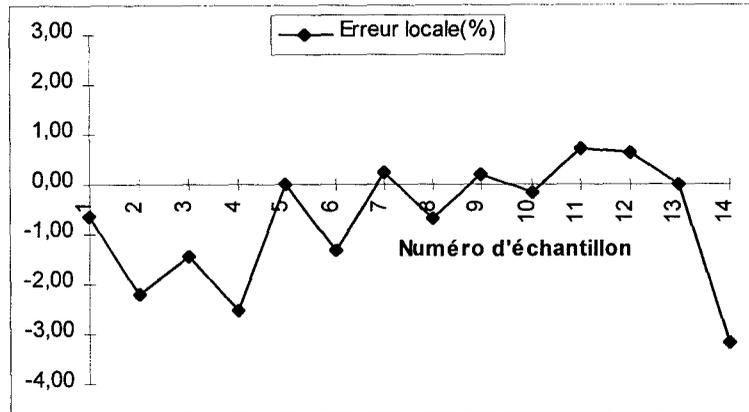
Tableau 7.2 : résultats globaux pour l'élément structurant rectangulaire

7.2.2.2 L'analyse des erreurs

L'analyse des erreurs locales représentées sur le graphique 7.1 montre que dans de nombreux cas, l'algorithme a tendance à sous-compter. Ceci est dû aux couples de piétons qui ont généré des silhouettes connexes difficiles à séparer.

Une forte erreur négative (-3.16%) est observée pour l'échantillon 14, mais elle est calculée sur un échantillon réduit.

Il faut noter aussi que pour les rares cas où un sur-comptage est observé, l'erreur positive ne dépasse pas 1% et est liée dans la plupart des cas à un comptage d'objets, larges cartons par exemple que nous avons déjà vus dans les séquences présentées aux chapitres V et VI.



Graphique 7.1 : erreurs locales pour l'élément structurant rectangulaire

7.2.3 Élément structurant ellipsoïdal dynamique

7.2.3.1 Les résultats globaux

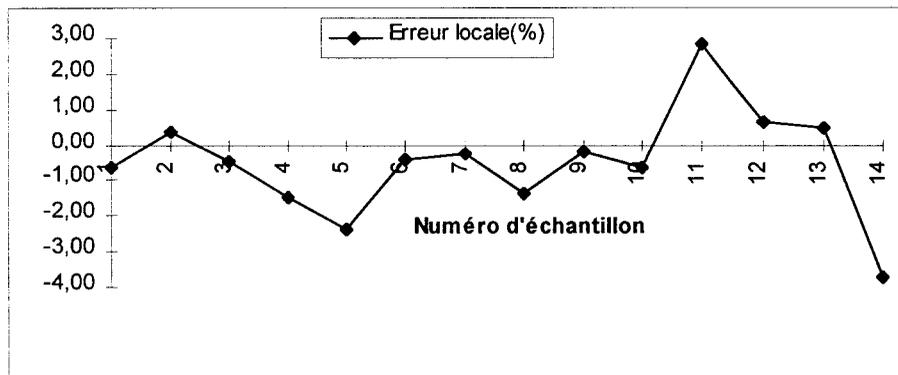
Nous obtenons une erreur globale de -0.56% pour l'algorithme à élément structurant ellipsoïdal. Cette erreur est obtenue en comparant les 3941 piétons comptés manuellement et les 3919 comptés par l'algorithme. Cet algorithme présente les mêmes performances en termes de comptage que celui à élément structurant rectangulaire (Tableau 7.3).

Numéro d'échantillon	Comptage manuel	Comptage automatique	Erreur locale (%)
1	160	159	-0,63
2	277	278	0,36
3	210	209	-0,48
4	202	199	-1,51
5	174	170	-2,35
6	228	227	-0,44
7	405	404	-0,25
8	452	446	-1,35
9	495	494	-0,20
10	640	636	-0,63
11	137	141	2,84
12	153	154	0,65
13	212	213	0,47
14	196	189	-3,70
Somme	3941	3919	
Erreur globale (%)	-0,56		

Tableau 7.3 : résultats globaux pour l'élément structurant ellipsoïdal

7.2.3.2 L'analyse des erreurs

L'analyse des erreurs locales (Graphique 7.2) montre que la seule différence entre l'algorithme à élément ellipsoïdal et celui à élément rectangulaire concerne l'échantillon 11 pour lequel la différence de comptage entre les deux est de 3 piétons (138 pour le rectangulaire contre 141 pour l'ellipsoïdal).



Graphique 7.2 : erreurs locales pour l'élément structurant ellipsoïdal

7.2.4 Élément structurant directionnel dynamique

7.2.4.1 Les résultats globaux

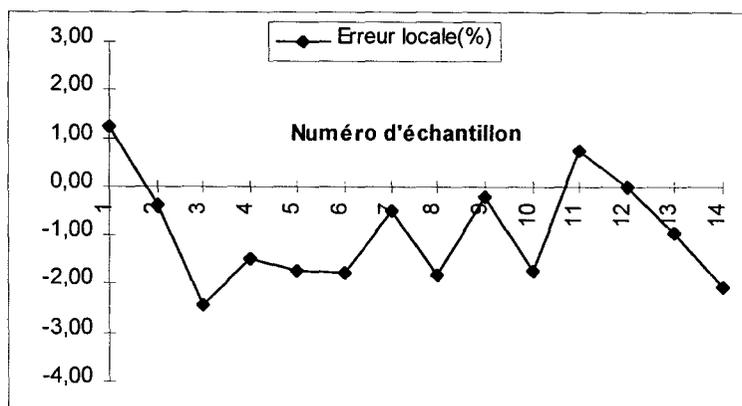
Pour ce dernier algorithme inspiré de la morphologie mathématique, l'erreur globale moyenne est un peu plus élevée, -1.05% sur le tableau 7.4, que pour les algorithmes précédents. Cet algorithme, même s'il inclut dans les traitements les directions de mouvement des piétons, contribue moins fortement à la séparation des formes en silhouettes.

Numéro d'échantillon	Comptage manuel	Comptage automatique	Erreur locale(%)
1	160	162	1,23
2	277	276	-0,36
3	210	205	-2,44
4	202	199	-1,51
5	174	171	-1,75
6	228	224	-1,79
7	405	403	-0,50
8	452	444	-1,80
9	495	494	-0,20
10	640	629	-1,75
11	137	138	0,72
12	153	153	0,00
13	212	210	-0,95
14	196	192	-2,08
Somme	3941	3900	
Erreur globale (%)	-1,05		

Tableau 7.4 : résultats globaux pour l'élément structurant directionnel

7.2.4.2 L'analyse des erreurs

Le pouvoir de séparation des formes en silhouettes de l'algorithme à élément structurant directionnel étant un peu moins fort que celui des deux précédents algorithmes, on observe un sous-comptage pour la majorité des échantillons (Graphique 7.3).



Graphique 7.3 : erreurs locales pour l'élément structurant directionnel

7.2.5 Aspects temps réel

Le temps de traitement est un facteur très important lorsqu'on envisage de développer un système devant fonctionner en temps réel. Les trois algorithmes, qui viennent d'être présentés, permettent de traiter 170 paires d'images lignes. Cette capacité de traitement ne concerne que le temps utilisé par les algorithmes. A cela il faut ajouter trois temps intermédiaires :

- le temps d'acquisition des images lignes par les caméras,
- le temps de stockage en mémoire des images lignes par les cartes d'acquisition,
- le temps de binarisation des images lignes.

Ces temps intermédiaires pris en compte, le système complet fonctionne, en temps réel, en traitant 150 paires de lignes par seconde, c'est à dire avec une période d'acquisition de 7.5 millisecondes.

7.3 RESULTATS DE L'ALGORITHME ASSESA

7.3.1 Les paramètres utilisés

De la même manière que pour les algorithmes inspirés de la morphologie mathématique, un certain nombre de paramètres doivent être définis au préalable pour l'algorithme ASSESA (Tableau 7.5).

On note également que 128 pixels seulement sont utilisés pour le traitement.

Algorithme ASSESA					
Nom du paramètre	Définition	Valeur	Unité	Valeur	unité
NBPIX	Nombre de pixels utilisés sur les 1024 de la barrette CCD	128	pixels		
SEGX1	Largeur rectangle pour 1er moyennage	8	pixels		
SEGY1	Profondeur rectangle pour 1er moyennage	4	lignes		
SEGX2	Largeur rectangle pour 2ème moyennage	16	pixels		
SEGY2	Profondeur rectangle pour 2ème moyennage	8	lignes		
VIT1	Classe de vitesse 1	[1,4]	acquisitions	[10 ; 2,5]	m/s
VIT2	Classe de vitesse 2	[5,8]	acquisitions	[2 ; 1,25]	m/s
VIT3	Classe de vitesse 3	[9,12]	acquisitions	[1,11 ; 0,83]	m/s
VIT4	Classe de vitesse 4	[13,16]	acquisitions	[0,76 ; 0,62]	m/s
VIT5	Classe de vitesse 5	[17 et +]	acquisitions	[0,58 ; 0]	m/s
VAL1	Niveau de seuillage entre classe de vitesse 1 et seuillage	50	Niveau de gris		
VAL2	Niveau de seuillage entre classe de vitesse 2 et seuillage	100	Niveau de gris		
VAL3	Niveau de seuillage entre classe de vitesse 3 et seuillage	150	Niveau de gris		
VAL4	Niveau de seuillage entre classe de vitesse 4 et seuillage	200	Niveau de gris		
VAL5	Niveau de seuillage entre classe de vitesse 5 et seuillage	250	Niveau de gris		

Tableau 7.5 : paramètres de l'algorithme ASSESA

7.3.2 Les résultats globaux

On note, dans les résultats de comptage de l'algorithme ASSESA, une erreur globale moyenne -0.48 % (Tableau 7.6). Ceci est un résultat satisfaisant compte tenu de la difficulté que présente le traitement de certaines formes dans les échantillons de l'expérimentation. Nous avons vu dans les séquences présentées dans les chapitres V et VI que l'algorithme ASSESA sépare progressivement les formes en silhouettes.

L'erreur globale moyenne proche de 0 s'explique par le fait que, dans les échantillons, les valeurs des erreurs locales se distribuent bien autour de la valeur 0.

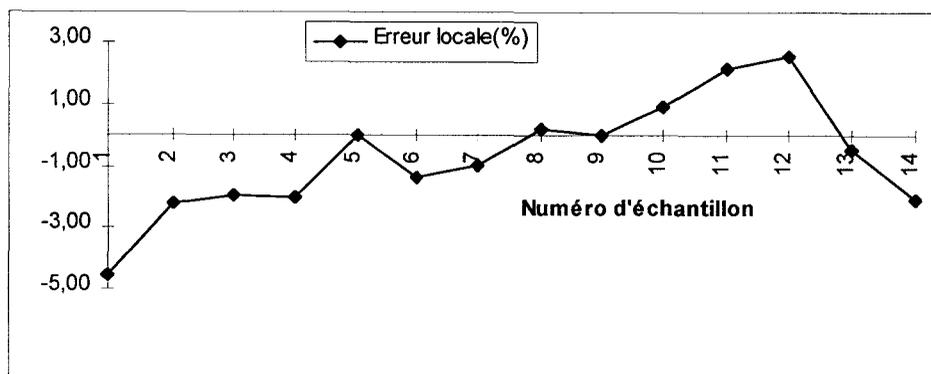
Numéro d'échantillon	Comptage manuel	Comptage automatique	Erreur locale(%)
1	160	153	-4,58
2	277	271	-2,21
3	210	206	-1,94
4	202	198	-2,02
5	174	174	0,00
6	228	225	-1,33
7	405	401	-1,00
8	452	453	0,22
9	495	495	0,00
10	640	646	0,93
11	137	140	2,14
12	153	157	2,55
13	212	211	-0,47
14	196	192	-2,08
Somme	3941	3922	
Erreur globale (%)	-0,48		

Tableau 7.6 : résultats globaux pour l'algorithme ASSESA

7.3.3 L'analyse des erreurs

L'analyse des erreurs locales (Graphique 7.4) montre que leurs valeurs sont toutes très proches de 0, à l'exception de celle calculée sur le premier échantillon et qui avoisine 4%. Il faut toutefois noter qu'elle est calculée sur un effectif de 160 piétons seulement. Deux aspects marquent bien la différence entre cet algorithme et les précédents :

- l'erreur globale moyenne est la plus faible,
- les erreurs locales ne révèlent pas systématiquement du sous-comptage, ce qui était le cas pour les algorithmes précédents.



Graphique 7.4 : erreurs locales pour l'algorithme ASSESA

7.3.4 Aspects temps réel

L'algorithme ASSESA qui fournit de meilleurs résultats en termes de comptage que les algorithmes précédents, est aussi plus rapide en temps de calcul. Ainsi, il est possible de traiter 250 paires de lignes par seconde. Si on ajoute les temps d'acquisition, le temps de mise en mémoire des lignes par les cartes d'acquisition et le temps de binarisation, on traite 230 paires de lignes par seconde.

7.4 COMPARAISON DES DEUX TYPES D'ALGORITHMES

Les deux types d'algorithmes, ASSESA et inspirés de la morphologie mathématique, peuvent être comparés selon deux critères :

- la vitesse de traitement,
- la précision du comptage.

En ce qui concerne le premier point, les deux types d'algorithmes ont été comparés sur les mêmes séquences et l'on a pu observer que l'algorithme ASSESA présente des performances meilleures en termes de vitesse de traitement. Sur les séquences de test, il est à peu près 1 fois et demi plus rapide avec une capacité de traitement de 230 paires de lignes par seconde contre 150.

Même si les deux types d'algorithmes sont exécutés avec la même résolution de 128 pixels par ligne d'acquisition, ceux inspirés de la morphologie sont moins rapides car ils font une analyse systématique de tous les pixels des séquences d'images lignes traitées. Ceci n'est pas le cas de l'algorithme ASSESA qui diminue progressivement cette résolution en découpant les images binaires d'origine en rectangles dans lesquels des calculs globaux sont effectués. Les vitesses de traitement des algorithmes inspirés de la morphologie sont moins rapides car ils incluent aussi les traitements de séparation directe par la vitesse et la procédure de suivi (chapitre V).

Sur le second point, la précision de comptage, l'algorithme ASSESA grâce, à sa capacité de bien individualiser les silhouettes contenues dans les formes, est plus performant en termes d'association piétons-formes. Par rapport aux autres algorithmes aussi, c'est le seul à ne pas présenter quasi-systématiquement du sous-comptage.

7.5 CONCLUSION

Dans ce chapitre, nous avons décrit une expérimentation en vraie grandeur pour l'évaluation de deux types d'algorithmes : ASSESA et inspirés de la morphologie mathématique.

Les deux types d'algorithmes ont été testés et comparés selon le même plan d'expérience.

Les résultats montrent que l'algorithme ASSESA offre de meilleures performances en termes de comptage et de rapidité d'exécution, même si avec les algorithmes inspirés de la morphologie mathématique, on obtient des résultats tout à fait honorables.

Nous avons vu aussi qu'une fois le système installé sur un site, un certain nombre de paramètres doivent être déterminés préalablement au comptage. Ces paramètres sont soit relatifs à la géométrie des lieux d'expérimentation, soit interviennent directement dans le fonctionnement interne des algorithmes.

Chapitre VIII

Conclusion et perspectives

Chapitre VIII

CONCLUSION ET PERSPECTIVES

Nous avons présenté, dans ce mémoire de thèse, un système de comptage automatique de piétons par vision artificielle linéaire. Après avoir défini les fonctions du système envisagé ainsi que son cahier des charges dans le chapitre I, une brève bibliographie sur les systèmes de comptage a été présentée dans le chapitre II. Dans le chapitre III, nous avons vu que grâce au dispositif développé, il est possible de disposer de trois grandeurs bien différenciées pour procéder au comptage :

- l'information spatiale qui correspond à l'axe défini par l'orientation des barrettes CCD,
- l'information temporelle liée à la fréquence d'acquisition des lignes,
- l'information vitesse obtenue grâce aux deux plans de surveillance.

Ensuite dans le chapitre IV, nous avons abordé le thème des prétraitements des données de base et notamment le filtrage des vitesses brutes. Le filtre conditionnel développé a permis de définir des vitesses homogènes pour les piétons.

Dans les chapitres V et VI, dédiés à la description des algorithmes de comptage, quatre méthodes ont été proposées. Les trois premières s'inspirent de la morphologie mathématique alors que la quatrième, ASSESA, utilise des procédures de sous-échantillonnages et de seuillages successives.

Les quatre algorithmes ont été évalués et comparés sur des séquences types choisies et l'on a pu noter les performances de chacun d'entre eux dans la séparation des formes pour l'individualisation des silhouettes qui les composent. Dans les résultats des séquences de test, on a pu observer la forte capacité de l'algorithme ASSESA dans la séparation des formes.

Dans le chapitre VII, nous avons décrit une expérimentation en vraie grandeur pour l'évaluation de deux types d'algorithmes : ASSESA et inspirés de la morphologie mathématique.

Compte tenu des résultats obtenus, le système de comptage répond aux exigences du cahier des charges défini en introduction de ce mémoire.

Les résultats montrent aussi que l'algorithme ASSESA offre de meilleures performances en termes de comptage et de rapidité d'exécution, même si avec les algorithmes inspirés de la morphologie mathématique, on obtient des résultats tout à fait satisfaisants.

Nous avons vu aussi qu'une fois le système installé sur un site, un certain nombre de paramètres doivent être ajustés préalablement au comptage.

Dans toute la technologie du système présenté, nous avons toujours mentionné l'utilisation de ce système avec deux caméras linéaires. Nous avons dit en début de ce mémoire que les caméras linéaires ont été préférées aux matricielles grâce à leur définition supérieure et à la vitesse de traitement qu'elles permettent. Cependant, nous avons vu, que pour la plupart des traitements présentés dans les chapitres V et VI, la résolution de 1024 pixels n'est pas nécessaire et que pour certains traitements 128 pixels suffisent.

Nous avons alors développé un autre système de comptage utilisant une caméra matricielle 512x512. Dans les images en deux dimensions obtenues, deux lignes d'acquisition sont sélectionnées à une période ΔT donnée. Ceci permet de simuler les deux capteurs linéaires des deux caméras linéaires du système précédent. Ce deuxième système présente plusieurs avantages :

- il permet de réduire le coût du système en ne nécessitant qu'une caméra matricielle, au lieu de deux caméras linéaires plus coûteuses,
- il permet de réduire l'encombrement du système sur le site dès lors qu'un recul suffisant permet de disposer de deux plans de surveillance quasiment parallèles,
- il permet de garder exactement les mêmes algorithmes de comptage utilisés dans le cas des deux caméras linéaires.

Dans cette optique et avec ce nouveau système, une seconde évaluation du système de comptage a été menée sur un autre site réel, dans un couloir très fréquenté à l'intérieur du métro de

Lille. L'algorithme ASSESA a été testé pendant plusieurs journées et pendant des périodes horaires très différentes : heures de pointe du matin et du soir, pointes de midi, heures creuses. Ainsi l'évaluation de l'algorithme ASSESA a été menée sur un échantillon global d'environ 30 000 personnes. Les résultats obtenus sur le site du métro de Lille montrent, toutes périodes confondues, qu'une précision de l'ordre du pourcent est obtenue en ce qui concerne le nombre de piétons comptés.

Devant l'intérêt des exploitants de transports publics pour ce système et des performances obtenues, une nouvelle action de recherche qui va plus loin que le comptage des passagers sera initiée en 1997.

En collaboration avec Transpole, exploitant des transports dans la métropole Lilloise, le comptage des piétons sera réalisé aux entrées et sorties des rames du métro VAL à l'aide de capteurs optiques. De plus, grâce aux méthodes qui vont être développées, les signatures des piétons, obtenues à l'aide d'un certain nombre de caractéristiques morphologiques, seront définies. Ceci devrait permettre de calculer à terme des matrices d'origine-destinations de déplacements des piétons dans l'enceinte du métro.

Un thésard travaille sur le sujet depuis début octobre, pour une durée de 3 ans, en collaboration avec le Centre d'Automatique de Lille (CAL).

Liste Bibliographique

Références bibliographiques

- [BRU 94] J.L. bruyelle, "Conception et réalisation d'un dispositif de prise de vue stéréoscopique linéaire - Application à la détection d'obstacles à l'avant des véhicules routiers", Thèse de doctorat, Université des sciences et technologies de Lille Flandres Artois, décembre 1994.
- [BUI 89] C. Buil, *Astronomie CCD - Construction et utilisation des caméras CCD en astronomie amateur*, Ed. Société d'Astronomie Populaire, Toulouse, 1989.
- [BUR 95] J.C. Burie, "Mise en correspondance d'images linéaires stéréoscopiques - Application à la détection d'obstacles à l'avant des véhicules routiers", Thèse de doctorat, Université des sciences et technologies de Lille Flandres Artois, septembre 1995.
- [CAR 96] J. Caron, L. Duvieubourg, J.J. Orteu, Ph. Révolte, "Real-time inspection of preweathered zinc based on computer vision", *Proceedings SPIE*, 29-30 January 1996, San-Jose California USA.
- [COH 90] S. Cohen, "INGENIERIE DU TRAFIC ROUTIER : Eléments de théorie du trafic et applications" Editeur presses de l'Ecole Nationale des Ponts et Chaussées , Novembre 1990.
- [COS 89] M. Coster, J.L. Chermant, "Précis d'analyse d'images", Presses du CNRS 1989.
- [COU 81] A. Couzy, "La télédétection", Coll. Que sais-je? Presses universitaires de France, Paris, 1981.

-
- [DEP 96/1] J.P. Deparis, S.A. Velastin, A.C. Davies, "Project CROMATICA", 5th International Conference on Automated People Movers , APM 96M.A, pp.227-239, Paris, June 1996.
- [DEP 96/2] J.P. Deparis, S.A. Velastin, A.C. Davies, "Telematic tools to improve quality of service in Public Transport. Example of an EC research project : CROMATICA", IBC European rail technology, London, June 1996.
- [DUV 91] L. Duvieubourg, "Analyse de séquences d'images linéaires stéréoscopiques - Application à la réalisation d'un système de détection d'intrusions intelligent pour les transports guidés", Thèse de doctorat, Université des sciences et technologies de Lille Flandres Artois, janvier 1991.
- [DUV 93] L. Duvieubourg, T. Carpentier et J.G. Postaire, "Monitoring changes in length with a linear CCD camera in mechanics of deformable solids, Proc. 5th Int. Congress on Condition Monitoring and Diagnostic Engineering Management, Bristol, Juillet 1993.
- [GLA 95] R. Glachet R, S. Bouzar and F. Lenoir , 'Estimation des flux de voyageurs dans les couloirs du métro par traitement d'images' Recherche Transport Sécurité, pp.15-22, 1995.
- [HOR 81] B.K.P Horn and B.G. Schunk, Artificial Intelligence, vol.17, pp. 185-203, 1981.
- [KHO 88] L. Khoudour, J.B. Lesort, ' Analyse du fonctionnement d'un carrefour sous différents modes de régulation' Rapport de recherche Inrets No 73, Juillet 1988.

-
- [KHO 96/1] L. Khoudour, L. Duvieubourg, et J.P. Deparis, ' Real-time passengers counting by active linear cameras' Proceedings SPIE, 29-30 January 1996, San-Jose California USA, pp 106-117.
- [KHO 96/2] L. Khoudour, L. Duvieubourg, et J.P. Deparis, "Linear image sequence analysis for passengers counting in Public Transport", International Conference on Public Transport Electronic Systems, Conference publication No 425, May 1996, IEE, London.
- [KHO 96/3] L. Khoudour, L. Duvieubourg, et J.P. Deparis, "A directional and size varying Structuring Element for isolating and counting moving deformable shapes", CESA'96 , Computational Engineering in Systems Applications; July 9-12 1996, Lille, France.
- [KHO 96/4] L. Khoudour, L. Duvieubourg, B. Meunier and J.P. Deparis, "A new fast algorithm using an adaptive structuring element applied to a counting device", IEEE International Conference on Image Processing, vol....., September 1996, Lausanne, Suisse.
- [KHO 96/5] L.Khoudour, L. Duvieubourg, et J.P. Deparis, "A device for counting passengers making use of two active linear cameras : comparison of algorithms", IEEE Conference....., Octobre 1996, Pekin, Chine.
- [MAC 93] L. Macaire, "Inspection d'états de surfaces métalliques par vision linéaire - Application au contrôle qualité continu en sidérurgie ", Thèse de doctorat, Université des sciences et technologies de Lille Flandres Artois, janvier 1993.
- [MEC 94] A.Mecoci A, F. Bartolini and V. Cappallini, " Image sequence analysis for counting in real time people getting in and out of a bus" Signal Processing 35, 105-116, 1994.
-

-
- [MIN 03] H. Minkowski : "Volume und Oberfläche", Math. Ann., Vol. 57, pp447-495, 1903.
- [OKS 89] J. Oksman, "Le traitement d'images : une technique en développement", La Revue des Laboratoires d'essais, pp. 7-13, février-mars 1989.
- [ROU 92] A. Rourke and M.G.H. Bell, "Video Image Processing Techniques and their Application to Pedestrian Data-Collection ", Research Report No. 83, Transport Operations Research Group, University of Newcastle, December 1992.
- [ROU 94] A. Rourke and M.G.H. Bell, "An image processing system for pedestrian data collection", in 7th International conference on road traffic monitoring and control, april 1994, pp. 123-126, IEE, London.
- [SAP 90] Saporta G. , "Probabilités, analyse des données et statistique", Edition Technip, 1990.
- [SAU 94] C. Saut . "Passengers counting systems - the latest advances" Transport Public International revue, pp 49-50, 1994.
- [SER 82] J. Serra : " Image Analysis and Mathematical Morphology " New York Academic, 1982.
- [SER 86] J. Serra, 'Introduction to Mathematical Morphology', Computer Vision, Graphics and Image Processing 35, 283-305, 1986.
- [TAN 90] P.M. Tancrez, "Un système de vision linéaire pour l'inspection et la caractérisation en temps réel de produits plats en défilement ", Thèse de doctorat, Université des sciences et technologies de Lille Flandres Artois, 1990.

- [VAN 96] P. Vannoorenberghe, CESA'96 , Computational Engineering in Systems Applications; July 9-12 1996, Lille, France.
- [VEL 93] S.A. Velastin and al., "Analysis of crowd movements and densities in built-up environments using image processing", in Int. Coll. on Image Processing for Transport Application, December 1993, pp. 1-6, IEE, London.

Autres articles consultés

- [AOK 84] M. Aoki : "Detection of moving objects using line image sequence" 7th Conference on Pattern Recognition, Montreal, Canada, pp 784-786, July 30-August 2, 1984.
- [BAR 94] F. Bartolini, V. Cappellini, and A. Mecoci, "Counting people getting in and out of a bus by real time image sequence processing", Image and Vision Computing 12(1), vol.12, no.1, pp. 36-41, 1994.
- [CRA 94] F. Cravino, M. Delluca and A. Tesei : " A multiple-hypothesis estimation strategy for crowding evaluation", Internationnal Symposium on Signal Processing, Robotics And Neural Networks, Villeneuve d'Ascq, France, pp 232-235, April 25-27,1994.
- [DAV 92] H.E.H. Davies, "The Puffin pedestrian crossing : experience with the first experimental sites. TRL research Report 364, Transport Reasearch Laboratory, Crowthorne, 1992.
- [FAT 95] M. Fathy and M.Y. Siyal, " A window-based edge detection technique for measuring road traffic parameters in real-time", Real Time Imaging, vol.1, pp. 297-305, 1995.

-
- [HEN 93] T. Hentschel, "Image processing techniques for the estimation of features of crowd behaviour in urban environments", MSc; Dissertation, King's College London, UK, 1993. (algorithme LAT)
- [KHA 89] M.A. Khan and F. Ince, Arabian journal of science and engineering, 14, pp. 541-549, 1989 (dans la biblio de l'article de sergio : comptage aérien).
- [KHO 96/6] L. Khoudour, L. Duvieubourg, et J.P. Deparis, "Real-time pedestrian counting by active linear cameras" Journal of Electronic Imaging/ October 1996/ Vol. 5(4).
- [KIE 94] A.F. Kieper, "Automated fare collection", Public Transport International, vol.6, pp 38-42, 1994.
- [LAB 88] M. Labarrère, J.P. Krief, B. Gimonet, " Le filtrage et ses applications", Editions Cepadues, année 1988.
- [LEU 95] M.K. Leung and Y.H. Yang, "First sight : A human body outline labelling system", IEEE trans. On PAMI, vol.17, no. 4, pp. 359-377, April 1995.
- [LU 94] Y.J. Lu, Y.Y. Tang, P. Pirard, Y.H. Hsu, and H.D. Cheng, "Measurement of pedestrian flow using image analysis techniques", Transportation research rec., no. 1281, pp. 85-96, 1994.
- [OKA 91] Y. Okawa and S. hanatani, "Determination of a pedestrian who does a prespecified body motion in a structured environment, IECON, 2591, pp. 2343-2348, 1991.
- [OSC 82] E. Oscarsson, "TV-Camera detecting pedestrians for traffic light control" Acta Imeko,275-282, 1982.

-
- [PEN 94] E.G. Pennie, "Counting for fair allocation of passenger income in tyne and wear", Public Transport International, vol.6, pp. 28-30, 1994.
- [POS 87] J.G. Postaire, "De l'image à la décision", Edition Dunod, Paris, 1987.
- [REA 95] I.A.D. Reading, K.W. Dickinson and D.J. Barker, "The Puffin pedestrian crossing : pedestrian-behavioural study", Traffic Engineering and control, 36(9), pp. 472-478, September 1995.
- [ROH 93] K. Rohr, "Towards model-based recognition of humen movements in image sequences", Computer Vision Graphics and Image Processing, vol. 59, pp. 94-115, 1993.
- [ROS 70] A. ROSENFELD, "Connectivity in digital pictures", Journal of the association for computing machinery (JACM), pp.146-160, January 1970.
- [ROS 94] M. Rossi and A. Bozzoli : " Tracking and counting moving people", Proceedings of the Second IEEE International Conference on Image Processing, 1994, pp 212-216.
- [TSU 95] M. Tsuchikawa, A. Sato, H. Koike and A. Tomono : " A moving-object extraction method robust against illumination level changes for a pedestrian counting system ", International symposium on Computer vision, Coral Gables, Florida, USA, November 21-23, pp563-568, 1995.
- [VEL 94] S.A. Velastin and al., "Automated measurement of crowd density and motion using image processing", in 7th International conference on road traffic monitoring and control, april 1994, pp. 127-132, IEE, London.

*ANNEXE 1 : Résumé des algorithmes
de comptage*

Annexe 1 : Résumé des algorithmes de comptage

Le schéma de la figure 1, fournit un résumé des procédures et algorithmes utilisés pour réaliser le comptage. Les références A et B concernent les traitements :

- inspirés de la morphologie mathématique et qui sont décrits en figure 2 à la page suivante,
- ASSESA, décrit en figure 3.

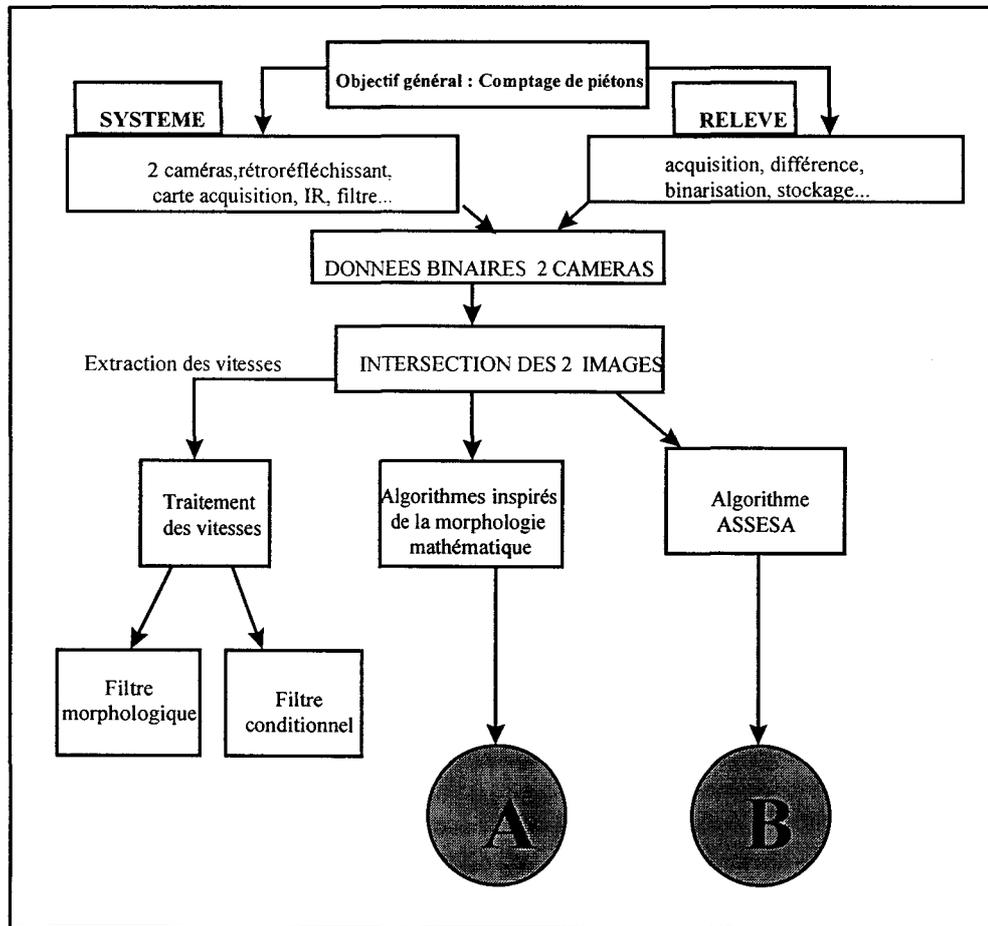


Figure 1 : résumé de tous les traitements utilisés dans le comptage

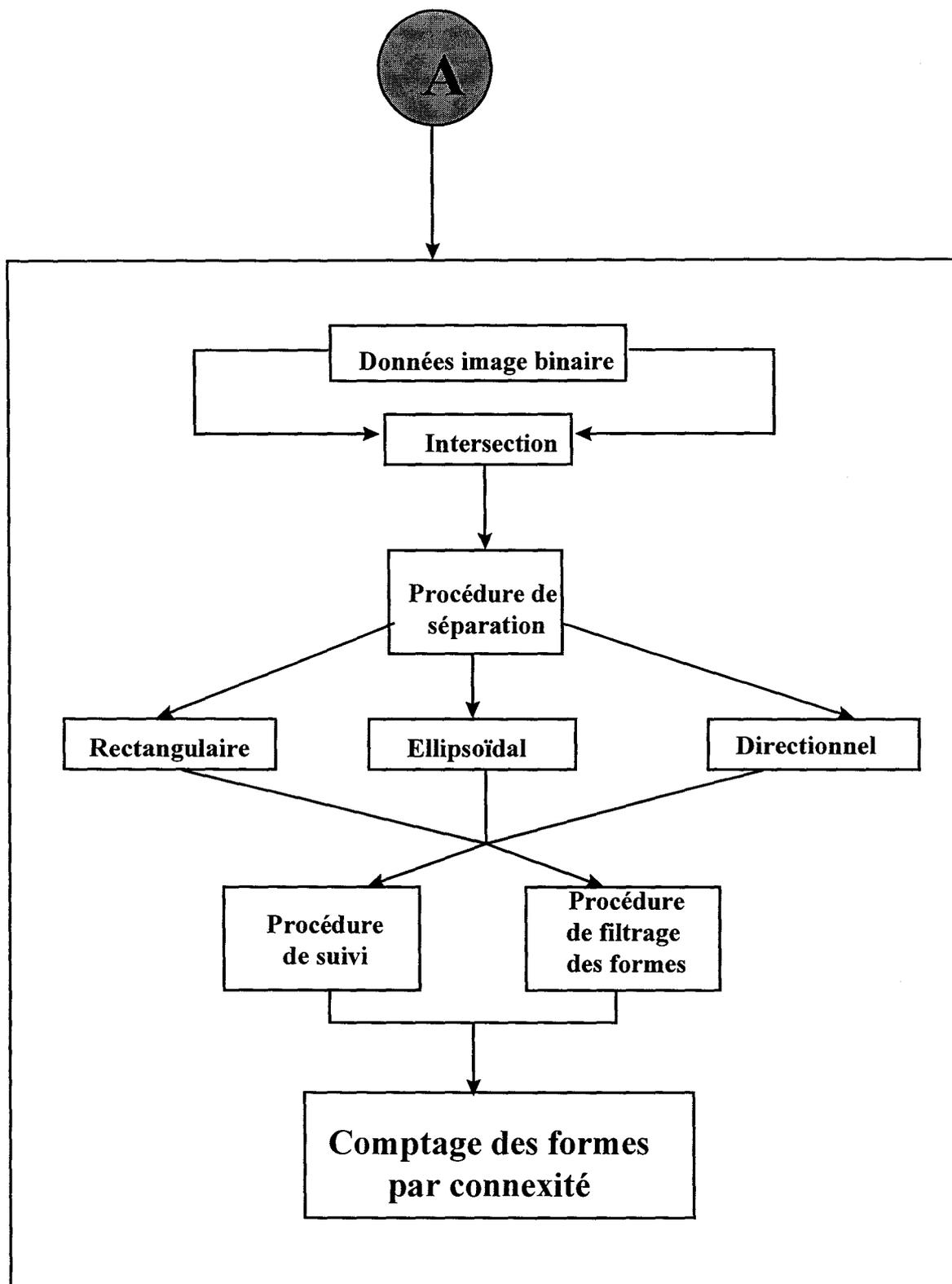


Figure 2 : résumé des procédures inspirées de la morphologie mathématique

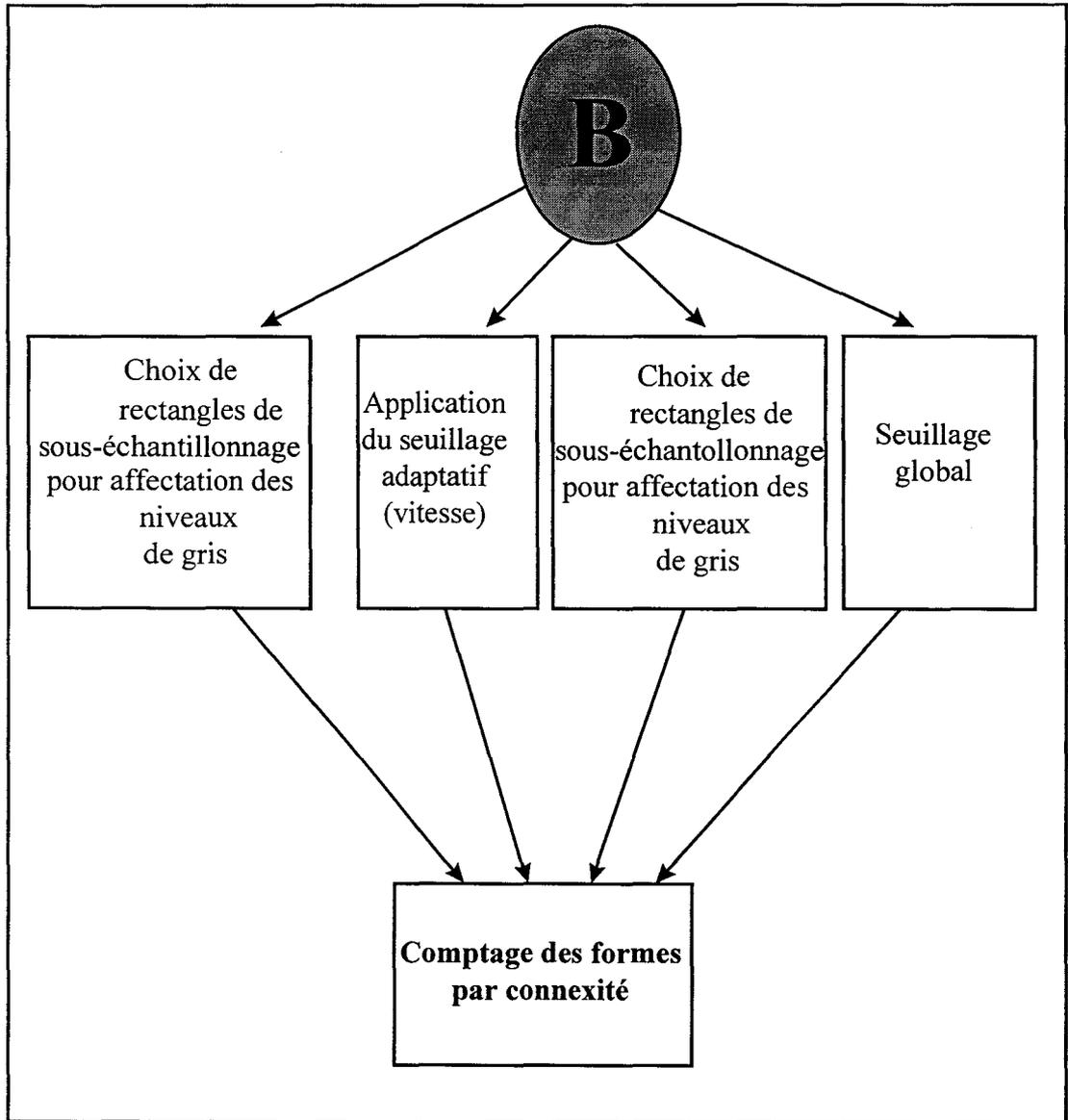


Figure 3 : résumé des procédures de l'algorithme ASSESA

