

CONFIDENTIEL



TITAN : SYSTEME DE MESURE DE TRAFIC

PAR ANALYSE D'IMAGE

INRETS

Blosseville J.M.

Lenoir F.

Espié S.

CMM/EMP

Beucher S.

Bilodeau M.

Octobre 1987

TITAN : SYSTEME DE MESURE DE TRAFIC

PAR ANALYSE D'IMAGE

INRETS

Blosseville J.M.

Lenoir F.

Espié S.

CMM/EMP

Beucher S.

Bilodeau M.

Octobre 1987

TABLE DES MATIERES

AVANT-PROPOS	5
Chapitre I : INTRODUCTION	6
I-1) LES OBJECTIFS	7
I-2) LES ENJEUX	8
I-2-1) En mesure	8
I-2-2) En surveillance	8
I-2-3) En régulation	9
I-2-4) Aspects financiers	9
I-3) LES RECHERCHES MENEES A L'ETRANGER	10
I-3-1) Principales recherches menées à l'étranger	11
I-3-2) Conclusions et implications de ces travaux	17
Chapitre II : LE PROJET INRETS/CMM. METHODE DE REPERAGE DES VEHICULES	18
II-1) LES CONTRAINTES	18
II-2) PRESENTATION DE L'ETUDE	19
II-2-1) La banque d'images-tests	19
II-2-2) Méthode d'analyse d'images utilisée	20
II-2-3) Matériel d'étude	20
II-2-4) Méthode d'approche	21
II-3) DETECTION DE LA CHAUSSEE, CALCUL DES DISTANCES AU SOL	21
II-3-1) Mise en évidence de la chaussée	22
II-3-1-1) Algorithme simplifié	24
II-3-1-2) Algorithme utilisant le marquage au sol	27
II-3-1-3) Algorithmes améliorés	27
II-3-2) Calcul des distances au sol	28

II-3-3) Transformations morphologiques prenant en compte la perspective	35
II-4) DETECTION DES VEHICULES	36
II-4-1) Détection des parties sombres	37
II-4-2) Détection des parties claires	38
II-4-3) Concaténation des marqueurs	40
II-4-4) Discussion de la procédure. Algorithmes additionnels.	42
II-5) TRAJECTOIRES ET MESURE DU TRAFIC	46
II-5-1) Constitution des trajectoires	46
II-5-1-1) Localisation des véhicules détectés	47
II-5-1-2) Construction des trajectoires spatio-temporelles	48
II-5-2) Mesures	52
II-5-2-1) Mesure de débit	52
II-5-2-2) Concentration	52
II-5-2-3) Estimation du taux d'occupation	53
II-5-2-4) Vitesse instantanée des véhicules	54
II-5-2-5) Vitesse moyenne temporelle d'un véhicule	54
II-5-2-6) Vitesse moyenne spatiale	54
II-5-2-7) Taux d'occupation spatial	54
II-5-3) Détection d'incidents	56
 Chapitre III : RESULTATS	 58
III-1) TRAITEMENT DES IMAGES: ILLUSTRATIONS ET COMMENTAIRES	58
III-2) PERFORMANCES DE CE TYPE DE CAPTEUR	69
 Chapitre IV : CONCLUSIONS ET PERSPECTIVES D'AVENIR	 70
IV-1) CRITIQUES DES DIFFERENTES ETAPES DU PROCEDE	70
IV-1-1) L'initialisation	70
IV-1-2) La localisation des véhicules	71
IV-1-3) La représentation spatio-temporelle	72
IV-2) PERSPECTIVES: TRAITEMENT EN TEMPS REEL	72

Annexe A: RAPPELS DE MORPHOLOGIE MATHEMATIQUE	73
Annexe B: CALCUL DES PARAMETRES DE DIMENSIONNEMENT AU SOL	80
Annexe C: IMPLANTATION DES TRANSFORMEES GEODESIQUES MULTI-MASQUES	84
Annexe D: PRESENTATION SUCCINCTE DU LOGICIEL ALADIN	87
BIBLIOGRAPHIE	90

AVANT - PROPOS

Ce rapport fait le point sur deux années de collaboration entre l'Institut de Recherche sur les Transports et leur Sécurité (INRETS) et le Centre de Morphologie Mathématique de l'Ecole des Mines de Paris (CMM/ENMP) dans le domaine de l'analyse du trafic autoroutier par traitement d'images.

Cette recherche a été entreprise après un appel d'offre du Ministère des Transports, appel d'offre resté sans suite, si ce n'est qu'il a permis, d'une part, au CMM d'appliquer les techniques d'analyse d'images issues de la Morphologie Mathématique à la résolution de ce problème original, et d'autre part, à l'INRETS d'appliquer cette technique à la gestion du trafic et de forger ainsi de nouveaux outils d'analyse et de mesure performants.

L'idée d'utiliser l'analyse d'images à l'analyse du trafic routier n'est pas nouvelle. De nombreux pays industrialisés s'intéressent à cette méthode. Les travaux présentés ici s'insèrent dans ce type de programme. C'est pourquoi on trouvera dans la suite de ce rapport un historique succinct des principales recherches menées à l'étranger. Cette bibliographie permettra également de mieux situer les travaux réalisés par le CMM et l'INRETS.

Les premiers résultats obtenus se sont avérés suffisamment intéressants pour envisager dans un avenir assez proche l'élaboration d'un capteur de trafic basé sur l'analyse d'images. D'ores et déjà, un prototype est en cours de fabrication. Evidemment, la route sera longue avant que des dispositifs automatiques envahissent le réseau routier français ou européen. Ces perspectives d'avenir, pour être concrétisées, nécessiteront encore un important travail de recherche, mais surtout la volonté de la part des industriels et des ingénieurs de trafic de faire sortir ce produit du laboratoire.

INTRODUCTION

Dans certaines conditions de prises de vue, l'image d'une scène de trafic routier est un excellent support d'information. L'idée d'utiliser cette information pour analyser et réguler la circulation s'est concrétisée vers 1970 avec le programme de recherche américain "SCAN". Depuis le nombre d'études et d'expérimentations conduites un peu partout dans le monde montre bien à la fois le grand intérêt de ce problème et sa difficulté d'approche.

La richesse de l'information recueillie par une caméra permet en effet de n'exclure a priori aucun des principaux domaines de gestion du trafic que sont les mesures, la surveillance et la régulation. De plus, la possibilité d'embrasser un champ beaucoup plus vaste que les capteurs classiques autorise l'accès direct à des variables actuellement impossibles à mesurer. Enfin, il est facile d'envisager des systèmes portables ne nécessitant pas de travaux de génie civil.

Cette potentialité de l'information contenue dans une image soulève toutefois de nombreuses difficultés. Entre autres problèmes, les déformations dues à la perspective, la grande variabilité des conditions d'éclairage et des formes de véhicules rendent indispensable l'utilisation de méthodes d'une grande robustesse. D'autre part, afin de pouvoir faire des mesures avec une bonne précision, il est nécessaire de traiter plusieurs images par seconde et donc d'utiliser des procédés très rapides.

Les difficultés d'approche de ce problème sont telles que le champ des recherches reste ouvert : aucun capteur de mesures de trafic suffisamment fiable dans toutes les conditions et utilisant ces procédés, n'existe pour l'instant, même à l'état expérimental.

Les résultats présentés ici sont issus des travaux menés en collaboration entre l'Institut de Recherche sur les Transports et leur Sécurité (INRETS) et le Centre de Morphologie Mathématique de l'Ecole des Mines de Paris (CMM/ENMP).

I - 1) LES OBJECTIFS

De façon à limiter la difficulté d'approche de l'étude, dans un premier temps, seules des scènes de trafic sur autoroute ont été considérées. Il s'agit de vues frontales couvrant des zones de 100 à 300 m d'autoroute à plusieurs voies. Les algorithmes développés concernent des flots de trafic s'éloignant ou se rapprochant de la caméra, éventuellement sur la même image.

Dans le cas de trafic fluide les mesures suivantes et leurs dérivées peuvent être obtenues à partir des algorithmes présentés ci-après :

- concentration : nombre de véhicules par unité de longueur (km) de la chaussée (voie par voie).

- débit : nombre de véhicules franchissant une ligne déterminée de la chaussée (voie par voie).

- vitesse individuelle des véhicules.

- Classification : sur des critères de taille et de longueur apparente, une répartition en deux classes peut être obtenue séparant les véhicules légers des poids lourds.

Une classification plus fine pourrait être envisagée en faisant appel à des techniques plus élaborées de reconnaissance des formes.

D'autres mesures directes ou dérivées sont également accessibles :

- Toutes mesures pouvant être faites sur les trajectoires : distances et temps inter-véhiculaires, et évolution de ces grandeurs : ralentissements, ondes de choc, détection d'accidents ou d'incidents.

- franchissement des limites des couloirs de circulation : changements de voie, arrêt sur voie d'urgence.

- taux d'occupation : pourcentage de temps pendant lequel il y a eu au moins un véhicule dans une surface donnée.

Des mesures concernant plus spécifiquement le trafic dense mais pouvant être également calculées en situation fluide sont actuellement à l'étude. On distingue :

- Vitesse moyenne de flot : c'est la vitesse voie par voie d'une file de véhicules indiscernés.

- Densité linéaire : proportion d'un segment de voirie occupé par une

file de véhicules. Cette mesure est une estimation de la concentration.

- Longueur de file d'attente.

Afin d'être élaborées avec précision, toutes ces mesures nécessitent le traitement de plusieurs images par seconde. Ceci implique l'utilisation de processeurs temps réel d'analyse d'image extrêmement rapides. L'INRETS et le CMM étudient à l'heure actuelle la conception d'un processeur spécialisé suffisamment rapide pour supporter les algorithmes présentés ici. Un prototype sera réalisé et des essais en vraie grandeur pourront alors être envisagés.

I - 2) LES ENJEUX

La richesse de l'information recueillie par une caméra permet a priori de n'exclure aucun des principaux domaines de la gestion du trafic que sont les mesures, la surveillance et la régulation.

I-2-1) En mesure

Le traitement d'images de trafic permet de mesurer en temps réel ou différé, sur des installations fixes ou mobiles la plupart des paramètres accessibles avec des moyens classiques : débit, vitesse, taux d'occupation. Il permet de plus l'accès aux variables spatiales de la circulation dont la mesure, impossible avec les dispositifs actuels, nécessite l'emploi d'un nombre important d'enquêteurs : la concentration ou densité, les mouvements directionnels, le stockage et les temps de dégagement aux carrefours, l'étude des points de conflits et des zones dangereuses.

I-2-2) En surveillance

Les systèmes de télésurveillance comprennent en général des ensembles de caméras qu'il serait possible de mieux exploiter sans doubler l'installation par des stations de détection. L'exploitation automatique des images permettrait, d'une part, de sélectionner les images les plus intéressantes à visualiser sur les écrans de contrôle (en utilisant par exemple des algorithmes de détection d'incident) et d'autre part de donner à l'exploitant, en même temps que l'image, les caractéristiques principales

chiffrées du trafic visualisé (débit, concentration, vitesse,...).

I-2-3) En régulation

Le principal problème que les systèmes actuels de régulation urbaine sont incapables de prendre en compte de façon satisfaisante, faute de moyens de mesure, est celui du blocage des carrefours, et en particulier du blocage dû au stockage des mouvements tournants. L'utilisation de capteurs vidéo pourrait de ce point de vue constituer une avance considérable dans la gestion des réseaux urbains, en permettant une commande en temps réel des carrefours évitant les blocages, car ils pourraient donner en permanence une mesure exacte de l'encombrement du carrefour lui-même, ainsi que de ses sorties. La modulation des temps de vert des feux de carrefours pourrait ainsi être ajustée avec précision.

De la même façon, la régulation en temps réel des accès de voies rapides serait facilitée et significativement améliorée par l'utilisation directe de mesures de concentration en remplacement d'estimations lourdes et peu précises effectuées actuellement avec des capteurs ponctuels.

I-2-4) Aspects financiers

Dans l'état actuel de la technique, l'installation d'un carrefour à feux (armoire et boucles) asservis au trafic de façon satisfaisante coûte de l'ordre de 100 à 300 KF, dont 60% représentent les travaux de génie civil nécessaires au rapatriement des informations de boucles jusqu'à l'armoire de carrefour. Une caméra, embrassant un champ plus vaste, pourrait être implantée beaucoup plus près de l'armoire de carrefour, et permettre ainsi une économie importante sur les travaux de génie civil. De façon analogue une caméra assortie d'un système d'analyse, installée sur une autoroute en rase campagne constitue bien, même financièrement, un moyen de mesure compétitif de la station de mesure traditionnelle (dont le coût avoisine les 300 KF).

I - 3) LES RECHERCHES MENEES A L'ETRANGER

Comparativement à l'explosion de la recherche en analyse d'image dans des domaines comme la biologie, la robotique, la géologie, etc... les tentatives de mise au point d'un système de mesures de trafic par traitement numérique d'image restent peu nombreuses.

On peut classer, brièvement, les méthodes susceptibles de détecter un objet en mouvement en 3 groupes. Elles utilisent, soit :

- a) la différence entre une image courante et une image de référence.
- b) la différence entre deux images successives.
- c) la segmentation et/ou classification d'images.

Le principe fondamental de la première catégorie de méthodes consiste à utiliser, sous diverses formes, le phénomène illustré par la pause photographique : la sommation d'images successives a tendance à conserver précisément les objets fixes et à faire disparaître les objets en mouvement. La soustraction entre une image de référence construite de cette manière et une image courante permet donc en théorie d'isoler les véhicules en mouvement. Dans la pratique, les difficultés dues à la variabilité de l'image altèrent l'image de référence et réduisent la portée de ce principe.

La seconde méthode participe de la même idée mais cette fois la soustraction est faite entre deux images successives. Sur l'image résultat les objets immobiles se sont estompés et une zone de mouvement combinant les positions des véhicules aux deux instants considérés par la capture des images peut être mise en évidence. C'est notamment cette méthode qu'utilise une équipe du CRESTA/INRETS (POSTAIRE, STELMASZYK, BONNET, 1986, [14]) pour localiser un mobile circulant dans le champ d'un véhicule sur rail automatique. La finalité de cette étude (sécurité) et la diversité des mobiles à détecter la placent dans un cadre plus général que celui qui nous intéresse ici.

Le troisième groupe de méthodes fait appel à des traitements d'images sophistiqués, à la fois longs et complexes. C'est la raison pour laquelle, dans le domaine de l'imagerie appliquée au trafic, leur emploi est resté, jusqu'à maintenant, quasiment nul.

Ceci explique pourquoi toutes les recherches antérieures ou actuelles

utilisent uniquement les deux premières classes de méthodes.

I-3-1) Principales recherches menées à l'étranger

L'une des premières tentatives date de 1973 et provient de l'université de Tokyo (ONOE & Al, 1973,[12]). Ce premier dispositif reposait sur le principe suivant : une image correspondant à la scène scrutée sans véhicule était enregistrée sur un tube à mémoire et soustraite de l'image courante. Seules les parties de l'image correspondant au mouvement étaient ainsi obtenues et stockées sur disque pour un traitement en différé. Il était possible de stocker une image par seconde et de faire de grossières estimations de la vitesse des véhicules. Des problèmes pratiques d'actualisation de l'image fond et de temps de calcul ont provoqué l'abandon de ce système.

Le même groupe (ONOE, 1976,[13]) développe ensuite un second système pour lequel l'analyse est réduite à des pixels de l'image sélectionnés sur une ligne de la chaussée étudiée. Les véhicules sont détectés le long de cette ligne lorsque l'intensité en niveaux de gris des pixels concernés s'écarte significativement d'une valeur pondérée des mêmes pixels aux instants précédents. Même en considérant cette réduction du nombre de pixels à analyser, les temps de calcul sur un mini-calculateur atteignaient un temps prohibitif de 10 secondes par image.

Deux autres groupes de recherche ont par ailleurs, en Allemagne et en Suède, travaillé à la mise au point de dispositifs similaires à celui développé en 1976 par ONOE & Al. Dans les deux cas (HOUKES, 1980,[10] et ABRAMCZUK, 1982,[1]), la détection des véhicules se fait le long de lignes fuyantes disposées longitudinalement sur chacune des voies (Fig. I.1). Les valeurs de niveaux de gris correspondant à l'image de fond sont soustraites de l'image courante, et les groupes de pixels fournissant une différence supérieure à un seuil sont estimés représenter des véhicules. Des problèmes surviennent en cas de densité élevée du trafic ou de changement brusque de l'éclairement.

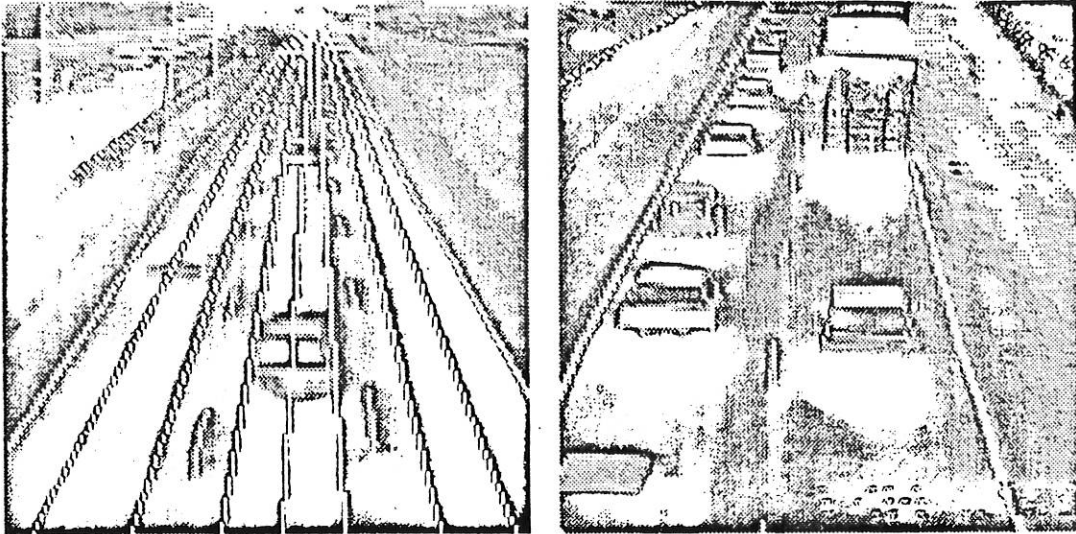


Figure I.1 : Détection des véhicules par le dispositif
de HOUKES et ABRAMCZUK

De façon à accroître la vitesse de calcul et rendre le système utilisable en temps réel, un troisième dispositif (TAKABA, 1982,[17]) fut mis au point à Tokyo. Ce système ne prend en compte que 64 points de l'image, cette fois analysés 30 fois par seconde. Le plus récent des systèmes étudié par cette équipe (TAKABA, 1984,[18]) ne considère que 16 points de l'image scrutés 50 fois par seconde. Pour ces deux derniers dispositifs, les points de détection sont situés le long de deux lignes transversales par rapport à la voie étudiée. Les véhicules sont détectés à partir des différences entre la luminance de l'image courante mesurée sur les points de détection et celle considérée comme étant représentative de la chaussée sans véhicule. Cette image de référence est actualisée par lissage exponentiel.

Les résultats obtenus avec ce dispositif sont assez bons mais sujets à caution si l'on considère qu'ils ont été acquis après réglage optimum des paramètres de l'algorithme. Ce manque de rigueur dans la présentation des résultats est d'ailleurs une constante dans l'ensemble de ces recherches; chacun produit des résultats sur un site particulier après un calibrage du dispositif totalement empirique dont on ne précise d'ailleurs pas l'importance. Pour ce qui est du dernier dispositif décrit, les paramètres mentionnés sont les valeurs de seuil utilisées pour obtenir l'image binaire, les paramètres de réduction du bruit spatial et les valeurs fixant la remise

à jour de l'image fond. Compte tenu de ces détails, l'auteur obtient dans les conditions de l'expérience, dans un cas, 6 véhicules non détectés sur 111, et dans un autre, 1 véhicule non détecté sur 181. Une "bonne coïncidence" est obtenue par ailleurs sur les mesures de vitesses.

Plus récemment une équipe de chercheurs australiens (DODS, 1984,[9]) publiait des résultats concernant un dispositif assez proche de celui développé par TAKABA.

Ce système détecte la présence de véhicules sur des points particuliers de l'image. 16 points de détection sont au préalable choisis par l'opérateur ainsi qu'un point de référence sur une partie de l'image non affectée par les mouvements des véhicules.

Le traitement est complètement analogique et effectué au rythme de 50 images par seconde. Le niveau de luminance enregistré sur le point de référence comparé à la valeur précédemment calculée fournit un décalage à répercuter sur les niveaux de luminance de chacun des points de détection. Ces niveaux corrigés sont comparés à un seuil pour décider de la présence d'un véhicule. Les différences des niveaux successifs sont également utilisées pour mettre en évidence des véhicules arrêtés sur un point de détection.

Le pourcentage de véhicules détectés par ce système dans une scène très ensoleillée et en utilisant les seuils de détection optimaux est égal à 97%.

Une autre façon de réduire l'information à traiter mais d'en assurer la validité consiste à réduire le champ de vision de façon à n'avoir à détecter qu'un ou deux véhicules à la fois. C'est l'approche utilisée par une équipe belge de Louvain (MAES, 1984,[11]). Une caméra CCD placée presque verticalement fournit des images d'une seule voie de circulation embrassant une zone de 8 à 15 m de profondeur. Quatre lignes d'image sont analysées 50 fois par seconde par différence avec un niveau de référence obtenu en l'absence de véhicule. Après seuillage et filtrage, la position de l'avant du véhicule est repérée sur chaque image (Fig. I.2). Les positions successives des véhicules permettent le calcul de la vitesse et des intervalles inter-véhiculaires. Tous ces calculs nécessitent un processeur principal et un processeur par ligne d'image inspectée. Certains problèmes subsistent: les ombres des véhicules sur une voie adjacente provoquent des détections erronées, les réflexions sur la chaussée en cas de pluie et les

scènes de nuit ne sont pas maîtrisées. Néanmoins un produit opérationnel peut être obtenu rapidement au détriment évidemment d'une certaine flexibilité.

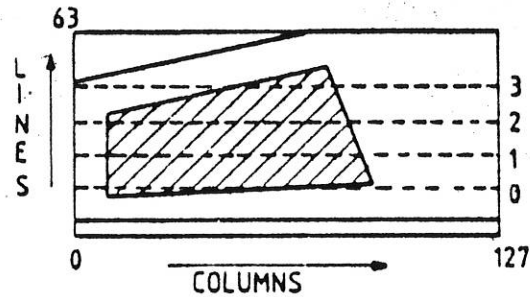
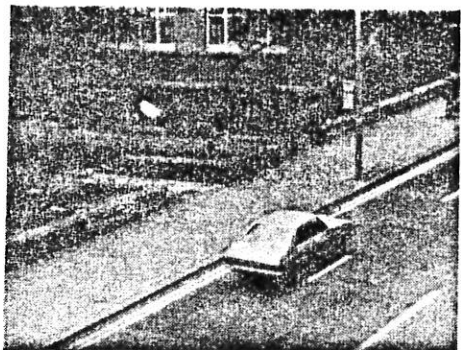


Figure I.2 : Dispositif de MAES

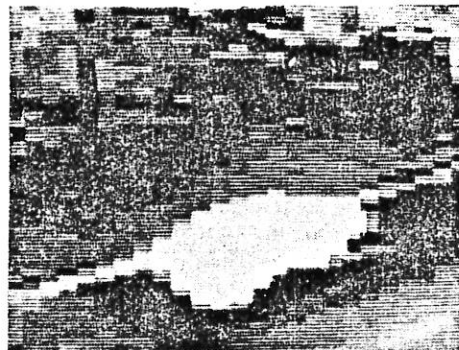
L'étude réalisée par l'équipe britannique (DICKINSON & Al, 1983, [8]) se place comme celle réalisée par l'équipe de Louvain dans un contexte de trafic relativement réduit (1 ou 2 véhicules présents sur l'image). Une caméra CCD fournit des images (8 par seconde) numérisées selon une trame 100x100 à une structure de 6 microprocesseurs disposés en pipe-line. Chacun d'eux effectue respectivement les opérations suivantes:

- 1) conversion analogique-digitale et convolution.
- 2) actualisation de l'image fond et soustraction avec l'image courante.
- 3) seuillage.
- 4) compression de l'image.
- 5) organisation de l'image en blocs à stocker sur disque.
- 6) contrôle du disque.

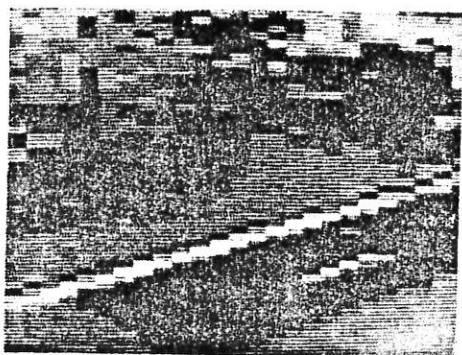
Ces données sont ensuite traitées en différé (Fig. I.3). 89% des véhicules passant dans la scène sont détectés. Dans une deuxième expérience citée par les auteurs, on ne traite que deux lignes d'image à une vitesse de 200 images par seconde. Vitesses et longueur des véhicules sont calculées.



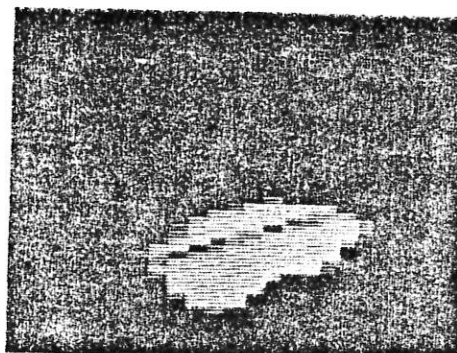
a) image initiale



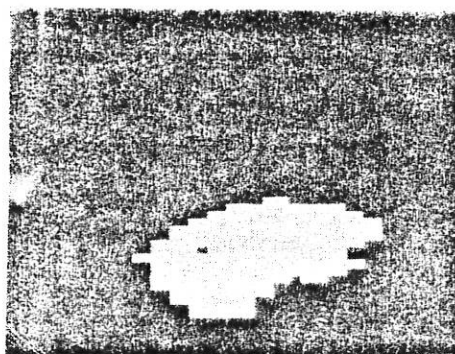
b) image numérisée



c) image fond numérisée



d) image différence



e) image binaire

Figure I.3 : Procédé utilisé par DICKINSON & Al

Depuis 1970, le Department of Transportation des Etats-Unis finance des travaux du Jet Propulsion Laboratory sur ce sujet. Le projet a consisté dans un premier temps à réaliser un dispositif véritablement capable de suivre un

véhicule après l'avoir repéré au passage d'une ligne d'image. Sur cette ligne de détection une signature du véhicule est obtenue et à l'aide d'une technique de corrélation croisée, les positions suivantes du véhicule sont estimées (Fig. I.4). Un rapport (SCHLUTSMEYER, 1982,[15]) fait état des problèmes liés à la poursuite de véhicule suivant cette technique. Des changements de luminosité au cours du processus, la déformation de l'apparence des véhicules sont les causes d'erreur de l'ordre de 25% sur la localisation et sur la vitesse. Finalement, le projet final retenu consiste à ne considérer que quelques lignes de détection sur l'image et à calculer la vitesse des véhicules au voisinage de cette ligne. Les résultats sont comparables à ceux d'une boucle magnétique et le paramétrage reste assez simple. Des mesures ont été faites dans des conditions de trafic très dense et également pendant la nuit.

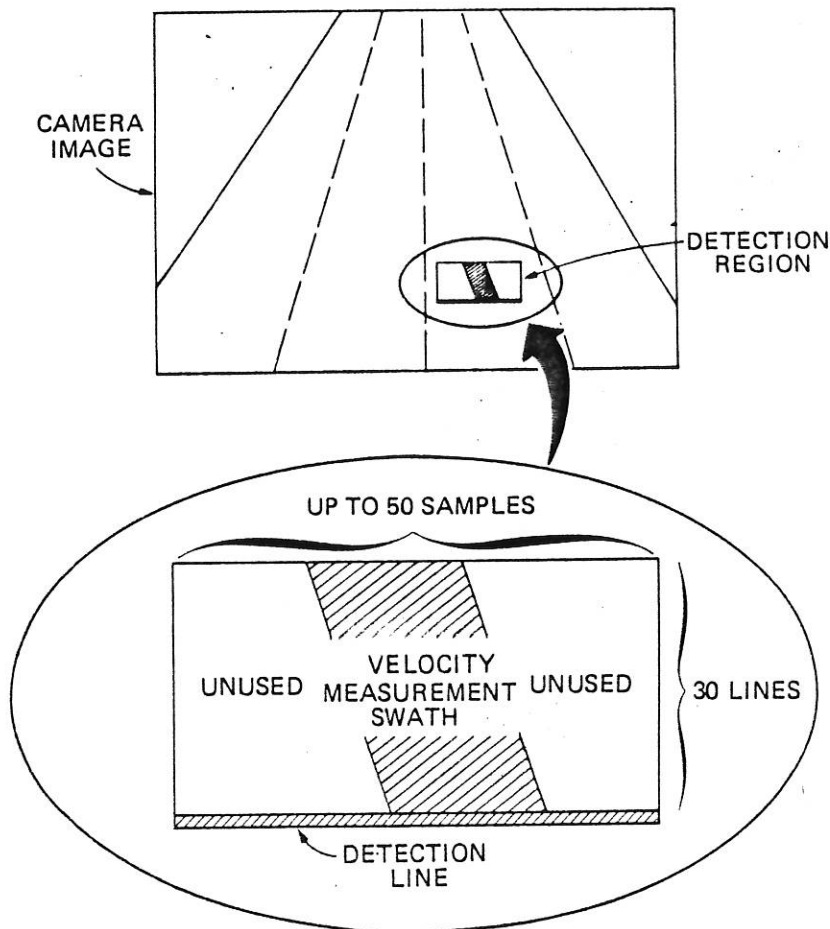


Figure I.4 : Principe du dispositif du JPL (USA)

I-3-2) Conclusions et implications de ces travaux

Les dispositifs issus de toutes ces recherches sont loin de fournir l'ensemble des mesures souhaitées. La difficulté majeure à laquelle se sont heurtés tous les systèmes présents est celle de la quantité d'information à traiter dans un temps très court. Chacun propose une manière particulière de limiter l'information faute de pouvoir disposer de moyens de calcul suffisamment rapides. Par rapport à toutes ces approches la caractéristique première d'un système plus performant réside donc dans une plus grande rapidité de calcul.

Toutes ces méthodes utilisant la différence par rapport à une image de référence ont montré une certaine efficacité à séparer les véhicules de leur environnement statique. Les problèmes à résoudre dans ce cas concernent essentiellement l'actualisation de l'image de référence et la séparation de véhicules apparemment connexes et même la délimitation précise des contours des véhicules. Toutes ces difficultés amènent des erreurs de détection que l'on peut éviter pour certaines conditions de prises de vues mais qui limiteraient considérablement le champ d'application d'un capteur élaboré suivant ces principes.

Avoir présent à l'esprit l'ensemble des difficultés et des contraintes rencontrées lors de ces travaux antérieurs a donc été d'un grand secours lors de l'élaboration et de la réalisation du projet de capteur d'images de trafic entrepris par l'INRETS et le CMM, projet dont la description et l'état d'avancement font l'objet de la suite de ce mémoire.

LE PROJET INRETS/CMM : METHODE DE REPERAGE DES VEHICULES

II - 1) LES CONTRAINTES

Etablir les contraintes devant être supportées par un analyseur automatique d'images de trafic est une tâche primordiale. De la liste de ces contraintes (celles imposées par la nature même du problème, ou celles que l'on s'impose a priori) dépendra l'ensemble de la chaîne de traitement, la méthodologie utilisée et finalement l'architecture d'un processeur apte à fonctionner en temps réel.

Les premières contraintes résultent de la grande diversité des situations rencontrées. Cette diversité a plusieurs origines :

- La nature des sites analysés. L'étude du trafic autoroutier amène des problèmes et des contraintes assez différentes de celles rencontrées en trafic urbain, la deuxième situation étant en général plus complexe que la première (analyse de carrefours, trafic plus dense, moins bien délimité dans l'espace, etc...).

- Le moment de la journée où se déroule l'observation (scènes de jour ou de nuit). De plus, la situation se complique grandement lors des périodes de transition (aube et crépuscule).

- Les conditions de trafic. Celui-ci peut être fluide ou dense, voire même saturé.

Si ces contraintes sont inévitables, il en est d'autres par contre que l'on a volontairement voulu introduire dès la première approche du problème. Ces contraintes que l'on s'impose constituent la deuxième caractéristique de cette étude. Parmi ces contraintes, la plus importante est de pouvoir utiliser délibérément pour l'acquisition des images l'infrastructure déjà existante (caméras vidéo déjà installées, réseaux d'acheminement des images vers les postes de commande, etc...). Ce choix est destiné essentiellement à réduire les coûts d'installation et d'exploitation d'un tel système, coûts qui deviendraient probablement prohibitifs si l'on utilisait des capteurs sophistiqués (caméras CCD, haute définition,

infra-rouge, etc...). Cependant, l'infrastructure existante est parfois ancienne, les caméras exposées aux intempéries se dérèglent facilement et les images fournies par de tels capteurs sont loin d'être de bonne qualité. De plus, et c'est là un point important, certaines des caméras implantées sont télécommandées. L'usage de ces caméras est intéressant notamment dans le cadre d'une utilisation du procédé en DAI (Détection automatique d'incidents). La télécommande permet en effet à l'opérateur une meilleure observation de la scène une fois qu'il a été alerté par le système. Malheureusement, il est absolument impossible, après observation, de replacer ces caméras dans une position exactement identique à celle qu'elles occupaient auparavant. La scène à analyser n'est donc pas la même. Le système doit donc être capable de se recalibrer de façon automatique après intervention de l'opérateur.

Les contraintes listées ci-dessus amènent différents niveaux de difficultés lors de la résolution du problème posé. Il convient donc de faire un choix parmi celles dont on veut tenir compte pour une étude de faisabilité. Ce choix doit prendre en compte un niveau de difficulté suffisant pour être représentatif des conditions réelles de l'étude, mais il doit cependant demeurer suffisamment simple pour rester dans le cadre d'une étude de faisabilité.

II-2) PRESENTATION DE L'ETUDE

Compte tenu des contraintes énumérées précédemment, plusieurs choix ont été faits concernant la banque d'images-tests, les méthodes d'analyse d'image, le matériel d'étude et la méthode d'approche.

II-2-1) La banque d'images-tests

Les scènes analysées sont des scènes de trafic sur autoroute couvrant une zone de 80 à 300 m, la caméra étant placée dans l'axe de la chaussée, les véhicules observés s'en éloignant ou s'en rapprochant.

Bien que l'on se soit limité à des scènes de jour, différents niveaux de luminosité dus aux conditions atmosphériques (soleil, pluie, neige) ont été considérés.

On s'est restreint à l'analyse de flux de trafic fluide à dense de

telle sorte que la connexité apparente des véhicules reste faible. Il est à noter que les conditions d'application des algorithmes développés seront d'autant plus larges que la caméra sera haut placée.

Les scènes de nuit ou de fortes densité de trafic font l'objet d'une autre étude en cours.

Différentes séries d'images ont servi à tester les algorithmes. Elles ont été prises soit à partir de caméras installées spécialement pour cette étude dans des conditions similaires à celles qui régissent l'implantation sur des sites réels de caméras destinées à la surveillance du trafic, soit à partir de ces caméras déjà installées dont les images sont collectées dans les postes centraux des autoroutes de la région parisienne: Autoroutes A6, A13, Boulevard périphérique.

II-2-2) Méthode d'analyse d'image utilisée

Les algorithmes d'analyse des images font largement appel aux outils de la Morphologie Mathématique. Cette méthodologie de traitement des images fournit des procédés rapides et souples d'analyse présentant de plus l'avantage d'être facilement implantables dans des processeurs câblés ([7],[16]). Les principales transformations utilisées dans cette étude sont succinctement décrites à l'annexe A.

II-2-3) Matériel d'étude

Les images traitées ont été enregistrées sur un magnétoscope au standard VHS et numérisées avec une résolution de 256x256 pixels sur 6 bits (64 niveaux de gris). Pour cette étude, un logiciel permettant l'écriture, l'exécution et l'enchaînement d'algorithmes d'analyse d'image a été conçu et développé. Ce logiciel constitue un produit de la recherche à lui seul et fait l'objet d'une commercialisation (cf annexe D). L'ensemble du traitement a été effectué soit en utilisant ce logiciel sur micro-ordinateur compatible IBM PC/XT, PC/AT, soit sur des machines spécialisées NS1500 (NACHET VISION) et MORPHOPERICOLOR (NUMELEC)¹.

¹ Produits sous licence ARMINES/Ecole des Mines.

II-2-4) Méthode d'approche

Le traitement des images s'effectue en trois temps : La première phase consiste à sélectionner dans la scène analysée, la zone d'intérêt (la chaussée et les différentes voies qui la composent). Cette délimitation de la chaussée a, on le verra, un double intérêt. La deuxième phase consiste à extraire les véhicules présents dans la scène, en utilisant différents critères de dimension et de luminance. A la fin de cette seconde phase, les véhicules sont supposés être repérés (il peut néanmoins y avoir des erreurs). La dernière partie du traitement propose de mettre en évidence les trajectoires des véhicules, vérifier leur cohérence, calculer, voie par voie, les divers paramètres accessibles (vitesses individuelles, vitesses de groupe, débits, concentrations, taux d'occupations, etc...).

Divers traitements additionnels peuvent également être effectués. Ils seront décrits plus en détail par la suite.

II - 3) DETECTION DE LA CHAUSSEE, CALCUL DES DISTANCES AU SOL

La détection de la chaussée a un double intérêt :

- La limitation en étendue de la zone utile de la scène analysée augmente les vitesses de traitement et réduit les problèmes liés à la détection d'artefacts. Le repérage des différentes voies s'avère également indispensable lors de l'analyse des trajectoires.

- L'autre intérêt de cette détection provient d'une contrainte inhérente à l'utilisation de capteurs d'images : la perspective. En effet, la caméra fournissant une vue de la chaussée sur plusieurs centaines de mètres, la perspective fait que la taille apparente des véhicules varie grandement de l'avant-scène à l'arrière-plan. Comme on désire néanmoins détecter les véhicules sur l'ensemble de la scène analysée, ceci afin d'avoir accès à des informations que ne fournissent pas les capteurs classiques (boucles magnétiques), la connaissance de la déformation apportée par la perspective est de première importance pour adapter les valeurs des paramètres des algorithmes à la position de chaque pixel dans l'image. Ce paramétrage consiste à définir des zones de l'image où la taille des transformations utilisées est constante. La génération de ces zones ou masques impose de bien connaître les distances au sol correspondant à chaque point de l'image. Cette connaissance de la perspective permet également de

positionner les véhicules repérés dans l'espace réel.

II-3-1) Mise en évidence de la chaussée

Pour mettre en évidence la chaussée, deux images sont générées :

- Une première image appelée image somme (ou IMS), obtenue en moyennant une séquence d'images successives de la scène :

$$IMS = \sum_{i=1}^n im_i / n$$

Cette opération fournit une image du site débarrassée des objets en mouvements, pourvu que le nombre n soit suffisamment grand (Fig. II.1).



Figure II.1 : Image somme.

- Une deuxième image appelée image différence (ou IMD), obtenue soit par moyenne des différences en valeur absolue de l'image IMS aux images successives, soit simplement par moyenne de différences d'images successives :

$$\text{IMD} = \sum_{i=1}^n |im_i - \text{IMS}| / n$$

ou :

$$\text{IMD} = \sum_{i=1}^{n-1} |im_{i+1} - im_i| / n$$

La deuxième formule est plus simple à mettre en oeuvre. Cependant les images obtenues sont équivalentes et correspondent à la zone de circulation des véhicules (Fig. II.2), l'application de la deuxième formule fournissant des images moins bruitées.

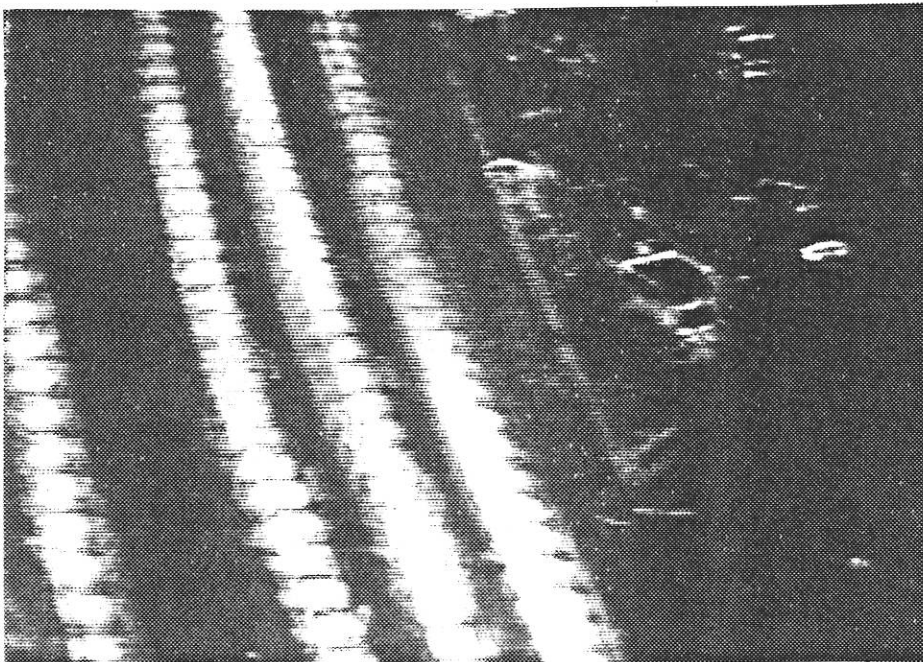


Figure II.2 : Image différence. Zones de circulation.

En fonction de critères liés à la qualité des images et à l'existence d'un marquage au sol séparant bien les voies de la chaussée considérée, plusieurs algorithmes ont été développés.

II-3-1-1) Algorithme simplifié

Dans le cas où les marquages au sol ne sont pas utilisables, seule l'image différence IMD est utilisée. Cette image est d'abord filtrée par ouverture morphologique (Fig. II.3).

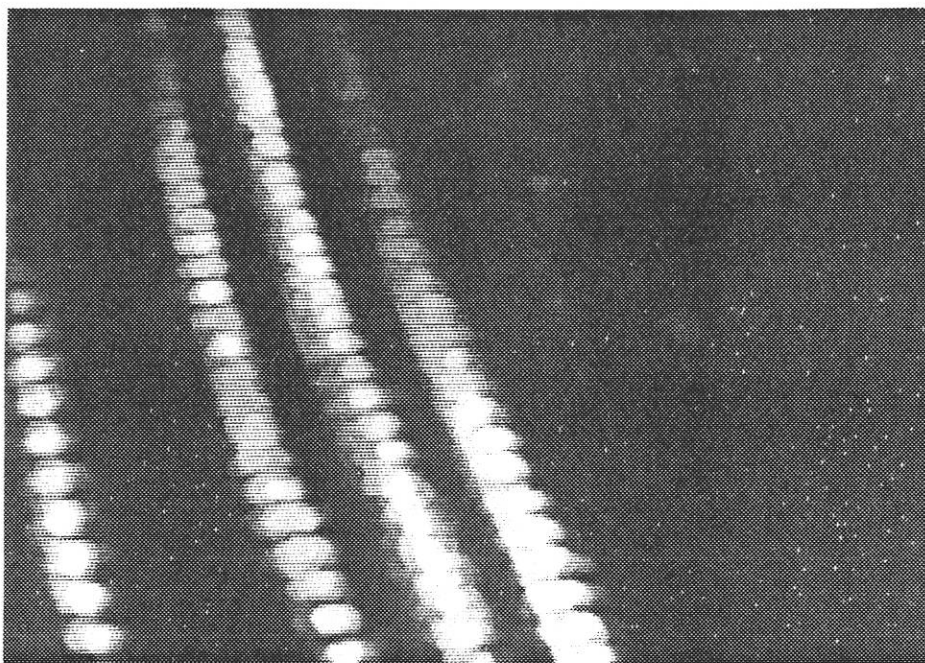


Figure II.3 : Filtrage par ouverture de l'image IMD.

L'image est seuillée, et l'image binaire obtenue est à nouveau filtrée, par érosion, reconstruction des grosses composantes connexes puis bouchage des trous. L'image obtenue est alors dans le meilleur des cas une image binaire où chaque voie est représentée par une composante simplement connexe (Fig. II.4).

La troisième étape du traitement consiste à générer l'extérieur de la chaussée. Pour ce faire, l'image précédente est dilatée, mais de telle façon que l'extérieur de la chaussée soit situé à une distance du bord de la route sensiblement égale à la distance séparant les zones de roulement (Fig. II.5).

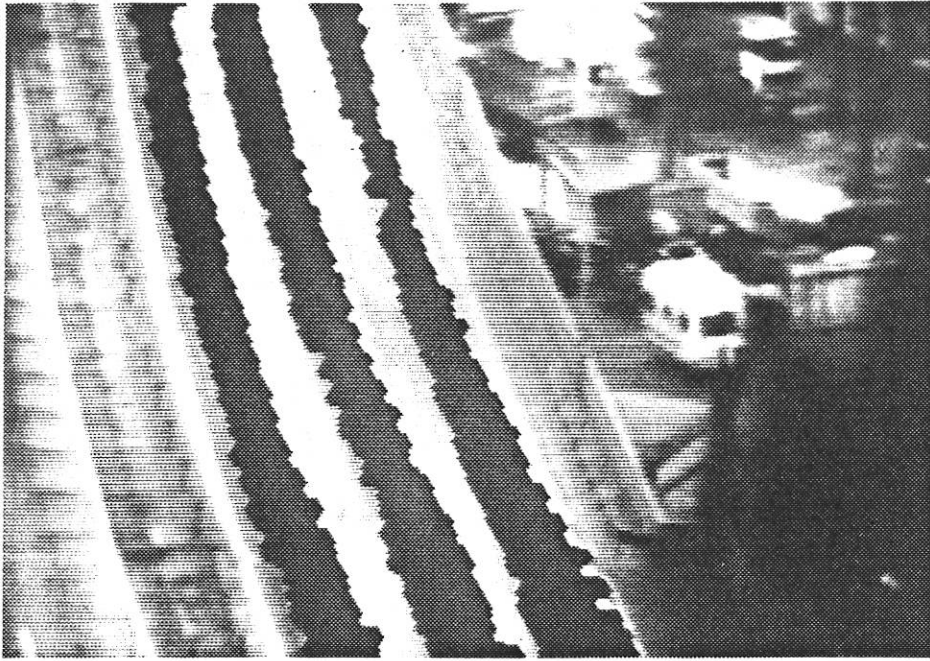


Figure II.4 : Image binaire des différentes voies de la chaussée.



Figure II.5 : Détermination de la région d'intérêt.

On y parvient en effectuant, pour chaque ligne de l'image, une dilatation linéaire jusqu'à ce que les différentes zones de roulement soient connectées. Cette connexion se produisant dans chaque ligne d'image pour une certaine taille de dilatation de taille n , on effectue alors une dilatation linéaire supplémentaire de même taille. Recommencant sur chaque ligne, on obtient l'ensemble complémentaire de l'extérieur de la chaussée.

On peut effectuer ce genre d'algorithme sur des groupes de plusieurs lignes pour gagner du temps. L'hypothèse sur laquelle il repose est que les véhicules occupent en moyenne le milieu de leur voie de circulation. l'union des zones de roulement et de l'extérieur de la chaussée fournit alors l'image suivante (Fig. II.6):

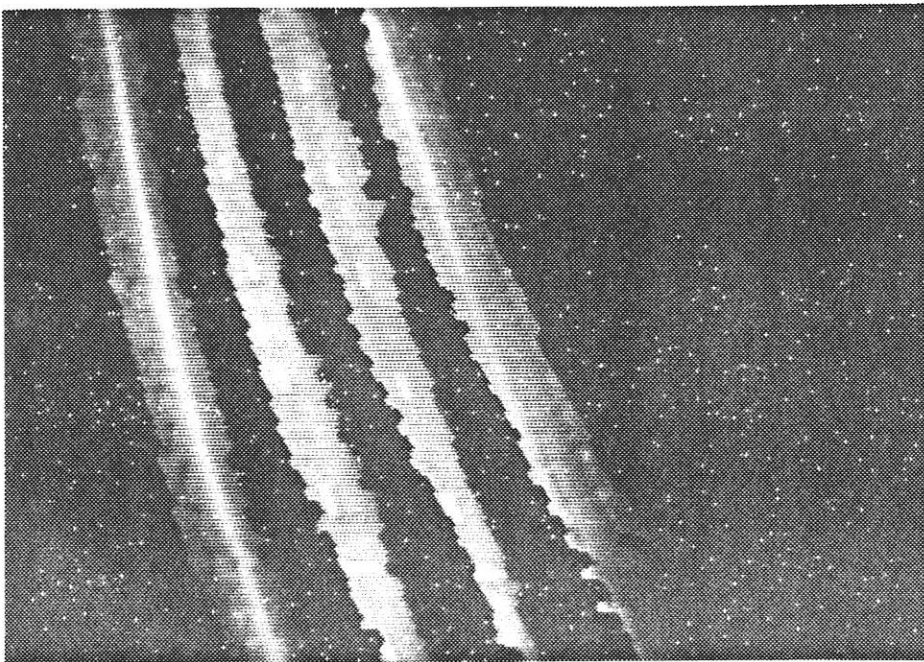


Figure II.6 : Marquage des voies et de l'extérieur de la chaussée.

Un squelette par zone d'influence de cette image permet ensuite de mettre en évidence les différentes voies de la chaussée (Fig. II.7).

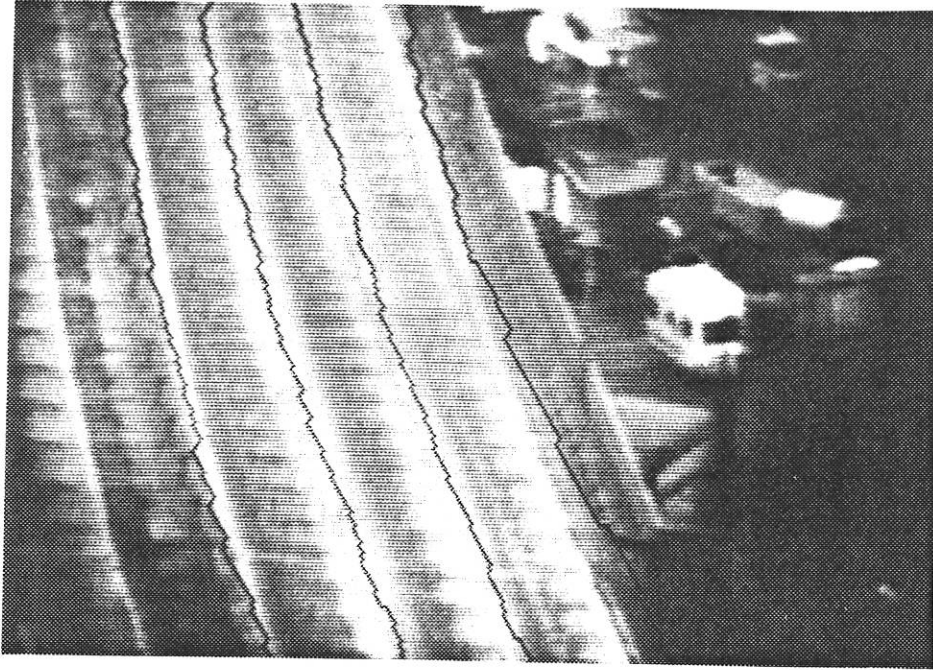


Figure II.7 : Mise en évidence des différentes voies par SKIZ.

II-3-1-2) Algorithme utilisant le marquage au sol

Dans le cas où le marquage au sol présente une bonne définition sur l'image IMS, il est extrait par une transformation appelée chapeau haut-de-forme (planche II.8).

L'image du marquage au sol après seuillage et filtrage est alors utilisée dans la dernière étape du traitement précédent où le squelette par zone d'influence devient un SKIZ géodésique. Cette procédure permet de placer les frontières des voies sur les marquages au sol.

II-3-1-3) Algorithmes améliorés

Plusieurs améliorations ont été apportées à la procédure afin d'augmenter sa robustesse. On peut tout d'abord travailler directement sur les images de gris sans passer par un seuillage initial. L'algorithme peut également être auto-adaptatif. En effet, il est évident que plus le nombre d'images utilisées pour générer l'image IMD sera grand, meilleur sera le résultat final. Cependant, afin de ne pas perdre trop de temps à élaborer le masque de la chaussée, on peut effectuer l'algorithme après une séquence de n images (n est variable, dépend des conditions de trafic du site étudié, n

est généralement compris entre 80 et 100), puis recommencer après $2n$, puis $3n$ images, etc... et à chaque étape tester la convergence du résultat obtenu. Cette façon de faire permet également d'éliminer des fausses séparations intra-voies, car celles-ci ne se produisent pas toujours dans la même zone de l'image (sauf si elles sont dues à des artefacts fixes). La planche II.9 explicite de façon détaillée la procédure adoptée.

La convergence de l'algorithme est testée en calculant à chaque étape le nombre de composantes connexes de l'image et en le comparant avec la valeur souhaitée.

D'autres procédures plus élaborées peuvent être utilisées, notamment lorsqu'il s'agit de générer le masque de la chaussée et les différentes voies présentant un accès à la voie principale (Fig. II.10).

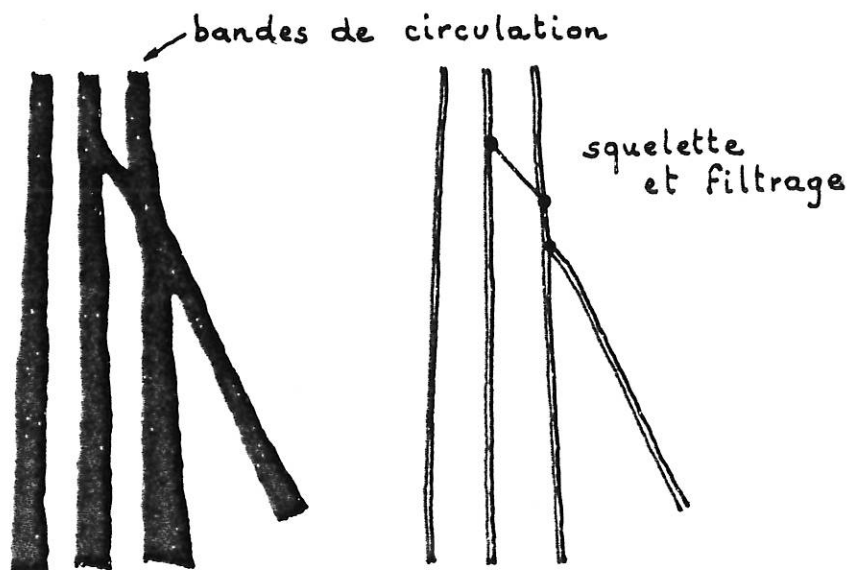
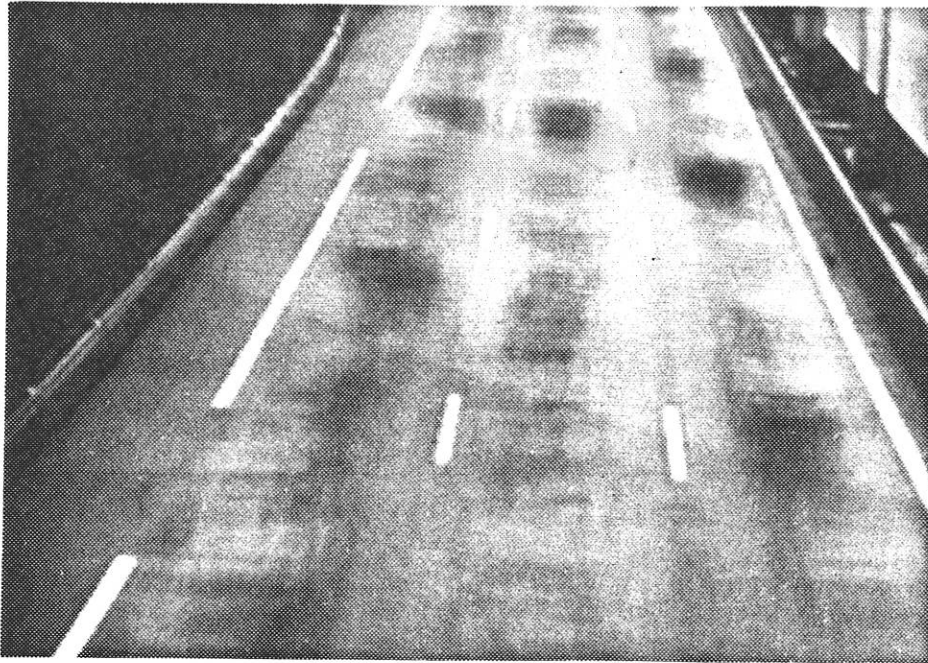


Figure II.10 : Principe de traitement des voies d'accès par segmentation des zones de circulation.

II-3-2) Calcul des distances au sol

Le second intérêt de la mise en évidence des masques de la chaussée est de pouvoir associer en tout point de l'image les distances au sol correspondantes. En effet, la perspective déforme la scène analysée. Les

(a)



(b)

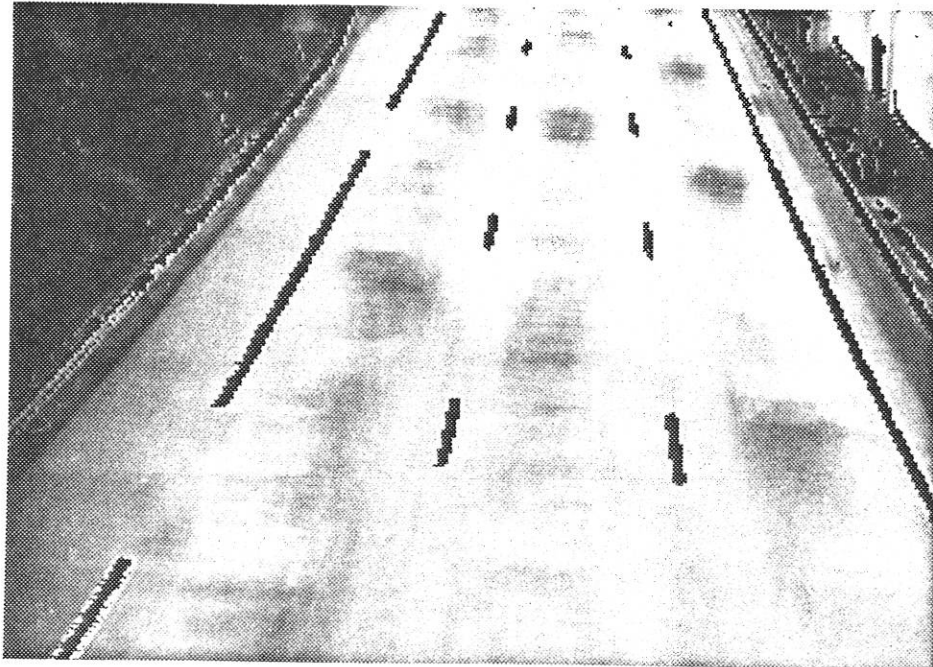
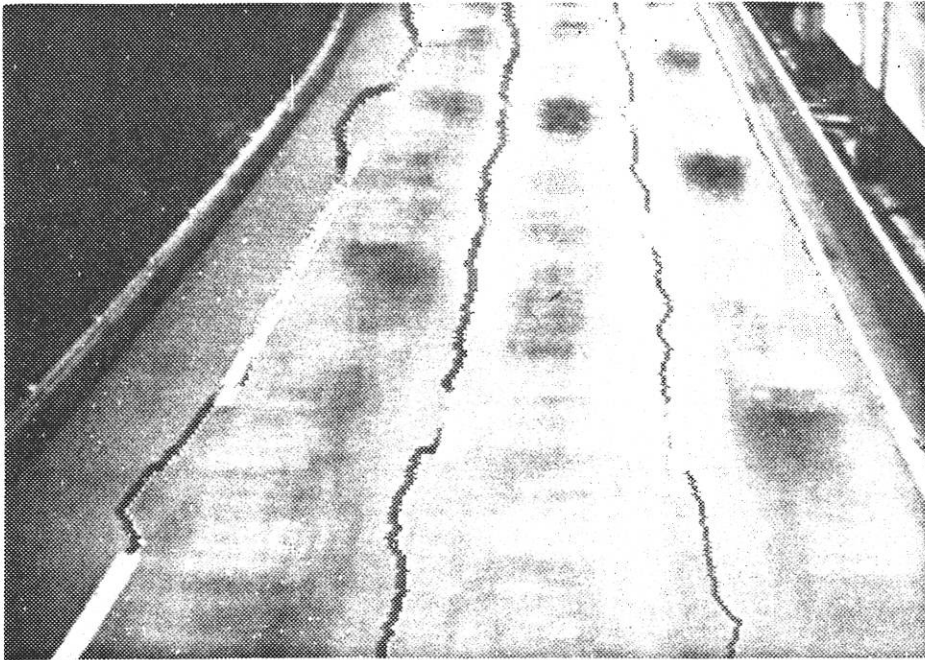


Planche II.8 : Détermination de la chaussée en utilisant l'image IMS.

(a) image IMS

(b) chapeau haut-de-forme, seuillage et filtrage

(c)



(d)

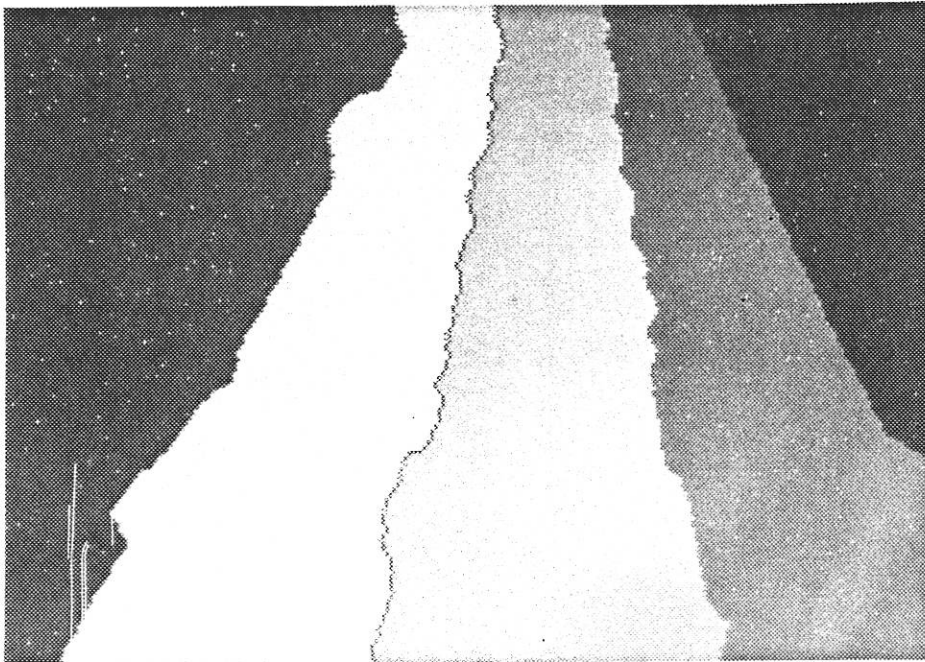
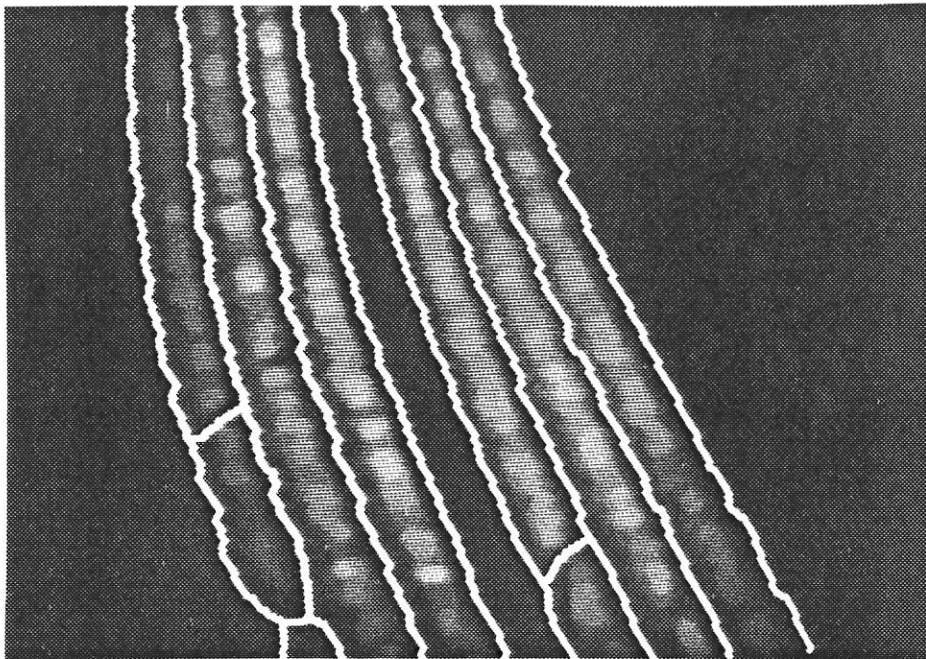


Planche II.8 (suite)

(c) SKIZ géodésique des marqueurs

(d) résultat final

(a)



(b)

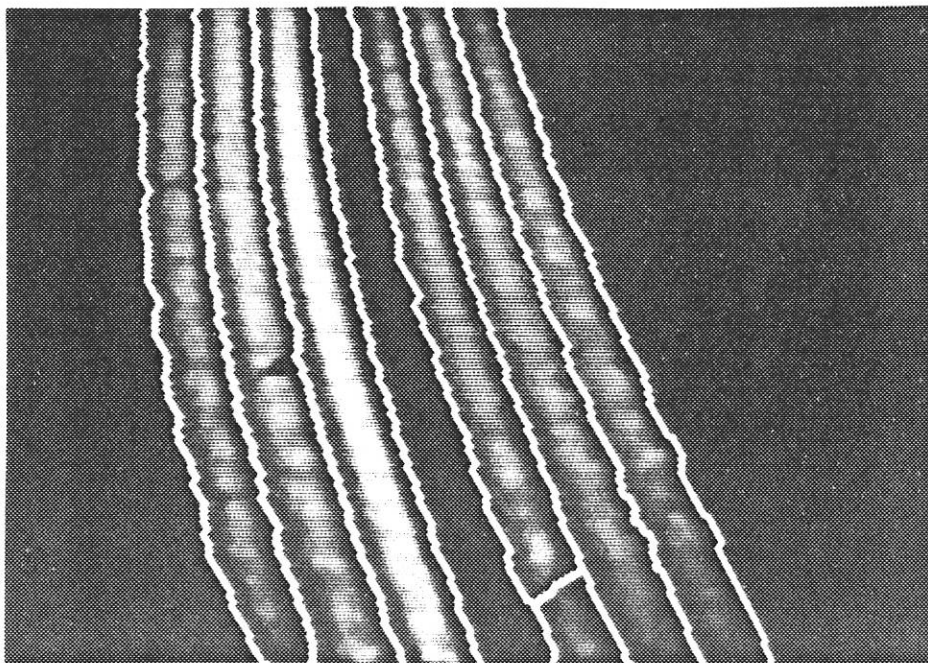
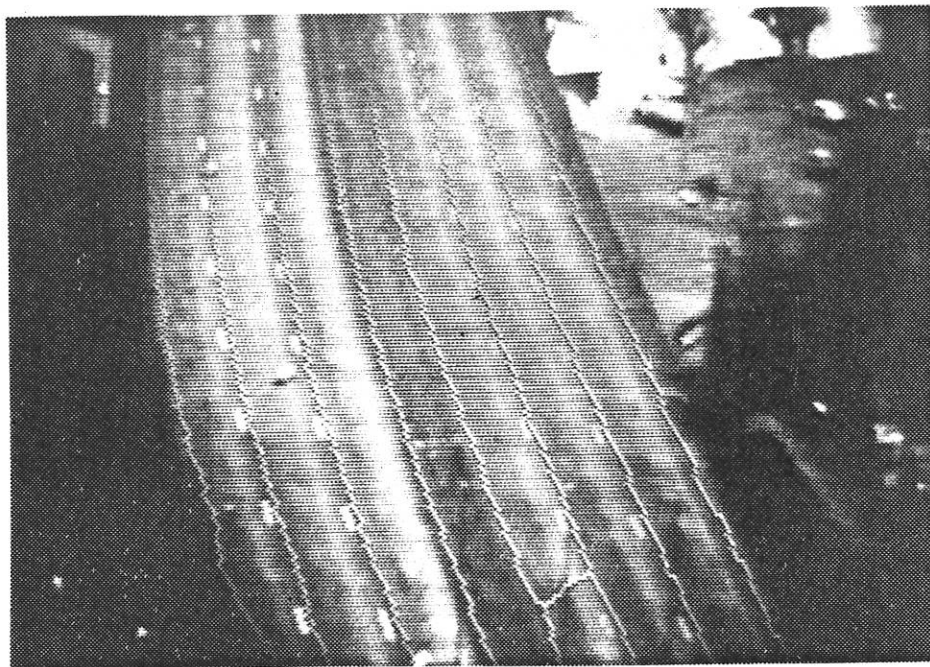


Planche II.9 : Masquage numérique adapté.

(a) 1ère séquence

(b) 2ème séquence

(c)



(d)

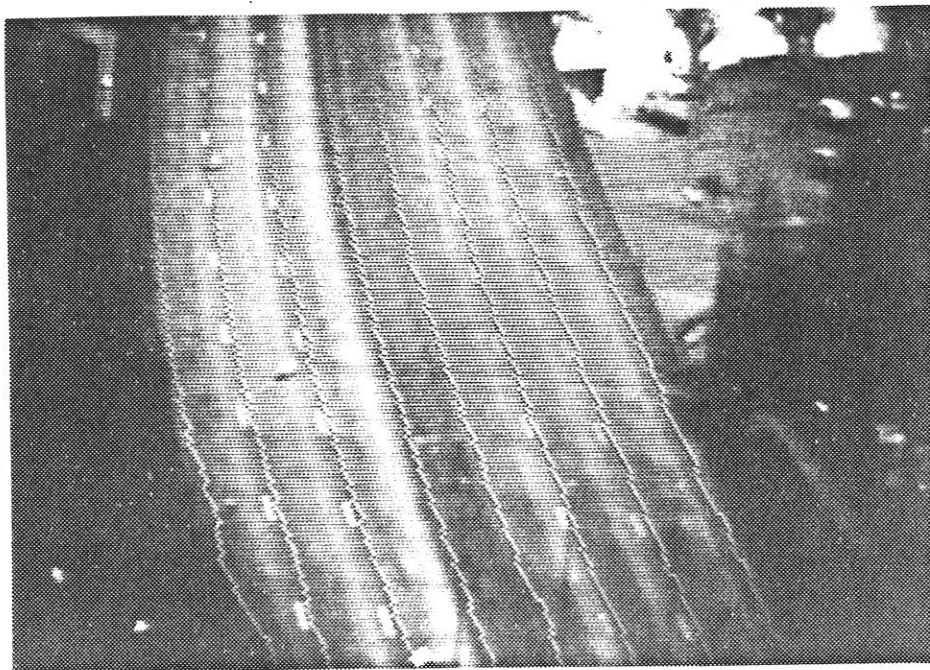


Planche II.9 (suite)

(c) élimination des segmentations intra-voies

(d) résultat final

dimensions apparentes des objets varient selon leur position dans l'image. Ce phénomène est donc gênant puisque les algorithmes utilisés pour mettre en évidence les véhicules font largement appel à des critères de taille. Il est donc nécessaire de calculer un facteur d'échelle permettant de faire correspondre, pour chaque point de l'image, à la dimension réelle exprimée en mètres une dimension apparente exprimée en pixels :

$$d_i(x) = e_x d_r(X)$$

x est un point de l'image

X est le point correspondant sur le site

d_i , dimension sur l'image

d_r , dimension réelle

e_x est le facteur d'échelle.

Ce facteur d'échelle, en toute rigueur, n'est pas un simple coefficient dépendant du point x , mais varie en fonction de l'orientation (l'objet mathématique le plus apte à représenter le facteur d'échelle est en fait un tenseur). Cependant, afin de simplifier la manipulation des formules de passage d'une dimension à l'autre, deux facteurs d'échelle scalaires e_v et e_h ont été calculés dans les directions verticales et horizontales (ces deux directions étant privilégiées dans les procédures utilisées). De plus, pour simplifier encore les formulations, on a supposé que e_v et e_h sont invariants le long d'une ligne horizontale (on se contente en fait de facteurs d'échelle moyens). Les équations fournissant les facteurs d'échelle e_v et e_h sont explicitées à l'annexe B. Elles peuvent être résumées ainsi (fig. II.11) :

$$e_h(x_i) = \frac{l_2(x_1 - x_i) - l_1(x_2 - x_i)}{(x_1 - x_2) L}$$

$$e_v(x_i) = \frac{[l_2(x_1 - x_i) - l_1(x_2 - x_i)]^2}{l_1 l_2 D (x_1 - x_2)}$$

L est la largeur de la chaussée et D la distance au sol entre les points b et c (ou a et d).

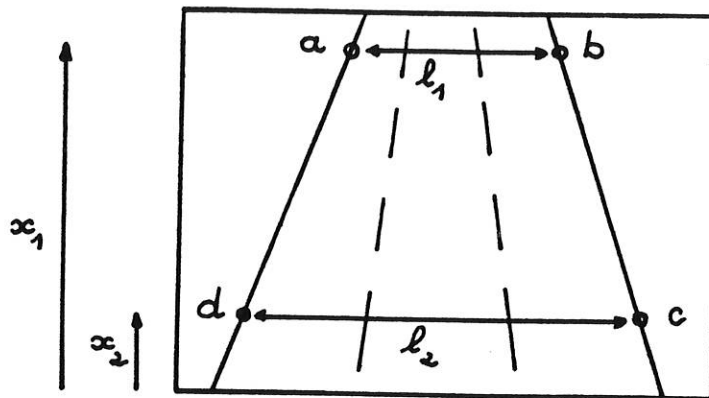


Figure II.11 : Dimensions de l'image

Les valeurs l_1 , l_2 , x_1 et x_2 peuvent être facilement calculées sur l'image du masque de la chaussée. La largeur de la chaussée est un paramètre facile à obtenir (il suffit de la mesurer). La longueur D est plus problématique. On peut cependant la calculer grâce à l'image du marquage au sol obtenue à partir de l'image IMS. Ce marquage au sol est en effet normalisé. Il suffit alors de positionner les lignes ab et cd de façon qu'elles encadrent un nombre entier de marques pointillées. Cependant, ce procédé étant un peu aléatoire (il implique que les marques au sol existent et qu'elles soient utilisables), on peut s'en dispenser, si la hauteur h de la caméra par rapport au plan de la chaussée est connue. En effet, on peut écrire :

$$e_v(x_i) = \frac{|[L^2(x_1 - x_2)^2 - (l_2 - l_1)^2 h^2]|^{1/2} [l_2(x_1 - x_i) - l_1(x_2 - x_i)]^2}{(x_1 - x_2)^2 (x_1 l_2 - x_2 l_1) L^2}$$

Cette formule un peu complexe a néanmoins l'avantage de n'utiliser que deux paramètres du site : la largeur L de la chaussée et la hauteur h de la caméra.

II-3-3) Transformations morphologiques prenant en compte la perspective

Lorsque les facteurs d'échelle sont calculés, il convient de les utiliser toutes les fois que les transformées morphologiques (érosion, dilatation ou toute combinaison de ces opérations élémentaires) dépendent d'un paramètre de taille.

Soit $T_1(X)$ une telle transformation. l est la taille exprimée en m au sol. Il ne suffit pas de subdiviser le masque d'analyse en zones où le facteur d'échelle e_x est relativement constant et d'effectuer une transformation de taille $e_x l$ dans chaque zone. On ne résoud pas ainsi les problèmes d'effets de bord, et de plus le résultat est faussé. La solution consiste à utiliser la notion de transformée géodésique et à la généraliser.

Une transformée morphologique est dite géodésique lorsqu'elle fait appel à une distance non euclidienne ([3]). Cette distance est définie dans un ensemble X , appelé ensemble géodésique : la distance entre deux points x et y appartenant à X est la longueur du plus court chemin inclus dans X et reliant x et y . Le dilaté géodésique d'un ensemble Y inclus dans X par une boule de rayon r est l'ensemble des points de l'espace situés à une distance géodésique inférieure ou égale à r de Y . D'après la définition donnée plus haut, ces points ne peuvent être qu'inclus dans X (voir annexe A).

La généralisation de cette notion consiste à définir la distance géodésique non pas simplement en terme de longueur, mais plus généralement en terme de longueur pondérée. Soit C un chemin reliant x et y dans X . La longueur de ce chemin est donnée par :

$$L(C) = \int_{xy} 1/e(s) ds$$

$e(s)$ est le facteur d'échelle le long du chemin. La distance entre x et y est la longueur du plus court chemin ainsi défini. Cette distance peut être utilisée pour définir de nouvelles transformées morphologiques de base, érosion ou dilatation (Fig. II.12).

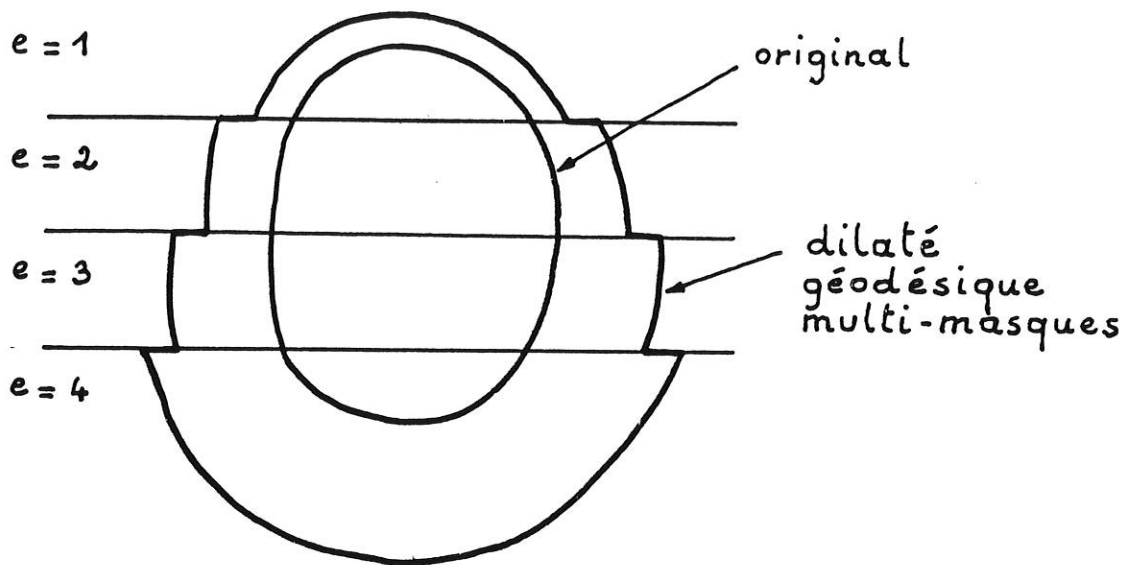


Figure II.12 : Dilatation géodésique et facteur d'échelle

La réalisation pratique de ce type de transformation soulève quelques problèmes dans le cas général. Cependant, dans le problème qui nous intéresse, cette mise en oeuvre est simplifiée car le champ des facteurs d'échelle est monotone. On effectue alors un découpage du champ d'analyse en zones où pour une taille de transformation l donnée, le nombre de pixels correspondant est constant. La transformée morphologique associée sera donc une combinaison de transformées géodésiques simples dans chaque zone. L'annexe C explicite ce type de transformation.

Ce procédé permet d'exprimer toutes les transformées en termes de distances au sol et donc de simplifier grandement l'élaboration des algorithmes de mise en évidence des véhicules.

II - 4) DETECTION DES VEHICULES

Plusieurs stratégies ont été envisagées pour extraire les véhicules. On pouvait tenir compte du mouvement, ou au contraire analyser la scène image par image. La deuxième solution a finalement été retenue, du moins dans le

cadre du traitement du trafic autoroutier, pour plusieurs raisons :

- Le premier groupe de méthodes est assez difficile à mettre en oeuvre sur des images brutes. La mise en évidence du mouvement par différence d'images ou par d'autres moyens n'est cependant pas une mise en évidence des véhicules eux-mêmes. Cette méthode n'apporte donc aucune simplification à la suite du traitement. De plus, ces procédés fonctionnent d'autant mieux que les vitesses des véhicules ne sont pas trop fluctuantes, ce qui est loin d'être le cas pour le trafic autoroutier.

- L'analyse de la scène image par image est une procédure qui permet de prendre en compte les véhicules même immobilisés. Cet avantage est considérable dans le cadre d'une utilisation du système en D.A.I.

La détection des véhicules consiste à les repérer en utilisant certaines caractéristiques visuelles et géométriques de l'image. Les véhicules se présentent en première analyse comme un assemblage de taches sombres et claires d'une certaine taille (entre 1 et 2 m de large). Ces taches sont constituées du toit, du capot, du pare-brise et de l'ensemble ombre portée-calandre (ou ombre-arrière). Cette dernière tache est souvent la plus importante. Elle est d'autant plus visible que les véhicules sont bien espacés. La détection des véhicules consistera donc, dans un premier temps, à mettre en évidence ces taches sombres et claires.

II-4-1) Détection des parties sombres

Les parties sombres sont de deux types :

- L'ombre portée à l'avant des véhicules. Cette ombre portée est toujours présente sur les images de jour, quelles que soient les conditions météorologiques, et à condition que le trafic ne soit pas trop dense.

- D'autres parties sombres, comme le pare-brise ou certaines zones des poids lourds.

Ces parties sombres sont mises en évidence par un algorithme de détection des minima régionaux de l'image (Fig. II.13).

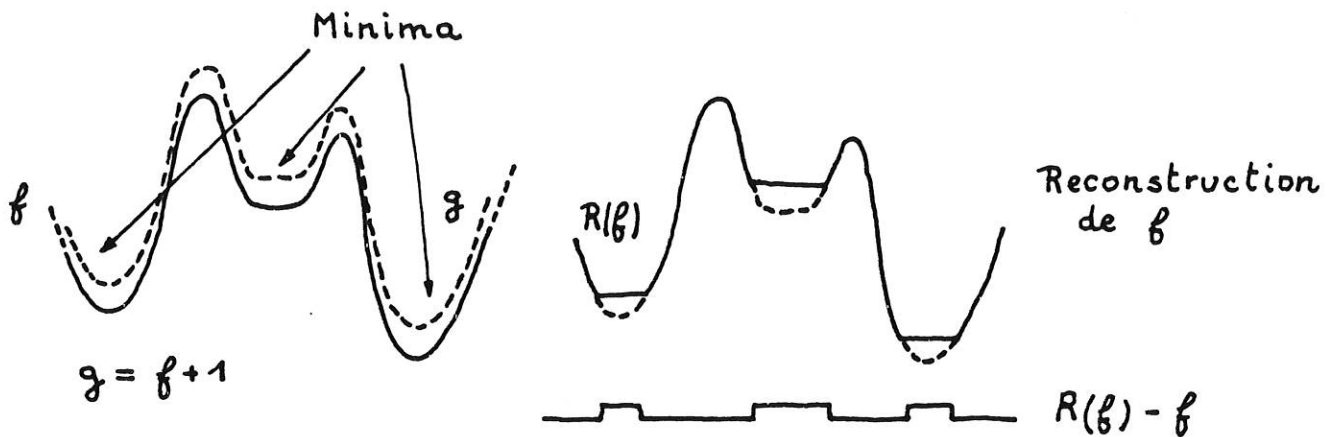


Figure II.13 : Principe de l'extraction des minima régionaux d'une image

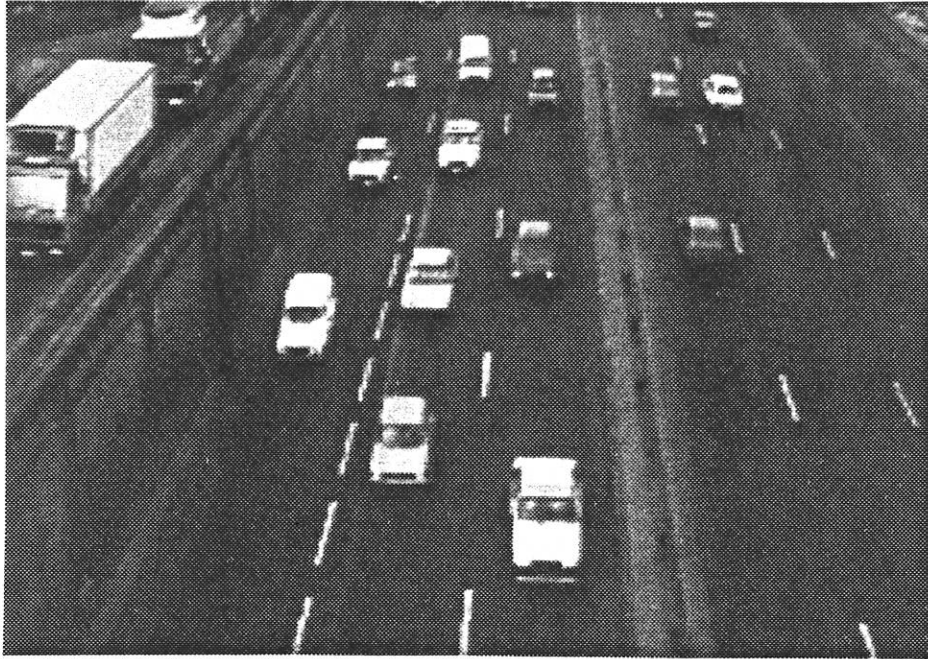
Seuls sont retenus les minima de taille suffisante, et dont le niveau de gris est inférieur au niveau moyen de la route (Fig. II.14).

Diverses procédures de seuillage des minima régionaux ont été testées. L'analyse des histogrammes de niveaux de gris et la recherche des modes de l'histogramme est la méthode utilisée à l'heure actuelle. Elle ne présente malheureusement pas toutes les propriétés de robustesse souhaitées. Diverses améliorations sont envisagées et sont décrites ci-après.

II-4-2) Détection des parties claires

Les capots et les toits peuvent être mis en évidence à l'aide d'un algorithme différent. En effet, à la différence des marqueurs sombres dont la taille et la forme peut être très variables (voir les résultats ci-dessous), les marqueurs clairs sont géométriquement mieux définis. C'est la raison pour laquelle on utilise la transformée chapeau haut-de-forme. La transformée chapeau haut-de-forme met en évidence tous les objets clairs de l'image (Fig. II.15). Un filtrage par ouvertures linéaires horizontales élimine sur l'image précédente les composantes de trop petite taille pour être des marqueurs caractéristiques.

(a)



(b)

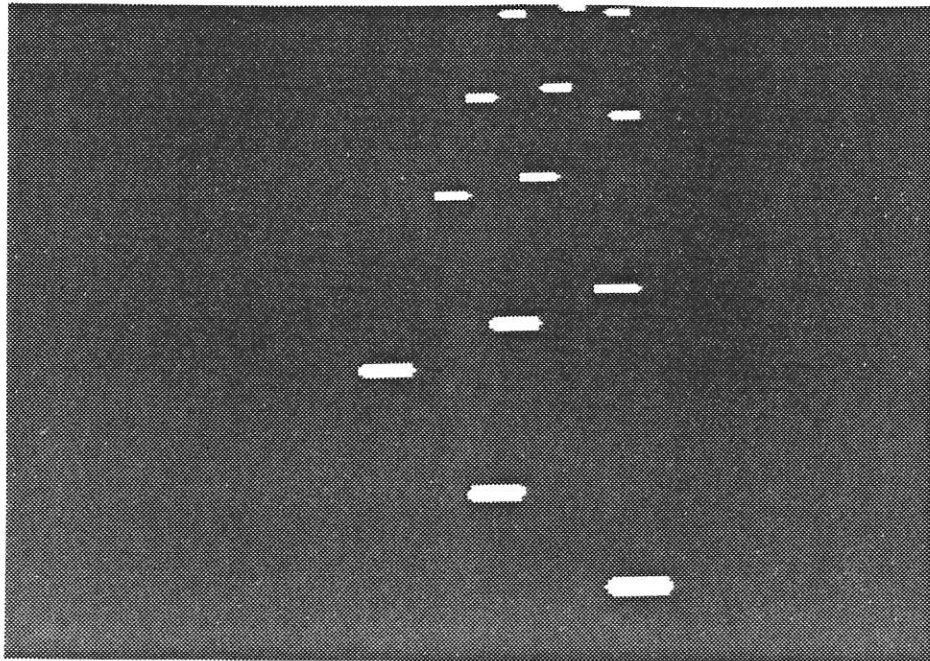


Figure II.14 : Détection des minima régionaux

(a) image originale

(b) minima extraits

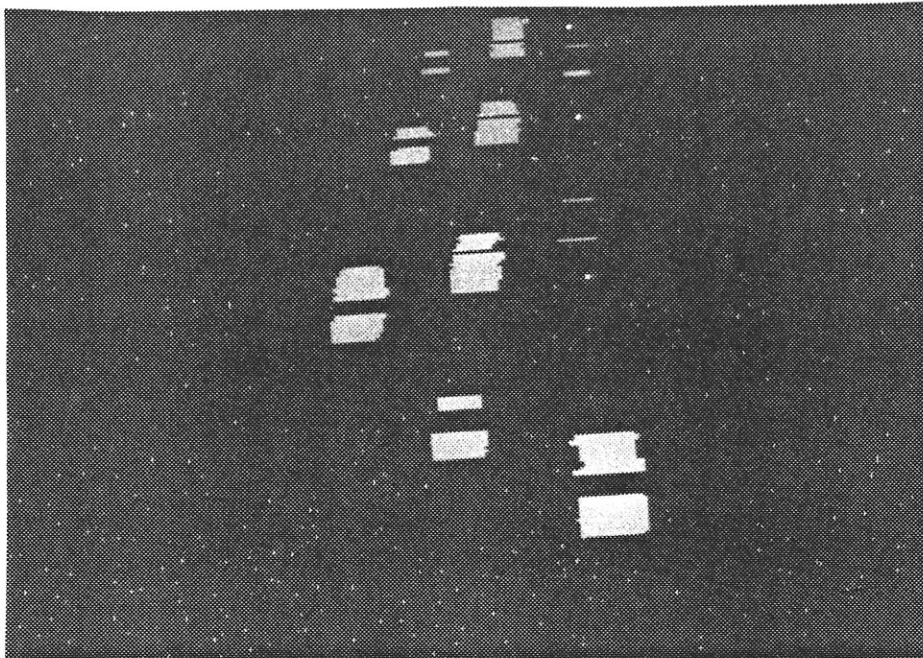


Figure II.15 : Transformation chapeau haut-de-forme.
Extraction des taches claires

II-4-3) Concaténation des marqueurs

La réunion des marqueurs obtenus par les deux précédents algorithmes doit permettre un marquage biunivoque des véhicules, c'est-à-dire qu'à chaque véhicule doit correspondre un et un seul marqueur. Ce n'est évidemment pas le cas, et il faut donc concaténer les marqueurs. Cette concaténation est une procédure hiérarchisée. Le marqueur correspondant à l'ombre portée est utilisé comme germe. Ce marqueur est toujours présent lorsque la densité de trafic n'est pas trop importante (lorsque c'est le cas d'autres algorithmes sont employés, algorithmes faisant l'objet de discussions ci-après). La concaténation s'effectue par fermeture linéaire verticale. Les marqueurs sombres sont connectés aux éventuels marqueurs clairs. La taille de cette fermeture est choisie de façon à correspondre approximativement à la longueur d'un véhicule.

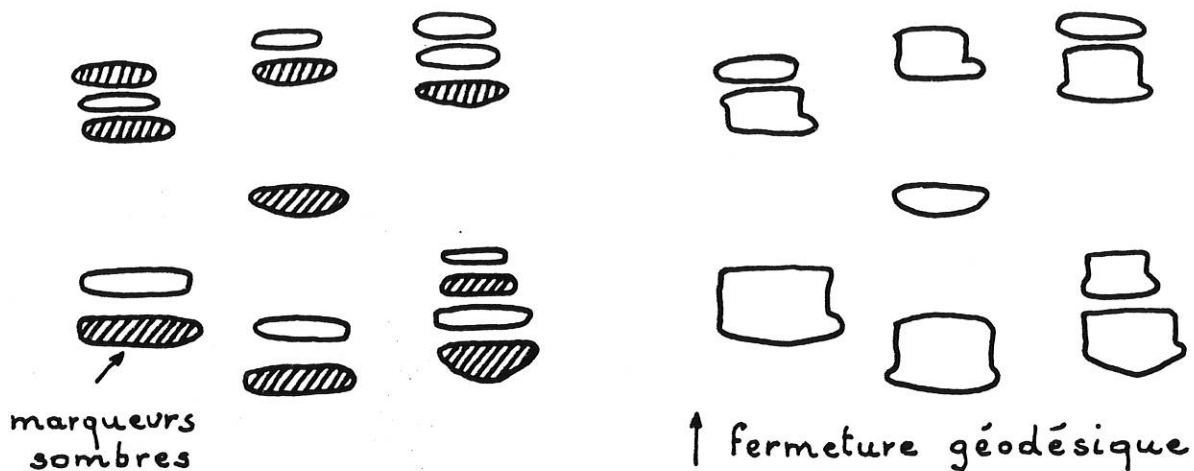


Figure II.16 : Concaténation des marqueurs 1ère phase

De plus, cette fermeture est géodésique dans le complémentaire des marqueurs blancs, ceci afin de disposer d'un garde-fou empêchant des reconnections de marqueurs trop éloignés. La figure II.16 décrit l'algorithme. Les marqueurs ainsi concaténés sont utilisés à nouveau pour concaténer d'éventuels marqueurs résiduels. La procédure est alors inversée, les marqueurs clairs servent de germes et la fermeture (dilatation et érosion géodésique dans le complémentaire des marqueurs précédemment concaténés) se fait en descendant (figure II.17).

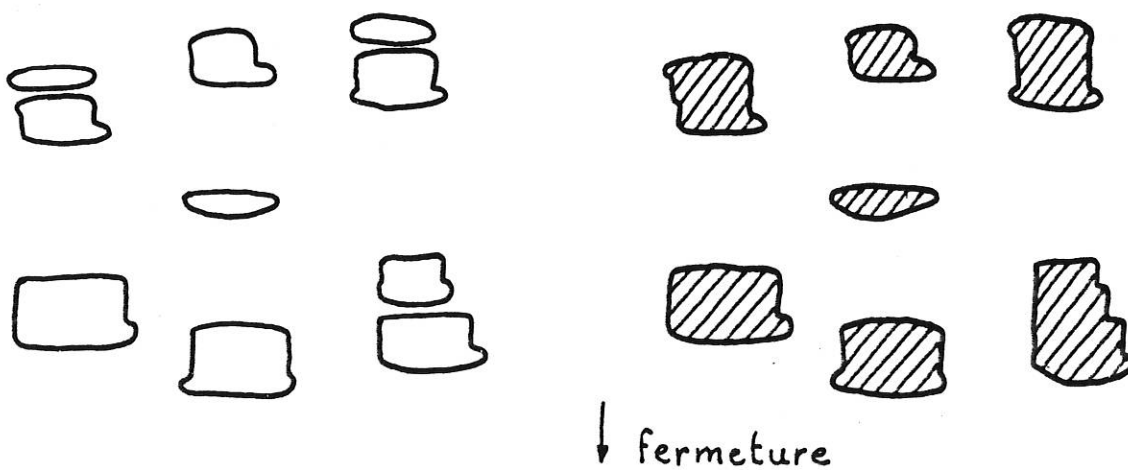


Figure II.17 : Concaténation des marqueurs 2ème phase

Le résultat final de cette procédure de concaténation est un marqueur unique par véhicule (fig. II.18).

Ce marqueur est positionné à l'avant du véhicule (c'est-à-dire au niveau du marqueur sombre ayant servi de germe) afin de pouvoir comparer aisément sa position d'une image à l'autre (Fig. II.19).

II-4-4) Discussion de la procédure - algorithmes additionnels

Plusieurs points dans la procédure actuelle de reconnaissance des véhicules méritent des améliorations. La détection des ombres à l'avant des véhicules utilise un seuillage. Ce seuillage doit être automatique. Il suppose que les minima régionaux de l'image se répartissent en deux groupes: les minima sombres correspondant effectivement à la zone d'ombre à l'avant du véhicule et des minima généralement plus clairs dus à des artefacts (taches d'huile, plaques de macadam, etc...) présents sur la chaussée. Si cette hypothèse est le plus souvent exacte, il n'en demeure pas moins que la répartition en deux modes des histogrammes des valeurs de gris des minima est loin d'être évidente (Fig. II.20).

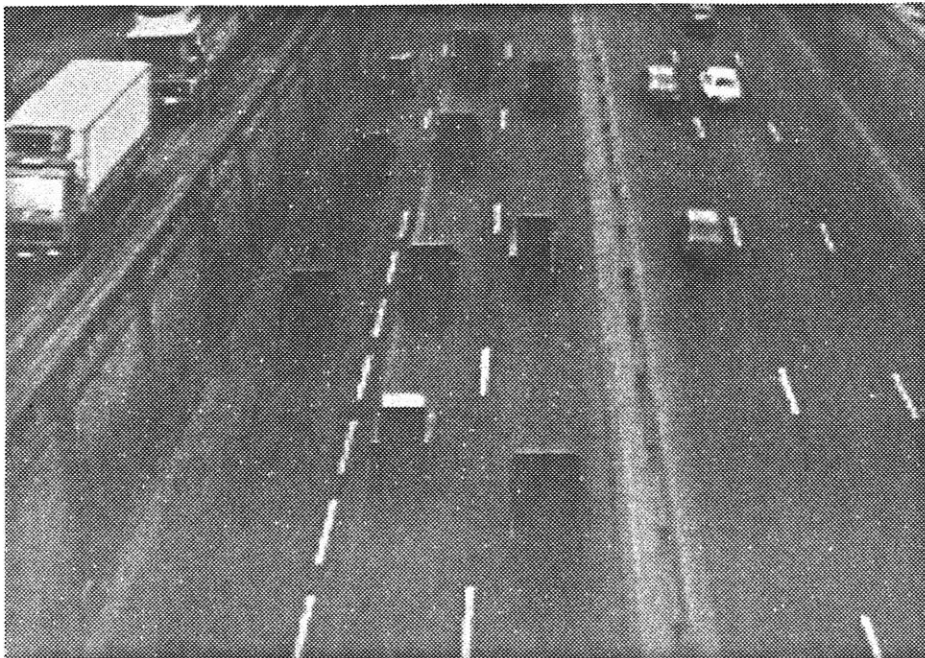


Figure II.18 : Marquage des véhicules. Résultat de la concaténation.

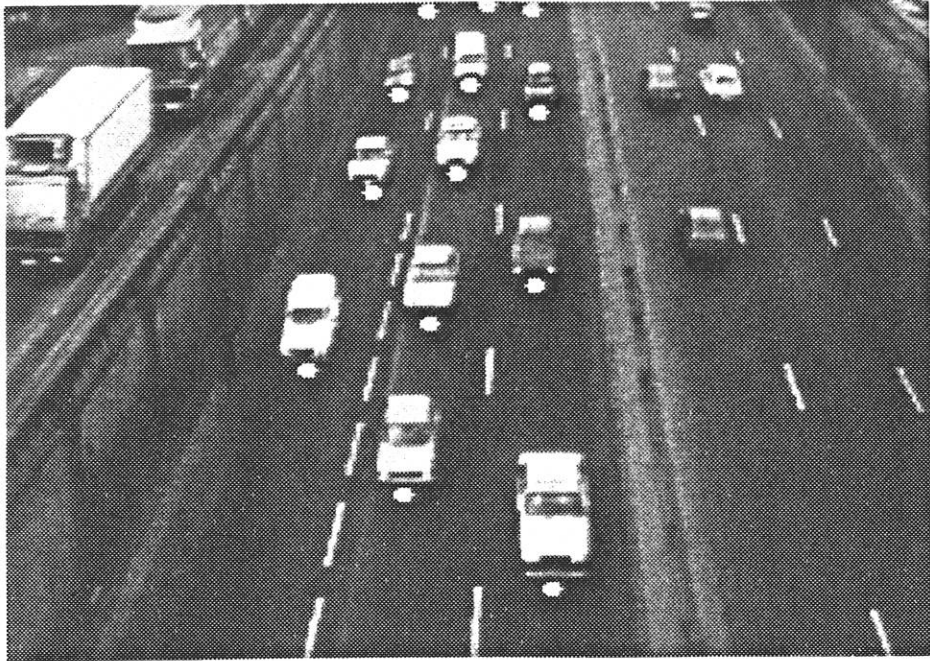


Figure II.19 : Marquage des véhicules. Résultat final.

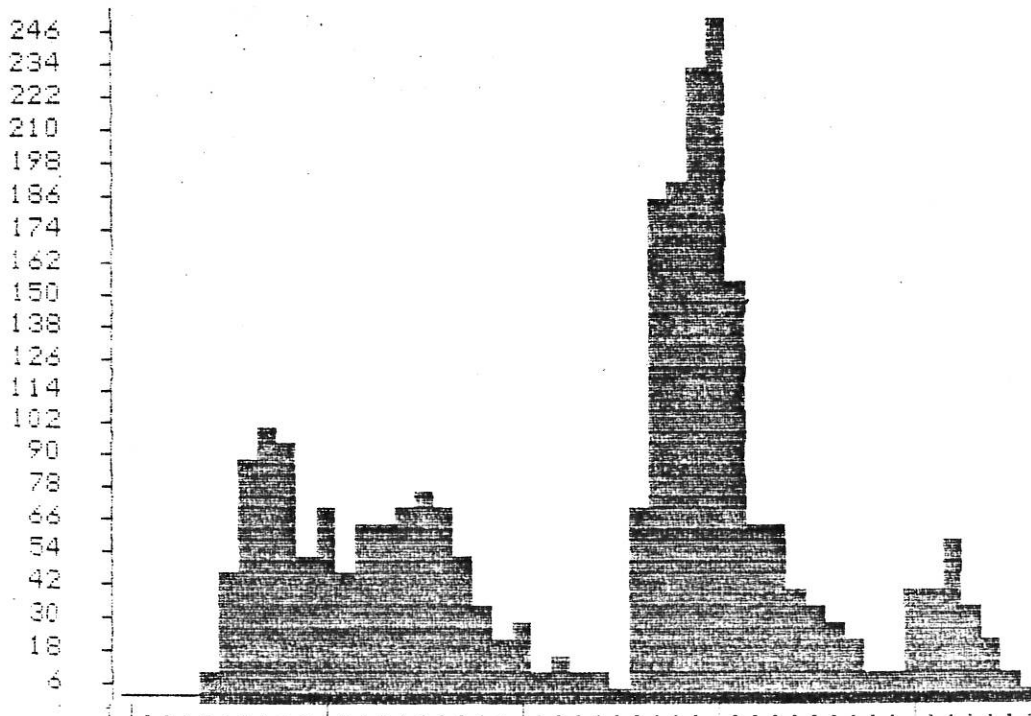


Figure II.20 : Exemple d'histogramme des valeurs de gris des minima régionaux.

D'autres méthodes de sélection sont à l'étude. Citons simplement un procédé de séparation basé non pas sur la valeur de gris des minima mais sur leur profondeur (les minima dus à l'ombre étant généralement plus contrastés que ceux dus à des artefacts). La profondeur du minimum peut être analysée par reconstruction (Fig. II.21).

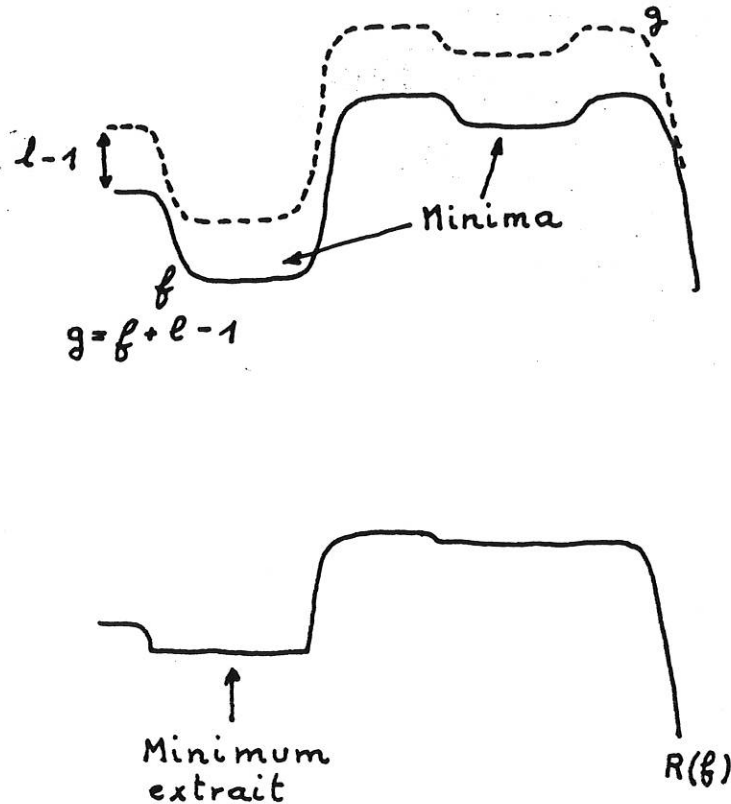


Figure II.21 : Principe de l'extraction de minima de profondeur 1.

Les calculs de facteurs d'échelle prennent en compte, on l'a vu, la hauteur de la caméra par rapport au sol. Cependant, certains marqueurs et en particulier les marqueurs clairs ne sont pas au niveau du sol. Les capots des véhicules sont à 80 cm du sol et les toits à 1,50 m. Les hauteurs des marqueurs sont encore plus importantes pour les camionnettes et les poids lourds. Ce biais introduit une différence importante dans les facteurs d'échelle, biais que l'on peut corriger en redéfinissant les algorithmes de reconnaissance des marqueurs clairs en tenant compte de cette différence de hauteur (Fig. II.22).

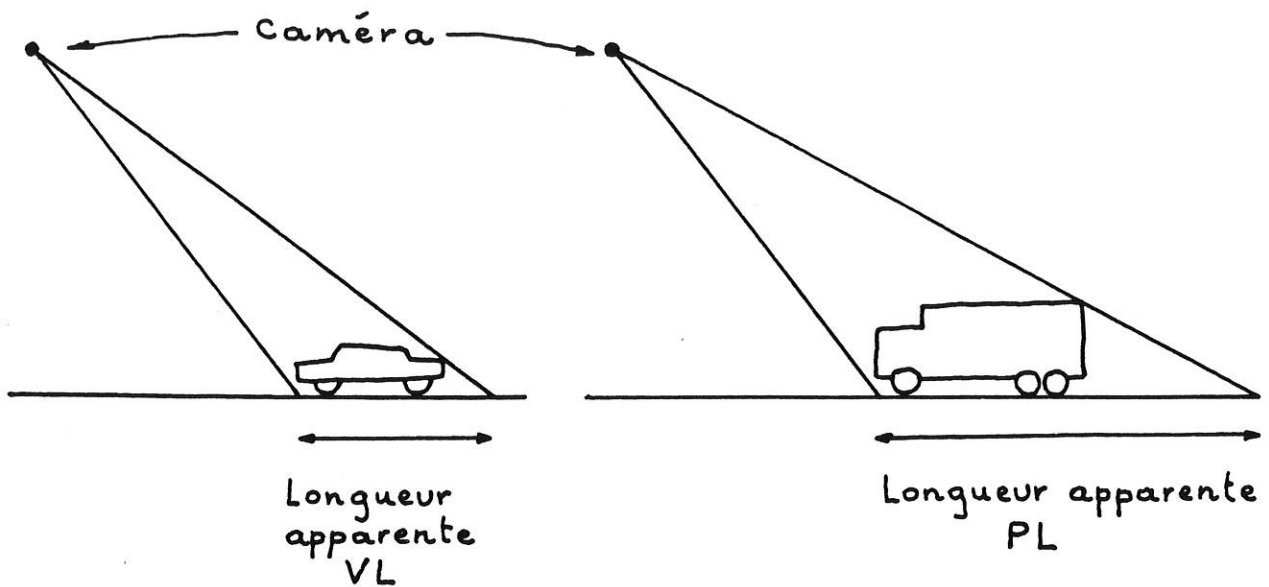


Figure II.22 : Biais dans les facteurs d'échelle
dû à la hauteur des véhicules

Les poids lourds sont plus difficiles à mettre en évidence. Le biais décrit précédemment est encore plus marqué. Il fait que souvent, seul le tracteur est correctement détecté. Il est donc nécessaire d'introduire des algorithmes additionnels pour mettre en évidence les poids lourds. Ces algorithmes sont basés sur une analyse granulométrique de la scène. Décrivons succinctement l'un d'eux :

L'image $|im_1 - IMS|$ est générée. Tout objet de niveau de gris différent du niveau moyen de la chaussée est ainsi mis en évidence. On procède alors à un filtrage alterné séquentiel sur l'image afin d'améliorer la détection. Enfin, une granulométrie par ouverture (exprimée en dimensions réelles) permet de séparer les gros objets (supposés être des poids lourds) des petits (supposés être des véhicules légers). Il ne reste plus alors qu'à vérifier que ces gros objets sont associés à des marqueurs détectés par l'algorithme décrit plus haut.

Toutes ces procédures ne sont pas fiables à 100% et l'analyse des trajectoires qui fait l'objet de la troisième phase du traitement apporte des informations fort utiles à l'élimination des fausses détections et à la

classification des types de véhicules.

II - 5) TRAJECTOIRES ET MESURES DE TRAFIC

Pour cette dernière étape de mesure des paramètres de trafic ,on dispose, pour tout instant t d'analyse des positions des différents véhicules repérés sur chaque voie (coordonnées des marqueurs précédemment mis en évidence). En pratique on se limitera dans l'historique des positionnements en choisissant une fenêtre temporelle pas trop importante et on bâtit l'espace des mesures au fur et à mesure des traitements.

Une manière synthétique de représenter les résultats consiste à marquer, pour chaque voie, les différentes positions des véhicules détectés, projetées sur l'axe médian de la voie. Cette opération répétée pour tous les instants de détection permet de constituer la carte des positions successives des véhicules.

Si l'on veut avoir une description fidèle des emplacements des véhicules, il importe de faire la représentation décrite ci-dessus dans l'espace réel. Il convient donc de corriger la perspective de chaque image. Ceci se fait simplement en utilisant les facteurs d'échelle. Tout se passe alors dans la représentation corrigée comme si les véhicules étaient vus d'en haut.

A partir de la représentation de ces points censés correspondre à des véhicules on s'efforce, en suivant des règles précises, de reconstituer des trajectoires spatio-temporelles qui permettent de relier tous les points affectés au même véhicule.

II-5-1) Constitution des trajectoires

Pour pouvoir construire les trajectoires spatio-temporelles, il faut d'abord affecter une voie de circulation à chaque véhicule, puis convertir sa position dans l'espace image en une position dans l'espace réel. On ne conserve alors que la position du véhicule sur l'axe de la chaussée, la position latérale par rapport à cet axe étant de peu d'intérêt. En répétant ces opérations pour plusieurs instants de détection on peut alors construire les trajectoires selon les règles décrites plus loin.

II-5-1-1) Localisation des véhicules détectés

Les opérations décrites dans les chapitres précédents aboutissent à une image contenant des marqueurs uniques des véhicules, marqueurs présents sur toutes les voies de circulation considérées (Fig. II.23). Les traitements ultérieurs s'effectuant voie par voie, on utilisera les différents masques de voie définis à l'étape d'initialisation pour isoler chacune d'elle (cf Fig. II.4).

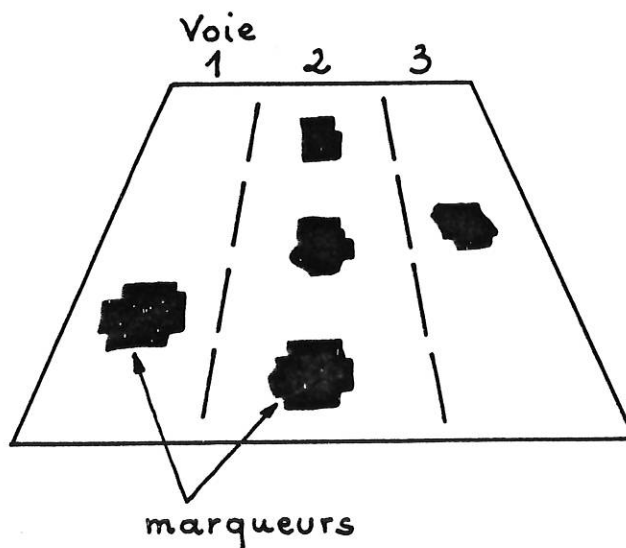


Figure II.23 : marqueurs complets des véhicules repérés

Après étiquetage des marqueurs sur chaque voie, on leur associe les coordonnées du centre de la partie inférieure du marqueur (Fig. II.24). Cette zone est préférentiellement choisie car elle correspond à l'ombre portée du véhicule. Ce marqueur est donc insensible à la hauteur du véhicule, et permet ainsi une plus grande précision des mesures ultérieures.

Comme on ne s'intéresse pas à la position latérale du véhicule à l'intérieur de sa voie de circulation, on peut réduire sa position à une seule coordonnée: celle mesurée sur l'axe de la voie (x_1 , x_2 , x_3 dans la Figure II.24).

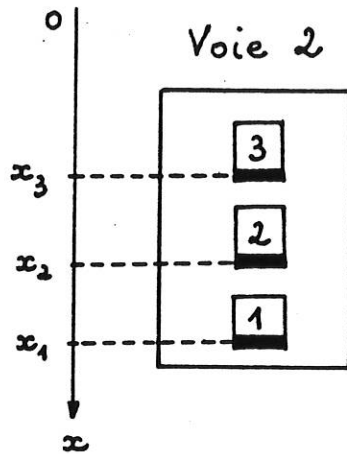


Figure II.24 : voie analysée et coordonnées retenues

La position réelle du véhicule sur l'axe de la chaussée se déduit facilement de cette unique coordonnée en utilisant le facteur d'échelle vertical e_v précédemment calculé (Fig. II.25).

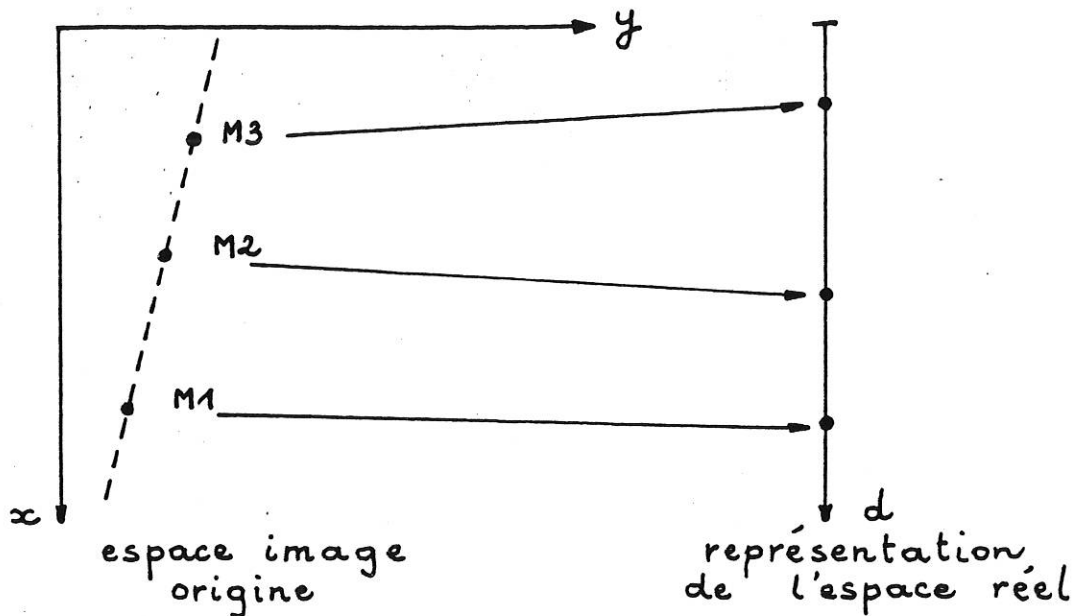


Fig. II.25 : Changements de coordonnées dans les différents repères

II-5-1-2) Construction des trajectoires spatio-temporelles

Après avoir répété les différentes opérations décrites ci-dessus pour la séquence d'images (c'est-à-dire l'intervalle de temps) que l'on veut considérer, on représente les coordonnées des véhicules sur un même

graphique, en déroulant la séquence suivant l'axe horizontal (Fig. II.26).

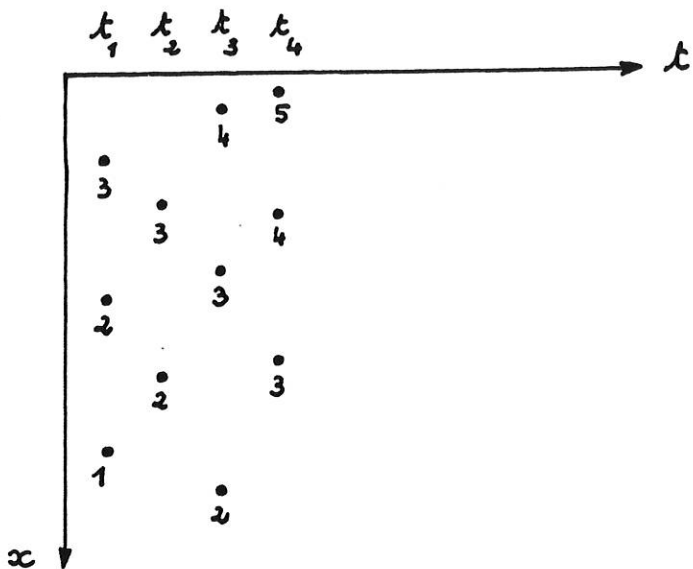


Figure II.26 :Construction de la représentation spatio-temporelle
 Les positions des mêmes véhicules sont repérées par les mêmes numéros.

Ainsi, pour une séquence d'une vingtaine d'images, on obtient un nuage de points (Fig. II.27) laissant déjà pressentir le graphe final.

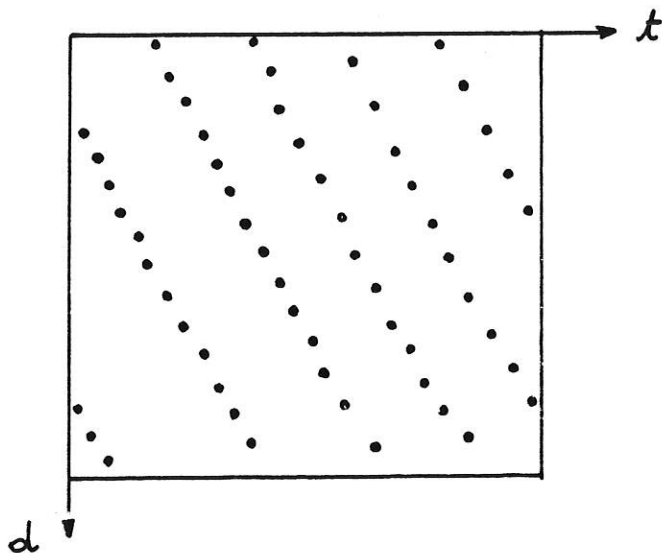


Figure II.27 : Image avant construction des trajectoires

La construction des trajectoires est alors effectuée. Elle utilise l'algorithme suivant:

Soit un véhicule i . Considérons un point $M(t)$ de coordonnées $(t, x_i(t))$ appartenant à sa trajectoire C_i . Un point $M(t+1)$, de coordonnées

$(t+1, x(t+1))$ ou à défaut $M(t+2)$ ou $M(t+3)$ sera considéré comme un point de la courbe C_i , si la vitesse du véhicule estimée entre les instants t et $t+1$ en utilisant ce doublet de points n'est pas trop différente de la vitesse $V_i(t)$ à l'instant t , ce qui peut s'exprimer par:

$$\frac{1}{\Delta t} [x(t+1) - x_i(t)] \in [(1-\alpha) V_i(t), (1+\alpha) V_i(t)]$$

Si cette équation est vérifiée, alors le point $M(t+1)$ sera considéré comme un point de la trajectoire C_i :

$$x(t+1) = x_i(t+1)$$

Le coefficient α dépend de l'accélération ou décélération limites acceptables du véhicule.

Le processus peut alors être réitéré. La vitesse à l'instant $t+1$ est calculée selon la formule :

$$V_i(t+1) = \beta V_i(t) + (1-\beta) \left[\frac{x_i(t+1) - x_i(t)}{\Delta t} \right]$$

où β est un coefficient de pondération qui peut dépendre de la dernière accélération calculée.

Dans l'état actuel de l'algorithme, la vitesse $V_i(t_0)$ attribuée à un véhicule à son entrée dans la scène est une vitesse moyenne de flot calculée en utilisant d'autres critères. Il serait plus performant de fixer cette vitesse initiale comme une fonction de la concentration ρ dans la section en utilisant la formule suivante, déjà largement utilisée en théorie du trafic :

$$V = V_f \left[1 - \frac{\rho}{\rho_{\max}} \right]^{1/m}$$

avec ρ_{\max} , la concentration maximum dans la section

V_f , la vitesse moyenne dans la section en trafic fluide

l, m , deux paramètres à calibrer ($1/m \approx 0.5$)

La constitution des trajectoires sera d'autant plus aisée que

l'intervalle de temps Δt entre deux images successives analysées sera faible. Le choix de Δt peut être fait de façon à être très inférieur à l'intervalle temporel δ_t entre 2 véhicules.

$$\Delta t \ll \delta_t = \frac{\delta_x}{v}$$

Avec δ_t , l'intervalle temporel entre 2 véhicules
 δ_x , l'intervalle spatial entre 2 véhicules
 v , la vitesse du premier véhicule.

On minimise ainsi la possibilité qu'une portion de trajectoire du véhicule $i+1$ se trouve dans le quadrant de recherche du véhicule i . Le calcul d'une vitesse moyenne de flot par simple appariement est même possible (Fig. II.28).

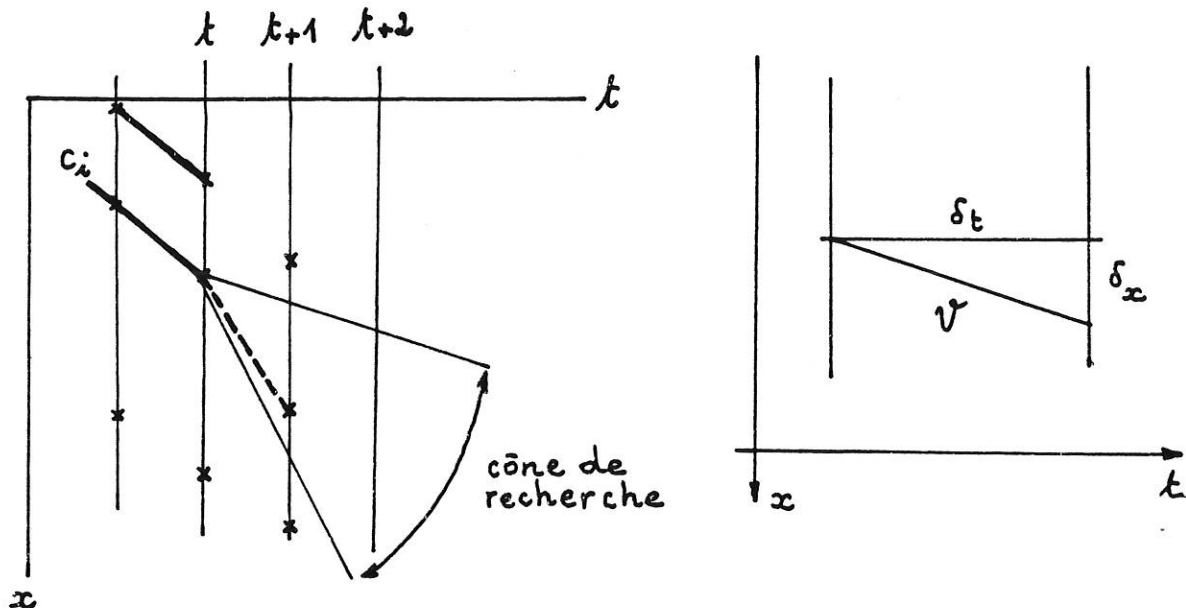


Figure II.28 : Relation géométrique entre la vitesse et la fréquence d'acquisition.

Par exemple, si Δt est égal à $1/4s$, l'intervalle temporel entre deux véhicules en supposant la vitesse v égale à 120 km/h serait également égal à $1/4s$, ce qui correspond à un débit de 14000 véhicules/heure par voie. Rappelons que les valeurs maxima sont dans la réalité de l'ordre de 2500 véhicules/heure.

II-5-2) Mesures

Cette représentation en coordonnées spatio-temporelles des marqueurs des véhicules voie par voie permet en reconstituant les trajectoires, de vérifier a posteriori la validité de l'algorithme de détection et de corriger d'éventuelles erreurs. Elle permet également la mise en évidence de certains évènements (changement de voie, véhicules arrêtés...). Enfin elle offre la possibilité d'extraire très facilement de nombreux paramètres de circulation. Les points (x,t) du graphe correspondant à un même véhicule i forment une courbe que nous désignerons par C_i . La figure II.29 illustre une telle représentation sur une séquence de 20 images.

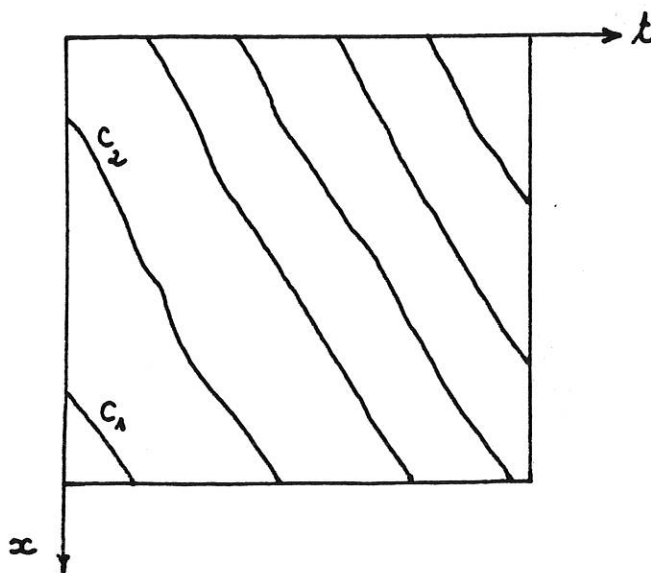


Figure II.29 : Trajectoires après reconstitution

II-5-2-1) Mesure de débit

Le nombre N_D de points d'intersection de l'ensemble des courbes C_i avec une droite horizontale d'équation $x=a$ représente le nombre de véhicules sur la voie qui sont passés au point $x=a$ pendant l'intervalle de temps considéré $[t_0, t_n]$. Le débit D en véhicules de la voie sera donné par la relation :

$$D = \frac{N_D}{t_n - t_0}$$

II-5-2-2) Concentration

Quant-au nombre N_C de points d'intersection de l'ensemble des courbes

C_i avec une droite verticale $t=a$, il représente le nombre de véhicules sur la voie qui se trouvent dans le champ de caméra à l'instant a . Connaissant la longueur L de la partie de voie qui figure sur les images, nous pouvons en déduire la concentration C de véhicules de la voie, c'est-à-dire la valeur moyenne du nombre de véhicules par kilomètre de voie :

$$C = \frac{N_c}{L}$$

II-5-2-3) Estimation du taux d'occupation

On peut sur la représentation spatio-temporelle estimer le taux d'occupation d'une boucle, grandeur traditionnellement utilisée en trafic. Ce taux d'occupation est égal au rapport du temps où la boucle est occupée au temps total.

Si on place sur l'axe vertical, à l'endroit où l'on veut mesurer le taux d'occupation (en abrégé T.O.), un segment de longueur L égale à la somme de la longueur moyenne des véhicules et de la longueur de la boucle, le rabattement sur l'axe des temps de cette bande de largeur L fournit pour chaque véhicule i l'intervalle de temps Δt_i qui lui a été nécessaire pour parcourir cette distance (Fig. II.30). Le taux d'occupation est alors donné par:

$$r = \frac{\sum \Delta t_i}{t_n - t_0}$$

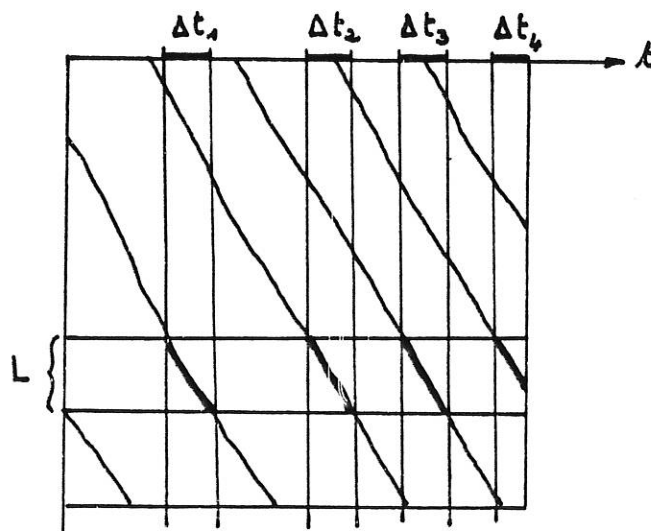


Figure II.30 : Estimation du T.O. sur l'image des trajectoires

II-5-2-4) Vitesse instantanée des véhicules

La vitesse instantanée, calculée entre deux instants de détection, est la pente de la trajectoire entre ces deux points consécutifs:

$$V_i(t) = (x_i(t) - x_i(t-\Delta t)) / \Delta t$$

Δt étant le pas d'échantillonnage temporel.

II-5-2-5) Vitesse moyenne temporelle d'un véhicule

C'est pour un véhicule, la moyenne arithmétique des vitesses calculées entre 2 instants de détection successifs t_j, t_{j-1} :

$$\bar{V}_i = \frac{1}{n} \sum_{j=1}^n \frac{x_i(t_j) - x_i(t_{j-1})}{t_j - t_{j-1}}$$

$[t_0, x_i(t_0)]$, $[t_n, x_i(t_n)]$ sont les points de détection correspondant aux débuts et fins d'une trajectoire C_i .

II-5-2-6) Vitesse moyenne spatiale

C'est entre 2 instants donnés la moyenne harmonique des vitesses d'un ou de plusieurs véhicules détectés. C'est aussi à un facteur près l'inverse du temps total de traversée de la scène :

$$U(t_j) = \frac{n_v}{\sum_{i=1}^{n_v} \frac{t_j - t_{j-1}}{x_i(t_j) - x_i(t_{j-1})}}$$

n_v est le nombre de véhicules.

II-5-2-7) Taux d'occupation spatial (T.O.S)

Le taux d'occupation spatial est une notion intéressante à développer. En effet, il peut être facilement mesuré par un capteur utilisant l'analyse d'image et est un bon estimateur de la concentration. Il peut aussi apporter une solution au problème de la validité des mesures dans le cas de fortes saturations où l'individualisation des véhicules n'est plus possible et par conséquent où les algorithmes exposés plus haut sont inopérants.

Le taux d'occupation spatial peut se définir ainsi: c'est le rapport de la longueur de la chaussée occupée par les véhicules sur la longueur totale de la chaussée observée (Fig. II.31).

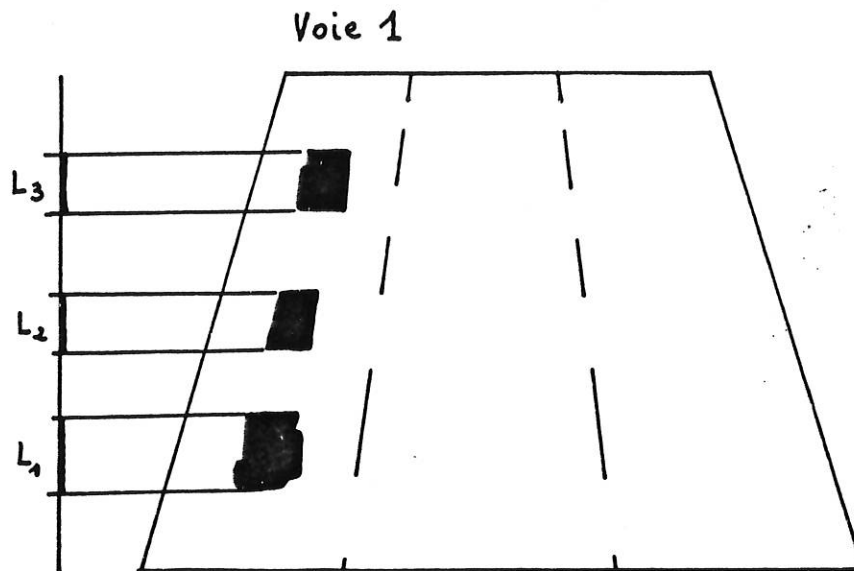


Figure II.31 : Image des marqueurs et taux d'occupation spatiale

Cette mesure se fera sur l'image des marqueurs obtenue à la fin des opérations de concaténation, afin de prendre en compte les longueurs apparentes L_i des véhicules.

La longueur réelle d'un véhicule est alors égale à :

$$l_i = \int_{L_i} e(v) dv$$

$e(v)$ est le facteur d'échelle vertical et le T.O.S. est donné par:

$$\text{T.O.S} = \sum l_i / l \quad (l, \text{ longueur réelle du tronçon})$$

Nous nous sommes délibérément limités ici à décrire les variables de trafic les plus communes, mais la richesse de l'information recueillie et sa représentation, qui permet notamment la saisie de plusieurs trajectoires, rend accessibles de nombreuses autres mesures (temps et distances inter-véhiculaires, accélération/décélération, etc...).

II-5-3) Détection d'incidents

L'autre intérêt de la représentation des trajectoires est qu'elle fournit une image d'une situation de trafic pendant un intervalle de temps donné. Cette image peut donc être analysée afin d'en extraire des caractéristiques d'événements particuliers. L'allure générale des trajectoires, leur répartition spatiale sont représentatives d'événements ou d'incidents précis se déroulant sur la voie de circulation inspectée. On peut ainsi mettre en évidence des véhicules arrêtés, les changements de voie, les ralentissements brutaux, les ondes de choc (voir par exemple la planche III.10).

L'analyse d'image et les transformées morphologiques sont là encore d'un grand secours. En effet, un événement particulier pouvant être associé à une forme caractéristique de la représentation des trajectoires, on conçoit immédiatement l'intérêt de méthodes de reconnaissance de formes pour quantifier l'image des trajectoires et en bâtir la typologie. La figure II.32 illustre l'"aspect" d'un tel événement sur la courbe des trajectoires. On imagine également facilement l'usage de transformations d'images pour étudier les relations de voisinages entre les véhicules.

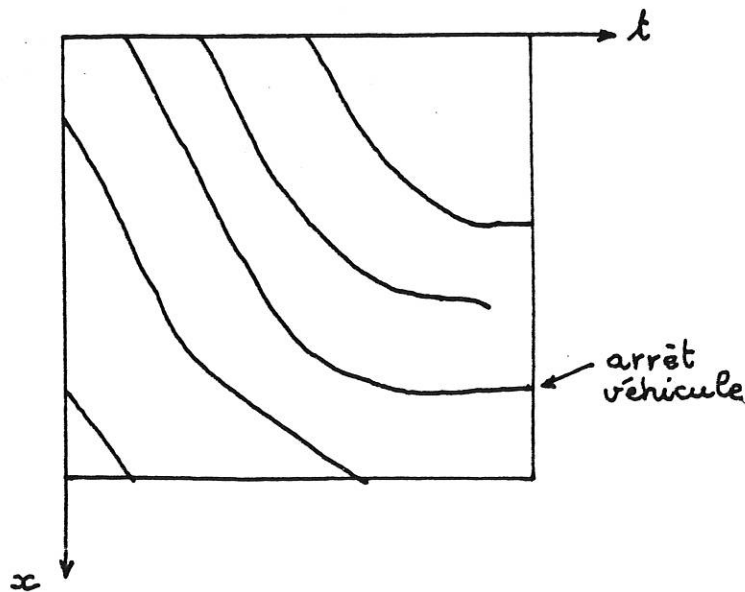


Figure II.32 : Arrêt d'un véhicule sur une voie
et représentation associée

(Le véhicule suivant s'arrête puis change de voie.)

La représentation des trajectoires, par la synthèse qu'elle donne des informations spatiales et temporelles, s'avère donc être un puissant outil en D.A.I.. C'est à terme son intérêt majeur.

RESULTATS

Les algorithmes décrits précédemment ont été appliqués à toutes les séries d'images-tests disponibles. Les trois grandes étapes du traitement (masquage de la chaussée, détection des véhicules, reconstruction des trajectoires et mesures) ont été testées dans les différentes situations rencontrées.

III - 1) TRAITEMENT DES IMAGES : ILLUSTRATIONS ET COMMENTAIRES

Les planches suivantes illustrent succinctement les résultats obtenus sur ces images. Un bref commentaire souligne dans chaque cas les particularités du site étudié et les difficultés rencontrées.

L'expérimentation sur des sites très variés permet ainsi de mieux apprécier les conditions d'application de ce type de capteur et de dégager un cahier des charges élémentaire pour son implantation. Les résultats les plus intéressants sont obtenus dans les conditions suivantes: caméra placée à une hauteur importante (20 m ou plus), au milieu de la chaussée, angle élevé par rapport à la chaussée.

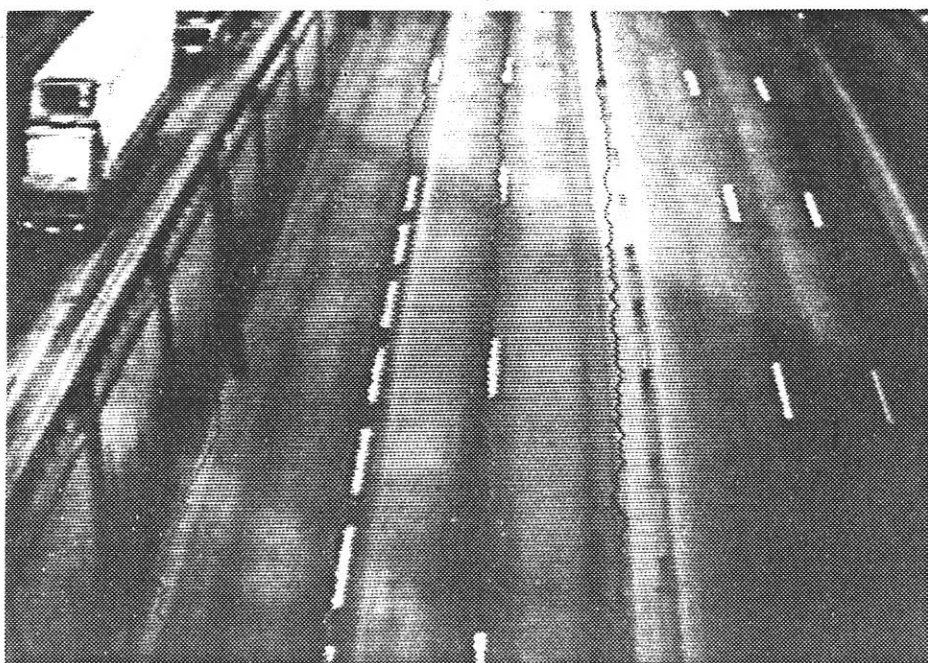
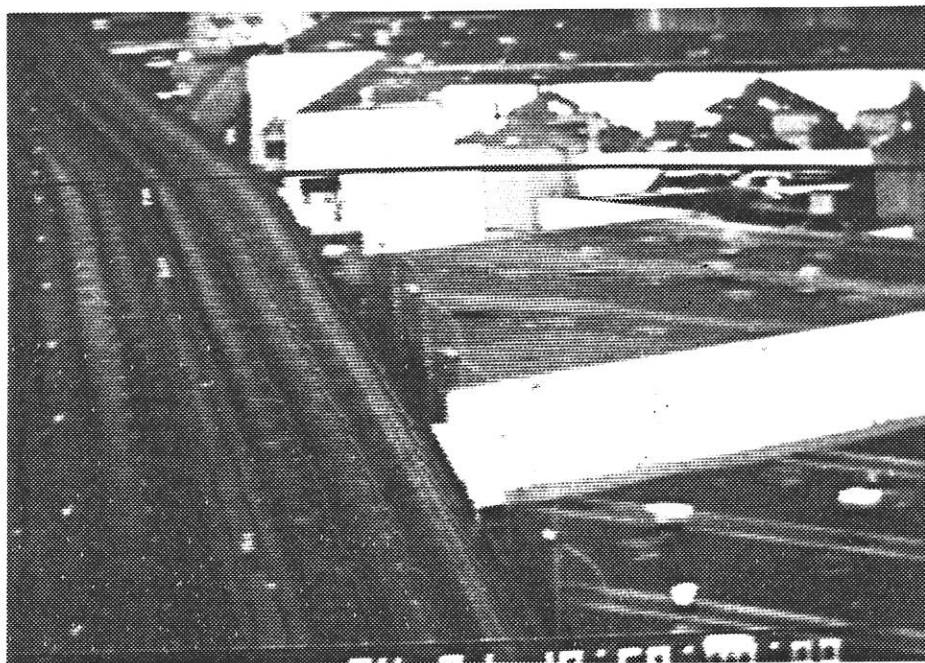


Planche III.1

Masquage de la chaussée, algorithme simplifié.

Cet exemple montre la difficulté qu'il y a parfois à travailler sur la totalité de la scène.

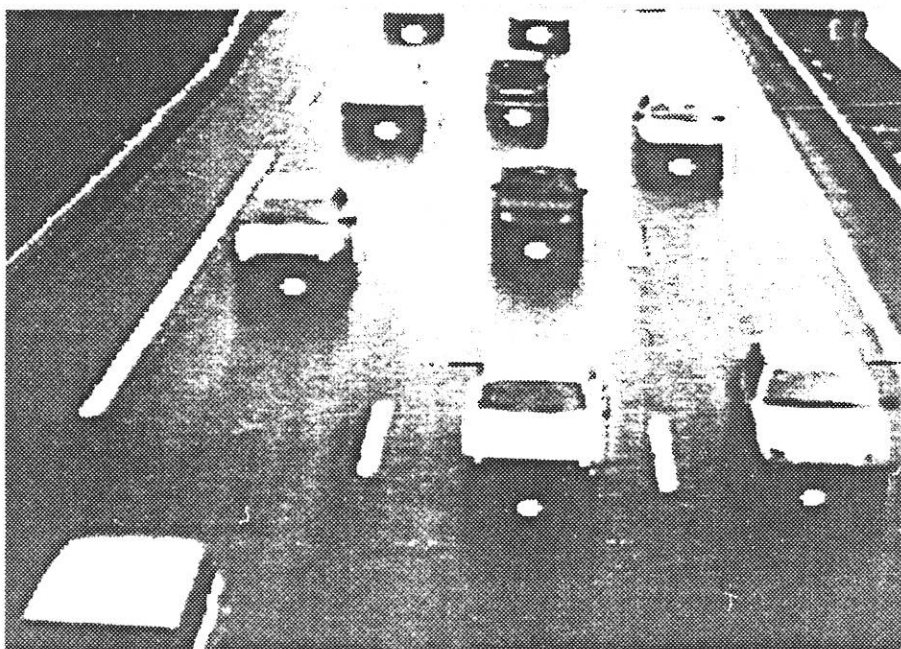
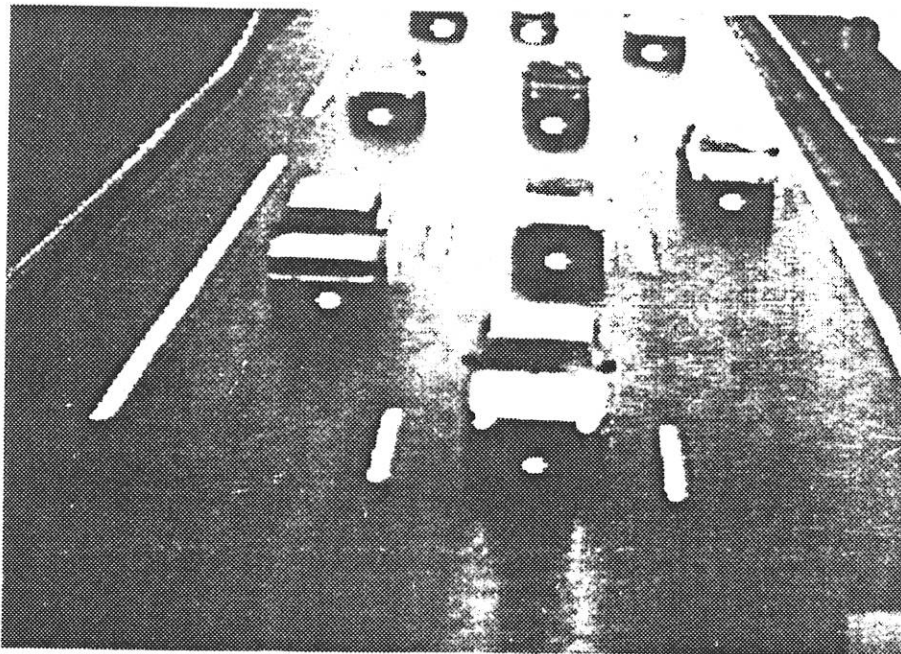


Planche III.2

Autoroute A6, champ de 55 m environ, caméra située à 7,5 m de hauteur, densité moyenne, images prises à une seconde d'intervalle.

Cette série illustre une situation où l'effet de la perspective se fait particulièrement sentir, et où les transformées géodésiques sont performantes.

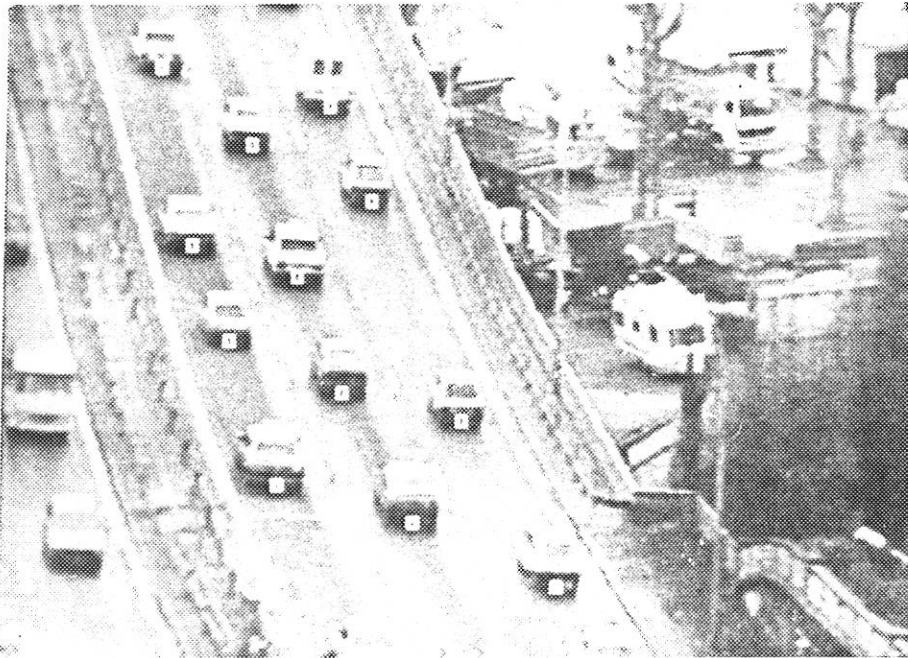
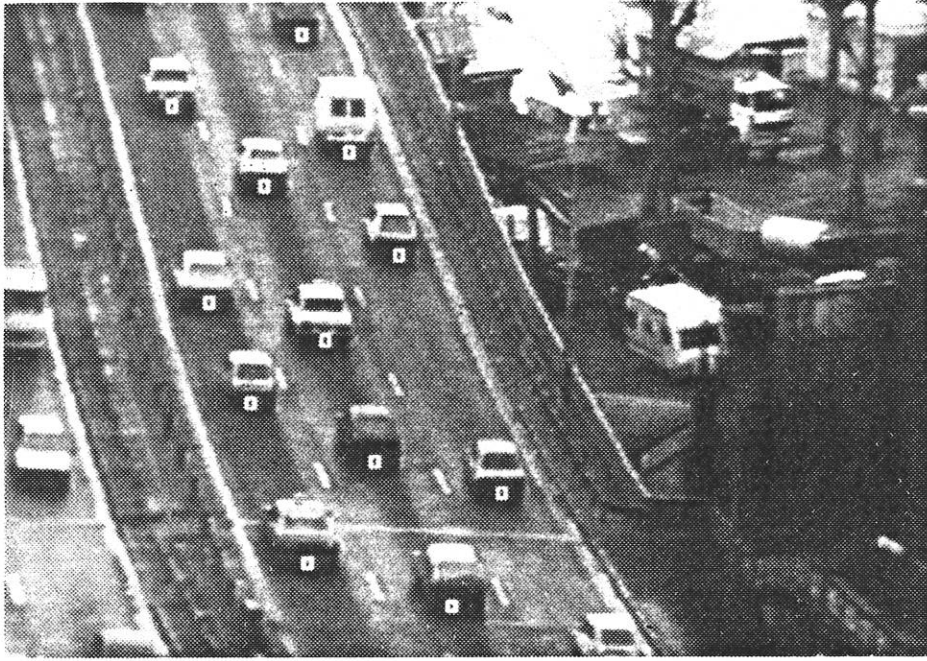


Planche III.3

Vue arrière, Boulevard Périphérique de Paris, champs d'environ 110 m, caméra à 30 m, densité moyenne (proportion de voie occupée par les véhicules $\approx 20\%$).

Le capteur est aussi performant pour les vues arrière que pour les vues de face. Plus la caméra est élevée, meilleurs sont les résultats même pour des densités assez élevées.

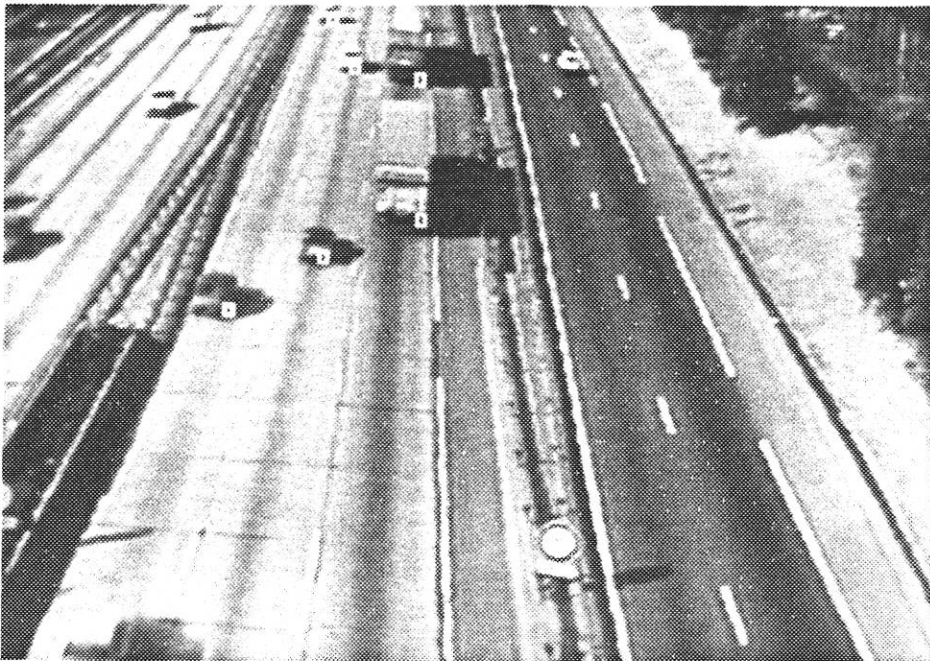
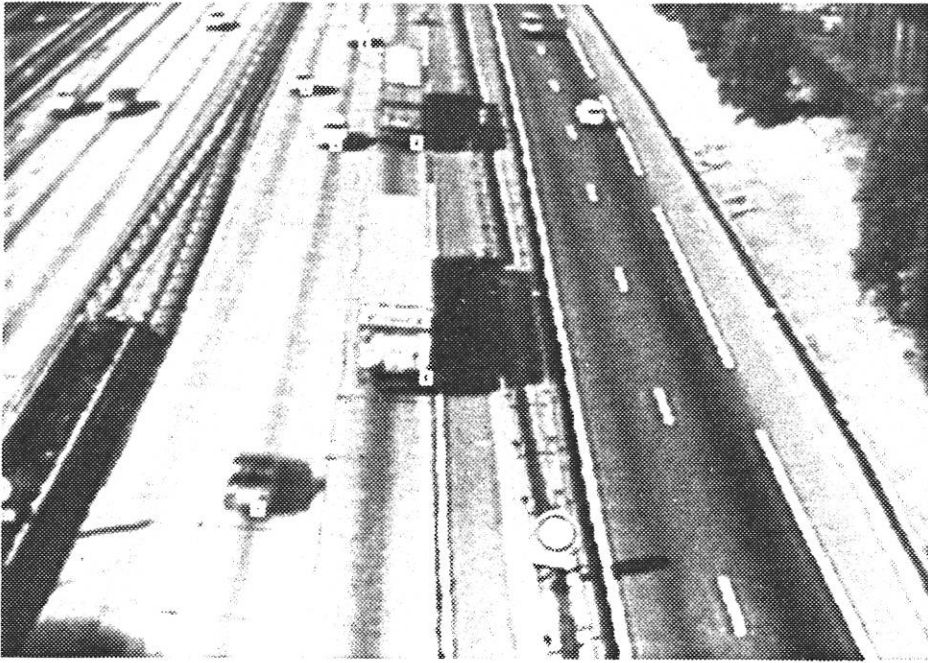


Planche III.4

Champs de 120 m, détection de véhicules avec ombre portée latérale. A noter que les poids lourds sont aussi bien détectés. La caméra est située à environ 15 m de haut par rapport à la chaussée.

L'éclairage rasant déporte les marqueurs. Ce décalage peut cependant être corrigé lors de la représentation des trajectoires.



Planche III.5

Champ de 70 m. Repérage de véhicules en conditions difficiles.
Camion à semi-remorque apparemment connexe au camion de chantier le suivant.

Cet exemple illustre bien comment la prise en compte d'une série d'images et la représentation des trajectoires peut compenser des erreurs de détection.



Planche III.6

Détection de véhicules de formes variées (camions ouverts) et en partie masqués (une fourgonnette derrière un poids lourd). Là encore, les véhicules arrivant sur l'avant-scène ont tendance à mieux se séparer.

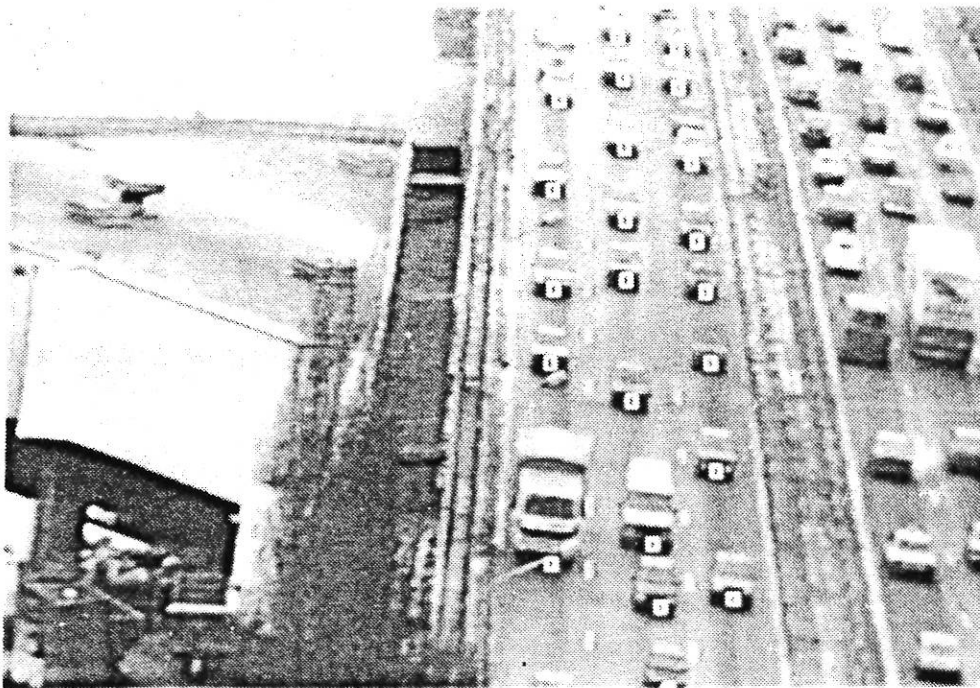
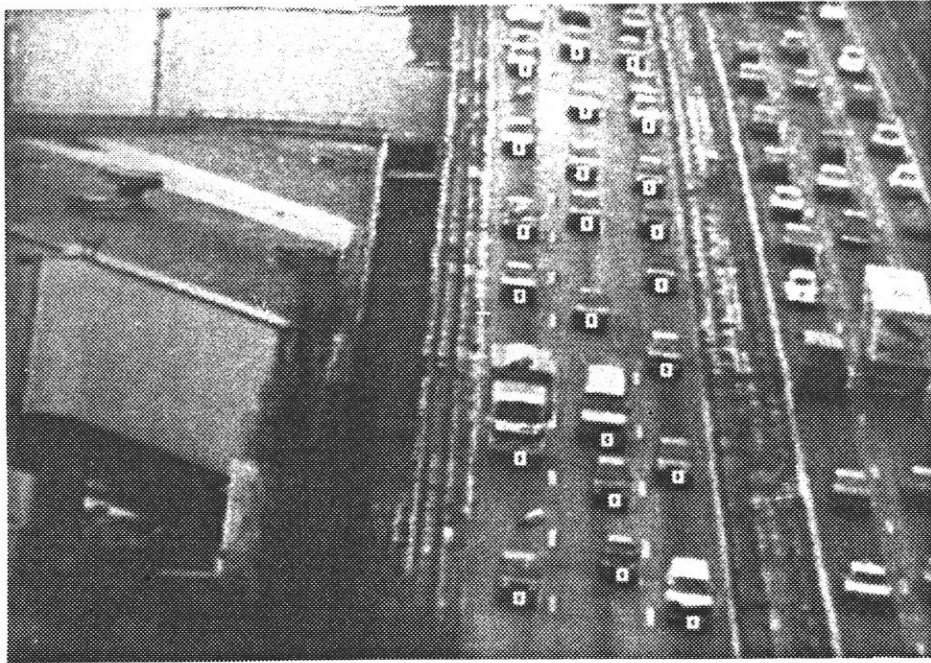


Planche III.7

Repérage de véhicules sur une étendue importante (≈ 220 m). La hauteur à laquelle est placée la caméra (30 m) permet une détection sur une grande distance même par densité assez forte (20 - 25%).

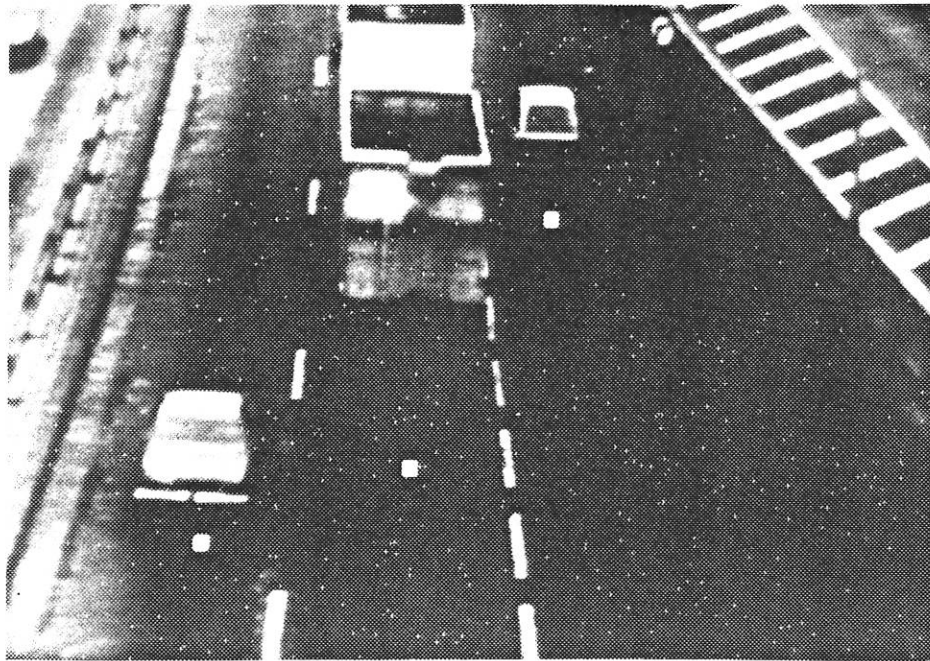
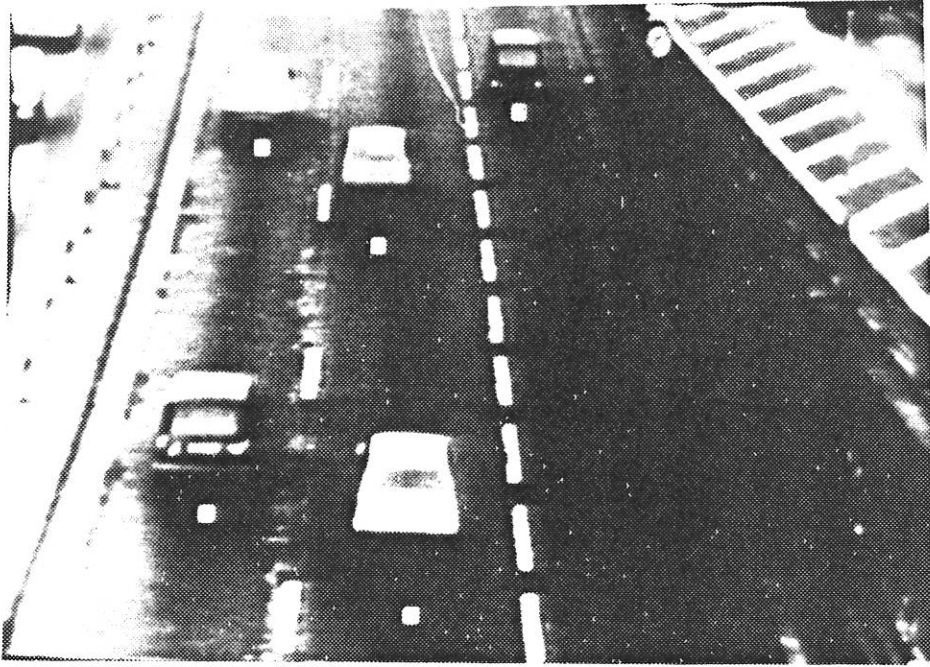


Planche III.8

Utilisation d'une caméra déjà installée (partie sud du B.P.). La faiblesse du champ embrassé est la conséquence de la faible hauteur des caméras (6 m). A noter le faible contraste des images et le mauvais centrage de la dynamique des niveaux de gris, du à l'ancienneté du tube vidéo et à la dégradation du signal pendant son transport jusqu'au poste de contrôle.

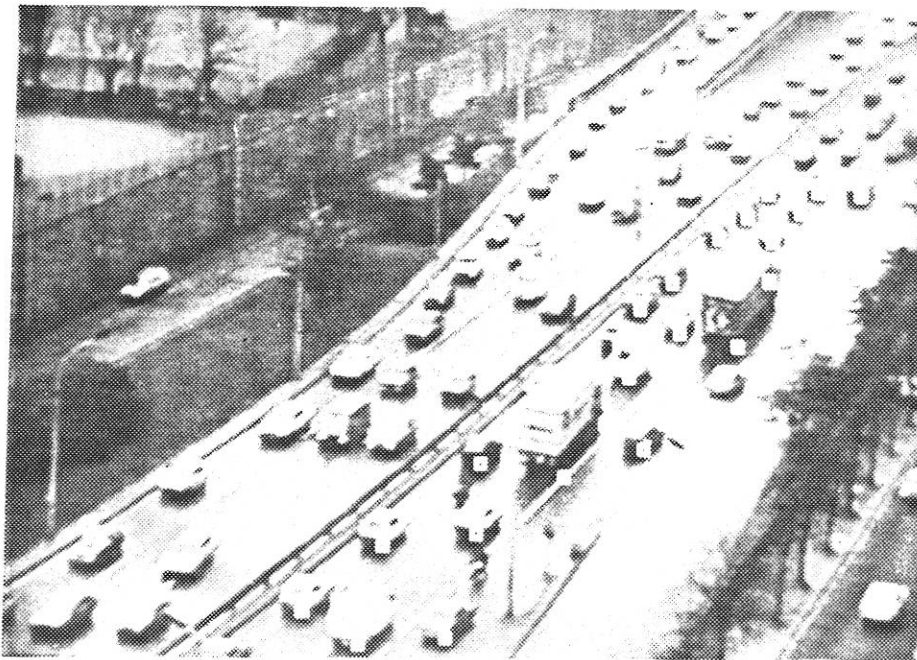
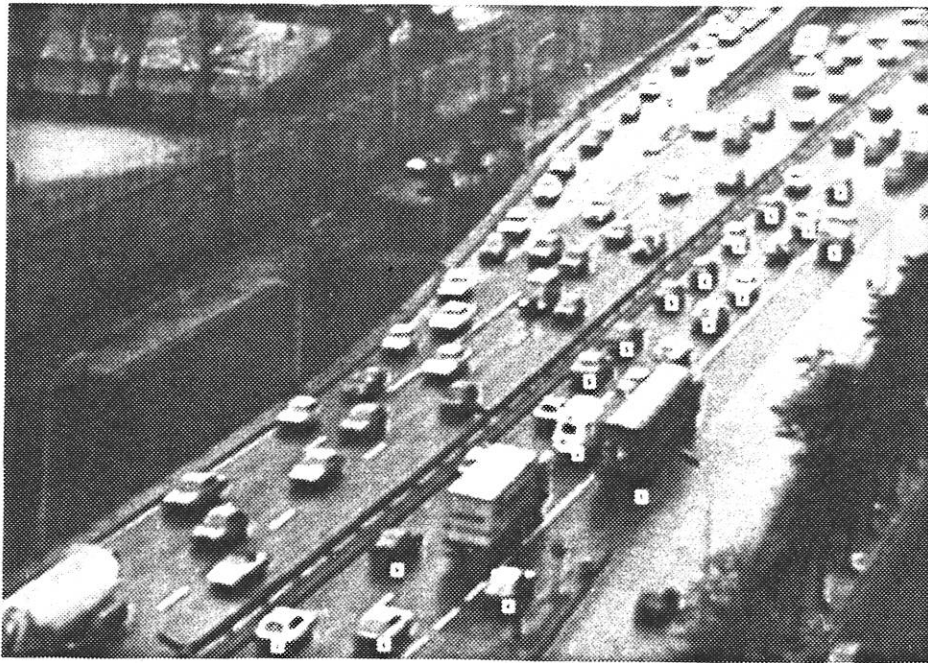


Planche III.9

Boulevard Périphérique, Porte Brancion, mauvaises conditions atmosphériques (pluie, neige).

Tentative de détection sur des images de fortes concentrations (>30 %) et de latéralité élevée. Malgré les bons résultats, on perçoit certaines limites de la méthode, dues à des occultations de véhicules légers par des poids lourds et à des amalgames de véhicules rendus inséparables du fait de la forte densité.

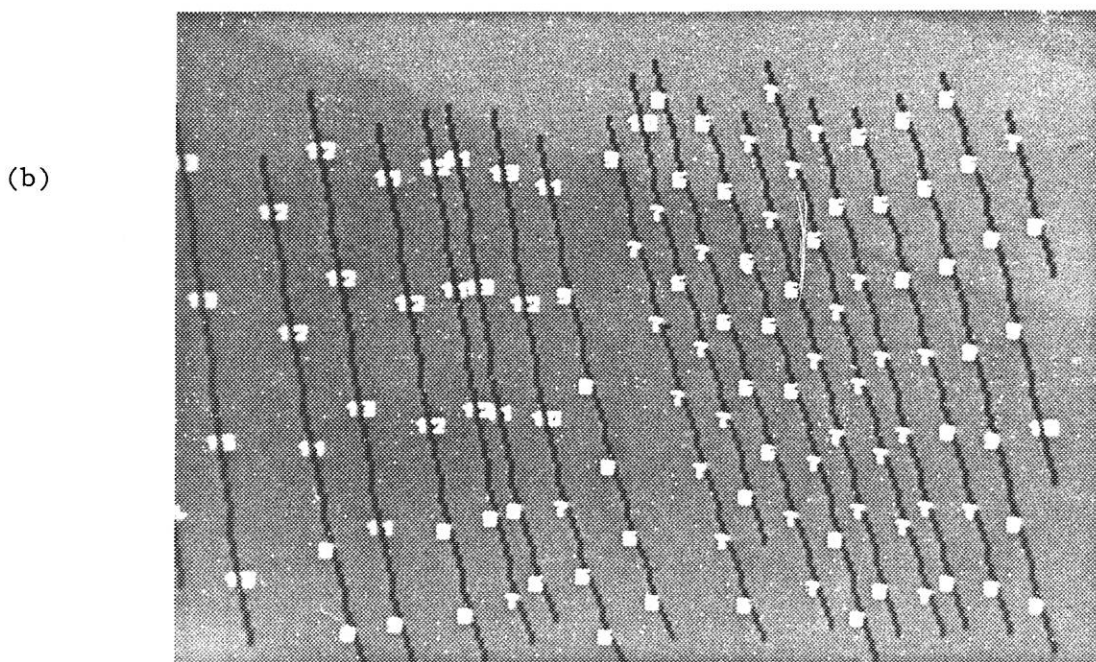
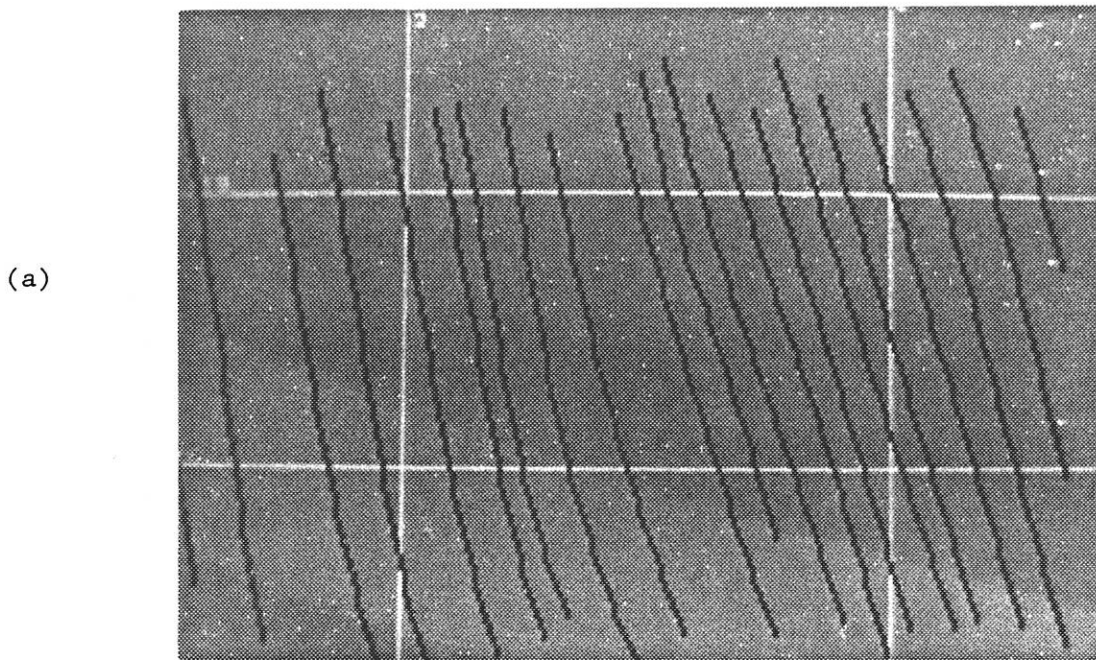


Planche III.10

Exemple de représentation dans l'espace-temps des trajectoires des véhicules. A noter que toute fausse détection a disparu après cette étape.

Figure a : Les intersections des trajectoires avec les lignes horizontales et verticales donnent respectivement les débits aux abscisses choisies et les concentrations aux instants choisis.

Figure b : Cette figure montre , outre les trajectoires des véhicules, les vitesses calculées entre chaque instant de détection. On peut ainsi, en une seule représentation, apprécier le déroulement d'un phénomène (ici une onde de choc aval) et appréhender les variations de toutes les variables impliquées (vitesse, concentration, débit).

III - 2) PERFORMANCES DE CE TYPE DE CAPTEUR

La banque d'images-tests comprend une dizaine de sites. Chaque site comprend une séquence d'une centaine d'images. Un millier d'images a donc été analysé à l'heure actuelle. Cet échantillonnage relativement important, compte tenu du temps de collecte des scènes et de leur temps de traitement, fournit une estimation assez précise des performances du procédé et de l'efficacité de chacune des phases de traitement. Ainsi le pourcentage de fausses détections (véhicules non pris en compte, double marquage, artefacts, etc...) est inférieur à 3% après la deuxième phase du traitement. Ces erreurs de détection sont dues pour une bonne part à l'effet perturbateur des poids lourds qui soit sont mal détectés du fait de leur taille (double marquage), soit cachent des véhicules légers qui les suivent.

Cependant, ce taux de fausses détections tombe à une valeur proche de 0 après la troisième phase d'analyse des trajectoires. La redondance de l'information intervient ici pleinement pour éliminer des erreurs inévitables lorsqu'un seul instant de détection est pris en compte.

Rappelons cependant que ces bons résultats ne s'appliquent pour le moment qu'aux scènes autoroutières, diurnes et dans lesquelles le trafic est fluide à moyennement dense.

CONCLUSIONS ET PERSPECTIVES D'AVENIR

Les résultats obtenus jusqu'à présent sont prometteurs et démontrent une certaine robustesse des procédés de détection. Pour les conforter et améliorer les algorithmes une expérimentation en situation est nécessaire. Mais avant cette expérimentation, de nombreuses tâches restent à entreprendre. On peut cependant dès maintenant et avant d'envisager les perspectives d'avenir, dégager les points forts et les faiblesses ou imprécisions de la méthode. L'amélioration du processus en éliminant ses faiblesses fait d'ailleurs partie des préoccupations majeures de la recherche menée actuellement.

IV - 1) CRITIQUES DES DIFFERENTES ETAPES DU PROCEDE

La critique des différents algorithmes sera effectuée dans l'ordre où ils apparaissent dans la description du procédé.

IV-1-1) L'initialisation

C'est une étape indispensable si la caméra doit se déplacer mais elle est de toute manière nécessaire si l'on veut pouvoir s'adapter facilement à un site quelconque.

La mise en évidence des voies de la chaussée renforce la fiabilité des résultats en permettant dans l'étape finale du traitement la prise en compte des véhicules voie par voie, et en évitant ainsi le chaînage de marqueurs appartenant à des voies différentes.

De plus, l'initialisation automatique permet de traiter toutes sortes de voies, droites, courbes (faible rayon), mais aussi les accès d'autoroute, à condition que la caméra demeure approximativement dans l'axe de la chaussée.

La détection des voies est d'autant plus facile que la densité de circulation est forte. La qualité du résultat (masque des voies) est liée au

nombre de véhicules ayant traversé la scène pendant la période d'initialisation. Il faudra donc sur un capteur temps réel, pour assurer une bonne initialisation des voies, adapter la durée d'initialisation à la densité du trafic.

Le traitement en général est rendu plus difficile si la prise de vue devient trop latérale (cf Planche III.9). La séparation des voies peut même devenir impossible si on ne distingue pas plus sur les images l'espace entre les voies adjacentes.

IV-1-2) La localisation des véhicules

Elle est calée, on l'a vu, sur les ombres portées ou les parties de véhicules proches du sol. Cette localisation n'est donc pas affectée par la hauteur des véhicules.

Les différentes incidences de l'éclairement et donc les différentes positions des ombres peuvent entraîner des erreurs de positionnement des véhicules. Cependant, ces erreurs peuvent être compensées lorsque les conditions d'éclairement restent stables. Elles peuvent néanmoins être gênantes lors de brusques changements de luminosité (passage de nuages). Ces brusques variations n'affectent pas réellement la position des ombres, mais modifie surtout leur contraste.

La précision des mesures de l'encombrement au sol (occupation spatiale), est liée au calcul correct des facteurs d'échelle (passage des distances mesurées sur l'image aux distances au sol). Elle dépend aussi de la qualité de l'extraction de toutes les parties des véhicules. Or l'intervention de la hauteur et sa projection dans le plan image, entraîne une erreur systématique dans le calcul de la longueur des véhicules. Il convient donc de l'estimer pour pouvoir corriger la mesure faite. Cette correction passe nécessairement par la reconnaissance de la catégorie des véhicules (VL, PL).

Néanmoins, avant toute évaluation, on peut affirmer que la mesure sera d'autant meilleure que la hauteur interviendra peu. Pour minimiser l'intervention de la troisième dimension, deux conditions sont nécessaires: une caméra d'une grande élévation, et une incidence acceptable par rapport à la verticale.

IV-1-3) La représentation spatio-temporelle

Cette représentation par son caractère synthétique est intéressante pour l'étude microscopique des événements. On voit s'y former les congestions, s'y illustrer les lois de poursuite. En élargissant l'espace sur laquelle elle s'applique cette représentation pourrait servir de synoptique des événements se déroulant sur une chaussée.

IV - 2) PERSPECTIVES : TRAITEMENT EN TEMPS REEL

Les résultats actuels sont obtenus par émulation du capteur sur une machine type IBM PC/AT. Le traitement complet d'une seule image, hormis la phase de tracé des trajectoires, se fait en environ 5 minutes. Pour suivre correctement la progression des véhicules, dans toutes les conditions possibles, un traitement tous les quarts de seconde est nécessaire. Un capteur temps réel devrait donc être 1000 fois plus rapide que l'émulation actuelle qui en est faite. Cette machine, réalisant une opération morphologique élémentaire sur une image 256x256x6 bits en 1,7 ms, est en cours de réalisation. Les premiers essais de la maquette sont prévus pour le début de l'année 1988.

De nombreuses situations de trafic restent encore à analyser. Pour n'en citer que deux, rappelons le cas particulièrement complexe des situations de trafic dense et le trafic nocturne. La mise au point d'algorithmes adaptés à ces situations constitue, avec la réalisation du processeur, l'actuel travail de recherche à court terme ([19]).

A plus long terme, un travail de supervision et d'intégration des différents algorithmes développés sera nécessaire si l'on veut disposer d'un capteur utilisable pour toutes les situations rencontrées. Cependant, cette nécessité d'utiliser à bon escient divers algorithmes ne remet pas fondamentalement en cause l'architecture du processeur temps réel. Celle-ci a été conçue de façon suffisamment souple pour supporter de nouveaux développements logiciels.

RAPPELS DE MORPHOLOGIE MATHÉMATIQUE

Cette annexe regroupe, pêle-mêle, les définitions et/ou les algorithmes des transformations morphologiques les plus souvent utilisés tout au long de l'étude entreprise. Nous rappellerons d'abord les définitions ensemblistes correspondant à la morphologie dite binaire, puis nous donnerons les définitions équivalentes dans le cadre de la morphologie dite numérique ou à niveaux de gris.

MORPHOLOGIE BINAIRE

Erosion, dilatation

Soit X un ensemble $\in \mathbb{R}^2$ et B un autre ensemble muni d'une origine et appelé élément structurant (B peut être simplement un disque ou un segment. On le supposera symétrique par rapport à son origine pour simplifier le formalisme). L'érodé de X par B noté $X \ominus B$ est défini par :

$$X \ominus B = \{x \in \mathbb{R}^2 ; B_x \subset X\}$$

C'est donc l'ensemble des points de l'espace pour lesquels B centré sur l'un d'eux est inclus dans X (Fig. A.1).



Figure A.1 : Erosion d'un ensemble

On définit de la même façon la dilatation, notée $X \oplus B$:

$$X \oplus B = \{x \in \mathbb{R}^2 ; X \cap B_x \neq \emptyset\}$$

Dans ce cas, tout point de l'espace tel que B centré sur lui intersecte X appartient au dilaté (Fig. A.2).

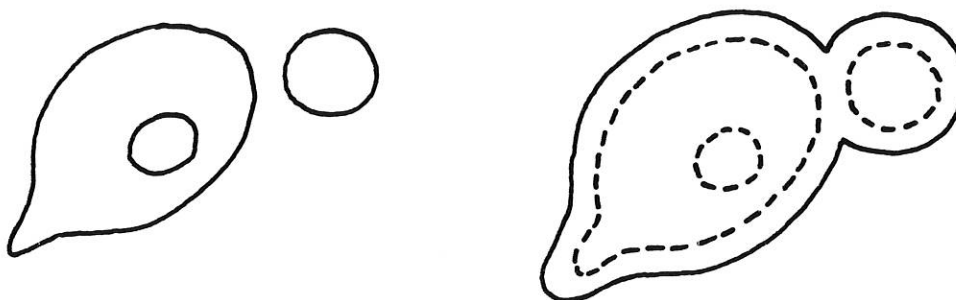


Figure A.2 : Dilatation d'un ensemble .

Ouverture, fermeture

L'ouverture d'un ensemble X par B, notée $(X)_B$, est définie par :

$$(X)_B = (X \ominus B) \oplus B$$

C'est une érosion suivie d'une dilatation.

La fermeture quant à elle, notée $(X)^B$ est égale à :

$$(X)^B = (X \oplus B) \ominus B$$

Ces transformées possèdent de bonnes propriétés de filtre (au sens restreint de tamis) qui expliquent leur usage fréquent tant en débruitage que pour simplifier les formes (Fig. A.3).

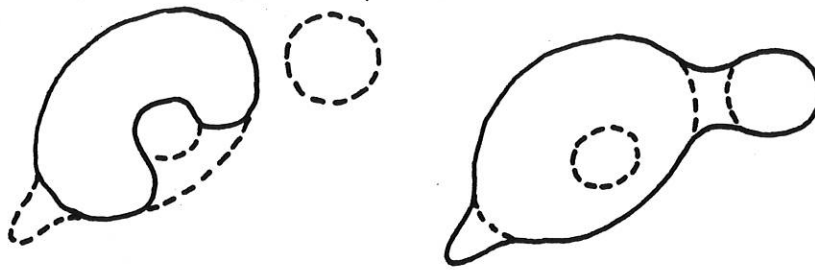


Figure A.3 : Ouverture et fermeture

Erosion, dilatation géodésiques

L'érosion et la dilatation géodésiques sont des transformations particulières :

- Elles sont définies dans un ensemble Z appelé ensemble géodésique à partir d'une distance particulière définie sur cet ensemble.
- L'élément structurant est une boule de rayon r .
- Leur définition est la suivante :

$$E_z(X,r) = \{x \in Z ; d_z(x, X^c) \geq r\}$$

$$D_z(X,r) = \{x \in Z ; d_z(x, X) \leq r\}$$

d_z est la distance géodésique définie dans Z . La distance géodésique entre deux points x et y de Z est égale à :

$$d_z(x,y) = L(x,y)$$

où L est la longueur du chemin minimal inclus dans Z joignant les points x et y (Fig. A.4).

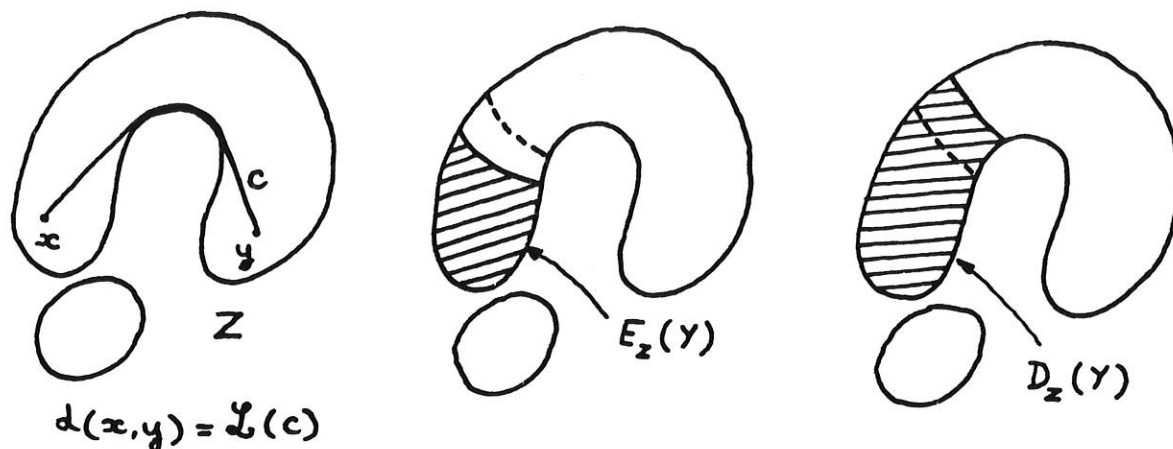


Figure A.4 : Distance géodésique, érosion, dilatation géodésiques

Squelette, Squelette par zones d'influence

Le squelette d'un ensemble X est défini comme le lieu des centres des disques maximaux inclus dans X .

Un disque est maximal lorsqu'il n'est pas lui-même inclus dans un disque inclus dans X (Fig. A.5).

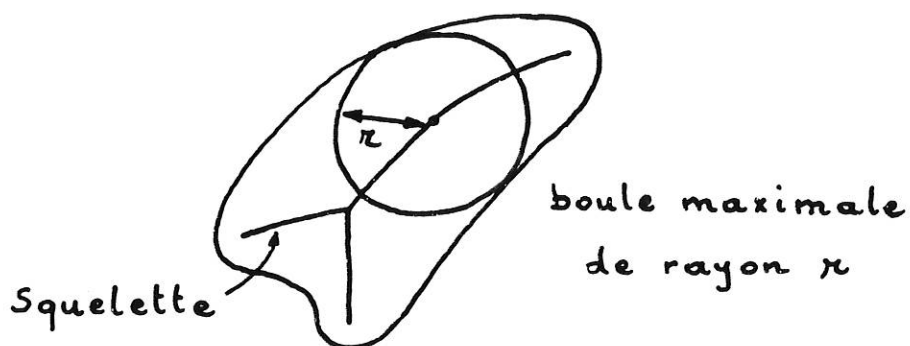


Figure A.5 : Boules maximales et squelette

Le squelette par zones d'influence (ou SKIZ) d'un ensemble X formé de plusieurs composantes connexes X_i est l'ensemble des points de X^o également distants d'au moins deux composantes connexes distinctes (Fig. A.6).

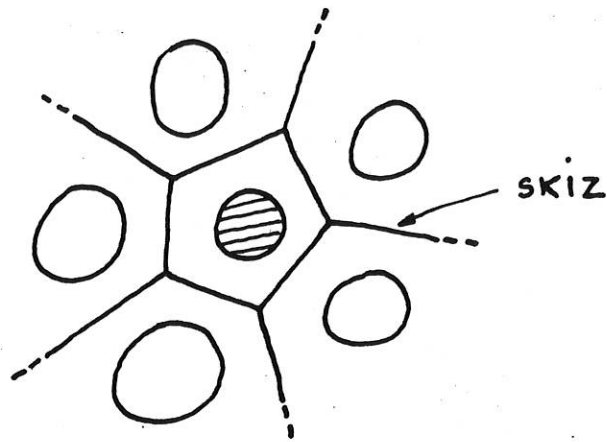


Figure A.6 : Squelette par zones d'influence

Ce squelette partitionne l'espace en polygones, chacun d'eux affecté à une composante connexe X_i et constituant sa zone d'influence.

MORPHOLOGIE NUMERIQUE

En morphologie mathématique numérique, les objets mathématiques utilisés ne sont plus des ensembles mais des fonctions (représentant généralement les valeurs de gris d'une image). Dans ce cas, on peut encore définir des opérations élémentaires comme l'érosion et la dilatation ainsi que des transformées plus complexes.

Erosion, dilatation, ouverture, fermeture

L'érosion d'une fonction f pour un élément structurant B est une fonction g égale à :

$$g(x) = f \ominus B(x) = \text{Inf}_{y \in B_x} f(y)$$

Quant à la dilatation, sa définition est :

$$g(x) = f \oplus B(x) = \text{Sup}_{y \in B_x} f(y)$$

L'ouverture $(f)_B$ et la fermeture $(f)^B$ s'obtiennent comme en binaire, en

combinant ces deux opérations élémentaires. La figure A.7 illustre ces transformations dans le cas monodimensionnel.

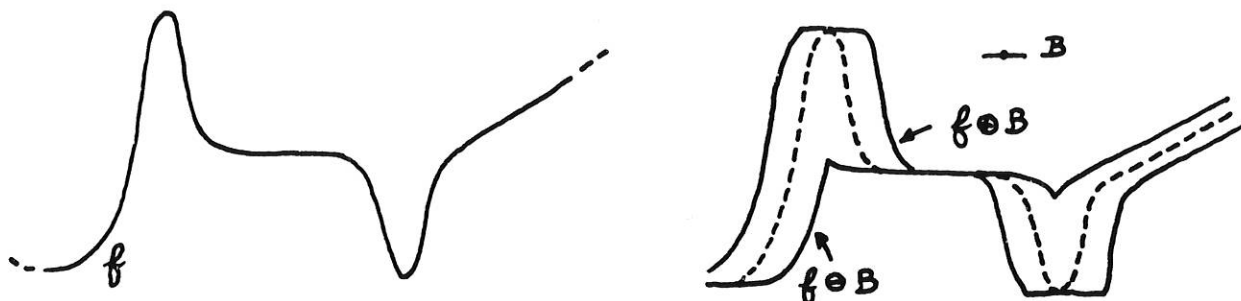


Figure A.7 : Erosion et dilatation numériques

Chapeau haut-de-forme

On appelle chapeau haut-de-forme (ou Top Hat) une transformation numérique consistant à soustraire d'une fonction f , son ouverture par un élément structurant B (Fig. A.8).

$$WTH(f) = f - (f)_B$$

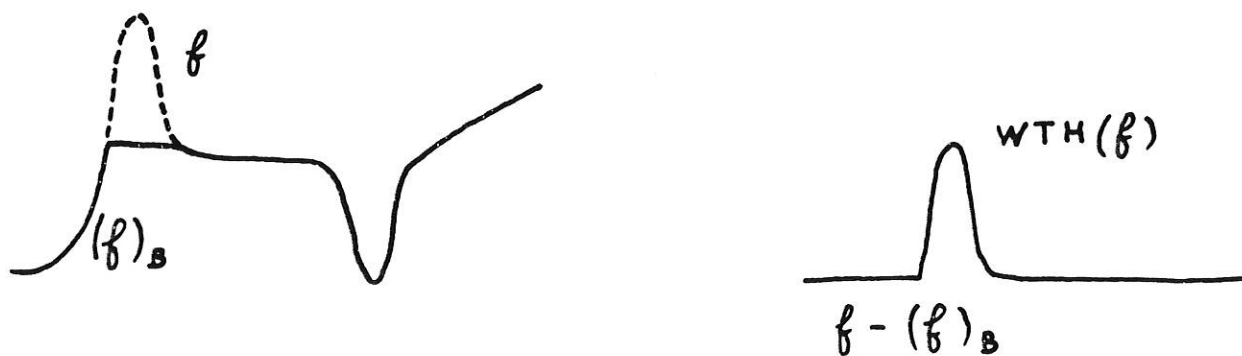


Figure A.8 : Transformée chapeau haut-de-forme

Cette transformation met en évidence dans l'image initiale les détails suffisamment fins et plus clairs que leur environnement. Une transformation similaire mais utilisant la fermeture peut également être définie pour extraire les objets sombres :

$$\text{BTH}(f) = (f)^B - f$$

Minima régionaux d'une fonction

On peut donner des minima régionaux d'une fonction une définition intuitive (mais cependant parfaitement exacte). Supposons que le graphe de la fonction représente une surface topographique et imaginons que nous nous déplaçons sur cette surface topographique. Si, ayant atteint un point x , à l'altitude $f(x)$, tout déplacement vers un point à une altitude inférieure quel qu'il soit, ne peut se faire en utilisant un chemin descendant, x appartient à un minimum régional (Fig. A.9).

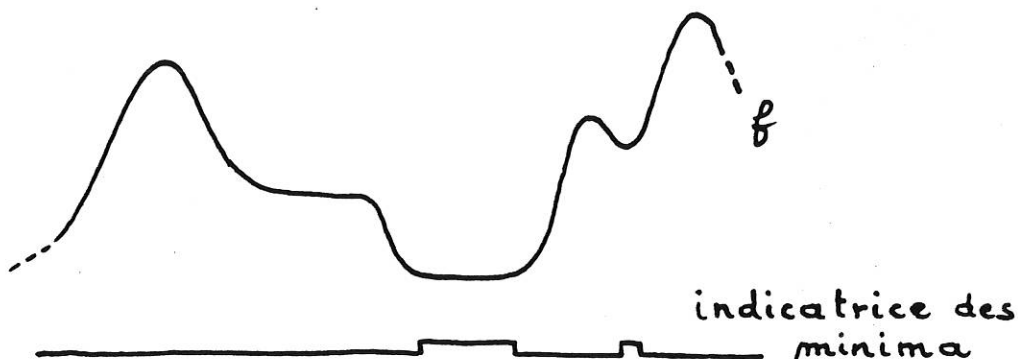


Figure A.9 : Minima régionaux d'une fonction

CALCUL DES PARAMETRES DE DIMENSIONNEMENT AU SOL

Ce calcul a pour but de connaître la valeur des facteurs d'échelle dans le sens horizontal et vertical, en fonction des coordonnées des points-image.

On supposera la caméra orientée parallèlement à la route (même si cette orientation est grossière).

La figure B.1 représente le dispositif de prise de vues.

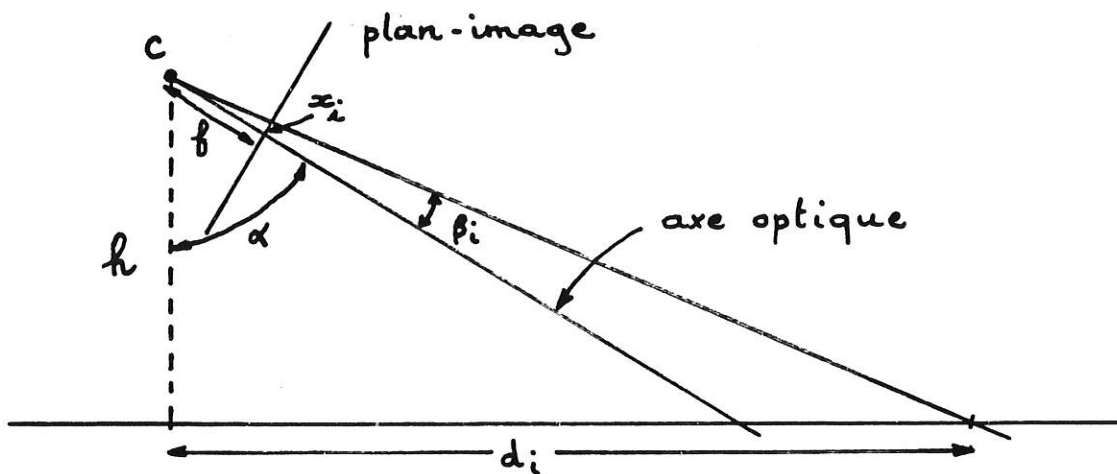


Figure B.1 : Dispositif de prise de vues

C est la caméra

h, son altitude par rapport à la chaussée

α , son orientation

f, est la distance focale de la caméra

Notons que les deux derniers paramètres sont généralement inconnus sur un site.

Si d_i est la distance d'un point de la chaussée au pied de la caméra, on peut écrire :

$$d_i = h \operatorname{tg}(\alpha + \beta_i)$$

(on supposera $\emptyset \leq \alpha + \beta_i < \Pi/2$, $\forall \beta_i$).

$$d_i = \frac{h(\operatorname{tg} \alpha + \operatorname{tg} \beta_i)}{1 - \operatorname{tg} \alpha \operatorname{tg} \beta_i}$$

comme $\operatorname{tg} \beta_i = \frac{x_i}{f}$, x_i coordonnée du point sur l'image, on obtient (Fig. B.2) :

$$d_i = \frac{h(\operatorname{ftg} \alpha + x_i)}{f - x_i \operatorname{tg} \alpha} \quad (1)$$

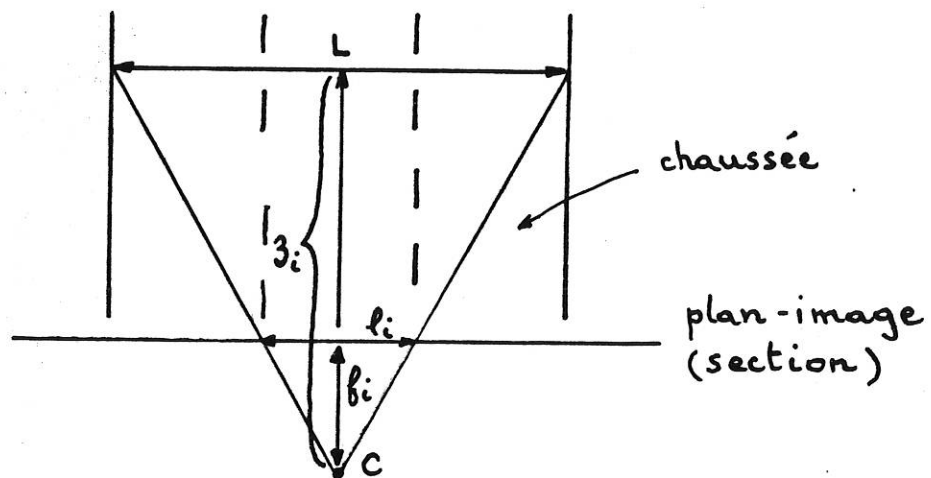


Figure B.2 : Paramètres frontaux du dispositif

L est la largeur de la chaussée

l_i est la largeur apparente sur l'image à la cote x_i

On a :

$$\frac{z_i}{f_i} = \frac{L}{l_i}$$

$$z_i = f_i \frac{L}{l_i} = \frac{fL}{l_i \cos \beta_i}$$

comme : $d_i^2 + h^2 = z_i^2$, on a :

$$d_i^2 + h^2 = \frac{f^2 L^2}{l_i^2 \cos^2 \beta_i} \quad (2)$$

On peut éliminer f entre les équations (1) et (2). On obtient :

$$\frac{x_i^2 (1 + \operatorname{tg}^2 \alpha)}{(d_i - h \operatorname{tg} \alpha)^2} = \frac{l_i^2}{L^2}$$

comme x_i et $(d_i - h \operatorname{tg} \alpha)$ varient dans le même sens, on peut écrire :

$$\frac{x_i}{\cos \alpha (d_i - h \operatorname{tg} \alpha)} = \frac{l_i}{L}$$

soit :

$$d_i = \frac{x_i h \sin \alpha + L x_i}{l_i \cos \alpha}$$

En posant :

$$d_i^2 + h^2 = \frac{L^2 f^2}{l_i^2 \cos^2 \alpha} (1 + \operatorname{tg}^2 \beta_i)$$

on obtient :

$$l_i = \frac{L}{h} [f \cos \alpha - x_i \sin \alpha]$$

Il y a donc une relation linéaire entre x_i et l_i (on s'en doutait d'après les images !), que l'on peut écrire :

$$l_i = A - B x_i$$

Les paramètres A et B peuvent être calculés directement sur l'image masque en déterminant les longueurs d'intercepts de la route pour différentes cotes x_i .

Le facteur d'échelle horizontal moyen est donné par :

$$e_h^i = \frac{l_i}{L} = \frac{(A - bx_i)}{L} \quad (\text{nombre de pixels/m})$$

Le facteur d'échelle vertical moyen est égal à

$$e_v^i = \frac{\partial x_i}{\partial d_i}$$

comme :

$$d_i = htg\alpha + \frac{Lx_i}{(A - Bx_i) \cos \alpha}$$

$$e_v^i = \frac{\cos \alpha (A - Bx_i)^2}{LA}$$

L'incidence α de la caméra n'étant pas connue, on peut éliminer ce paramètre si on connaît au sol un intervalle D entre deux points et les coordonnées x_0 et x_1 qui lui correspondent sur l'image.

On a alors :

$$D = \frac{LA |x_1 - x_0|}{\cos \alpha (A - Bx_1)(A - Bx_0)}$$

soit :

$$e_v^i = \frac{(A - Bx_1)^2}{D} \frac{|x_1 - x_0|}{(A - Bx_1)(A - Bx_0)}$$

On peut aussi plus simplement faire intervenir la hauteur de la caméra, si elle est connue (la formule correspondante est donnée au chapitre II).

IMPLANTATION DES TRANSFORMEES GEODESIQUES MULTI-MASQUES

Cette annexe décrit les algorithmes d'implantation des opérations morphologiques élémentaires (érosion, dilatation), lorsque leur taille est exprimée en mètres au sol.

Soit l , la taille de la transformée à effectuer. Il lui correspond en chaque point x de l'image une taille exprimée en pixels égale à $e(x)l$, où $e(x)$ est le facteur d'échelle.

Le champ d'analyse est découpé en différents masques, chacun d'eux correspond à une zone de l'image où la taille en pixels est approximativement constante et entière. D'un champ à l'autre, la taille augmente d'un pixel (Fig. C.1).

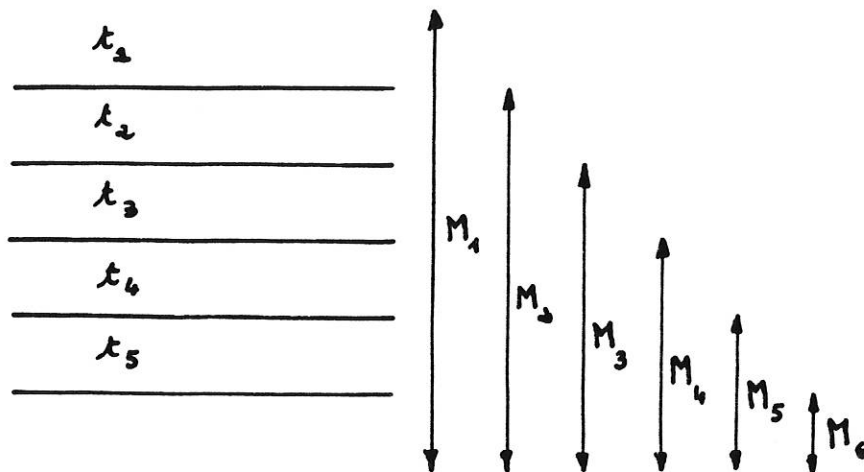


Figure C.1 : Masques géodésiques

On considère en fait les masques cumulés M_1, M_2, \dots . La séquence $\{M_i\}$ est monotone décroissante, et les tailles en pixels les plus faibles correspondent à l'arrière-plan. Soit t_i , la taille minimale correspondant au masque M_i : pour tout point du masque, la taille exprimée en pixels correspondant à la taille réelle l est au moins égale à t_i .

Effectuer une transformation $Y = T_\ell(X)$ revient donc à programmer la séquence suivante :

- entrer la valeur ℓ
- calculer et positionner les masques M_i
- si la transformée est extensive effectuer la séquence A. Effectuer la séquence B si l'opération est anti-extensive (érosion).

Séquence A :

- effectuer la transformée géodésique de la restriction de X dans le masque M_1 , la taille étant t_1 . Le résultat est noté $T_{t_1}^{\circ}(X)$,
- passer au masque suivant M_2 . Il lui correspond une taille minimum $t_2 = t_1 + 1$. Il suffit alors d'effectuer la transformée géodésique de taille 1 de la restriction de $T_{t_1}^{\circ}(X)$ dans le masque M_2 . Cette opération est notée $T_1^{\circ} \circ T_{t_1}^{\circ}(X)$.
- recommencer la dernière opération sur tous les masques M_i de l'image.

Séquence B :

La séquence B est similaire. Elle est seulement inversée, en ce sens que l'on commence par une transformée de taille 1 sur le dernier masque situé à l'avant scène. On incrémente ensuite de 1 la taille de l'opération, pour terminer par une opération de taille t_1 sur l'ensemble de l'image.

Cette façon de faire évite des problèmes d'effet de bord qui pourraient apparaître lorsque l'on combine ces deux opérations élémentaires.

Les figures C.2 et C.3 illustrent ces opérations.

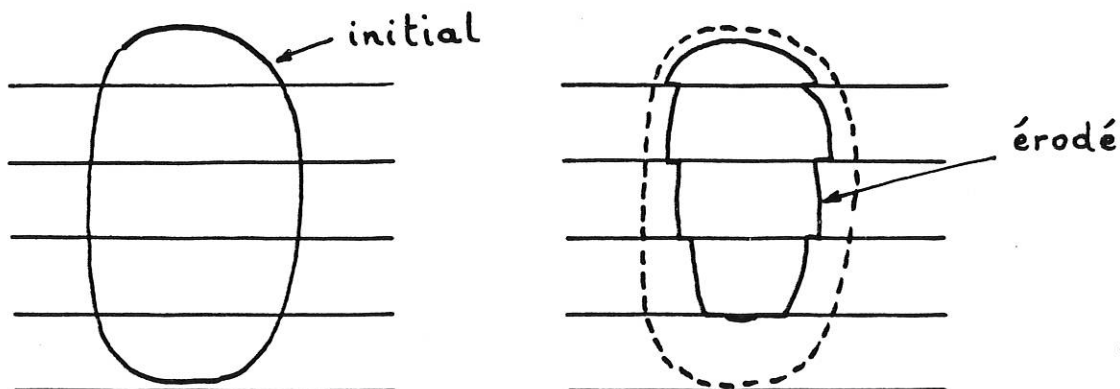


Figure C.2 : Erosion géodésique multi-masques

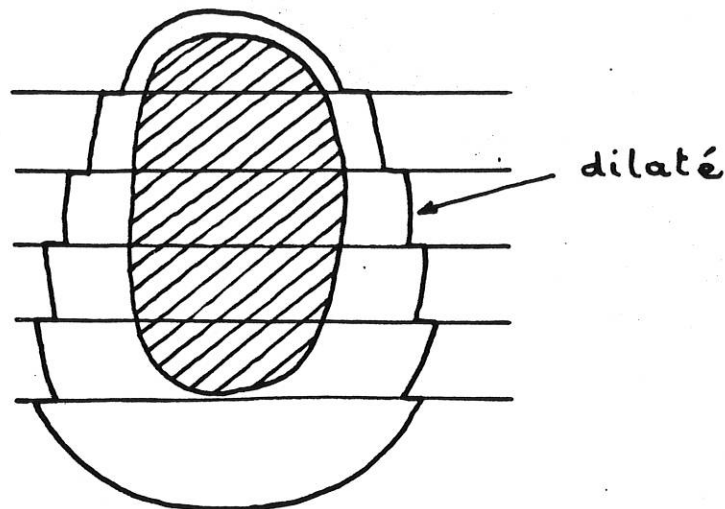


Figure C.3 : Dilatation géodésique multi-masques

Deux facteurs d'échelle ont été calculés, e_h dans le sens horizontal et e_v dans le sens vertical. Il est bien sûr évident que dans le cas des transformées linéaires horizontales, le facteur d'échelle à utiliser est e_h .

Pour les transformées bidimensionnelles et en particulier celles qui utilisent des éléments structurants carrés ou octogonaux, la rigueur voudrait que ces transformations soient découpées en une succession de transformées linéaires, chacune d'entre elles utilisant le facteur d'échelle approprié.

Cette procédure est en pratique longue et fastidieuse. Elle multiplie en particulier le nombre de masques à utiliser. Or, pour le problème qui nous intéresse, étant donné que le paramètre déterminant est l'épaisseur apparente des traits caractéristiques des véhicules, on utilisera simplement le facteur d'échelle e_v (en particulier pour les transformées chapeau haut-de-forme).

PRESENTATION SUCCINCTE DU LOGICIEL ALADIN

Le logiciel ALADIN développé sur systèmes IBM PC a été conçu pour répondre aux besoins suivants de l'utilisateur:

- Fournir un moyen d'initiation à l'analyse d'images et à la morphologie mathématique. Ce logiciel permet notamment l'exécution de toutes les opérations de base mises en oeuvre par cette méthode.

- Permettre l'écriture d'algorithmes complexes d'analyse d'images, ainsi que la gestion de ces algorithmes (stockage, modification, destruction). L'exécution de ces algorithmes doit évidemment être possible de façon interactive.

- Autoriser par l'exploitation du logiciel une évaluation de la complexité d'un algorithme, complexité exprimée en termes de nombre d'opérations élémentaires. Cette évaluation doit permettre de cerner les performances d'un éventuel processeur câblé.

DESCRIPTION DU LOGICIEL

Le logiciel ALADIN est composé de deux parties:

- Une bibliothèque d'opérations de base.
- Un environnement de développement qui fournit principalement à l'utilisateur les services suivants:

- Un langage de description d'algorithmes.
- Un éditeur interactif pour l'écriture de ces algorithmes.
- Un gestionnaire de la bibliothèque de traitements.

La bibliothèque de base

Cette bibliothèque comporte plusieurs types d'opérateurs. On distingue:

- Les opérateurs liés à l'interface de numérisation/visualisation.
- Les opérateurs d'entrées/sorties.
- Les opérateurs de traitement d'images. Ces opérateurs sont soit des transformées arithmétiques ou booléennes entre deux images, soit des

opérateurs morphologiques de base: transformées en tout ou rien, érosion, dilatation, amincissement, épaissement, etc...

L'environnement de développement

Le logiciel de développement étant trop long à décrire ici, on se contentera d'en donner un exemple commenté d'utilisation.

L'utilisateur veut réaliser un squelette par amincissement. Pour réaliser un tel squelette, il faut effectuer un amincissement dans les six directions de la trame hexagonale. Pour ce faire, on définit un élément structurant adéquat et on le fait tourner après chaque transformée élémentaire. On itère ensuite le processus jusqu'à stabilisation.

L'algorithme est le suivant:

```
REPETER
AMINCISSEMENT DANS LES SIX DIRECTIONS
JUSQU'A IDEMPOTENCE
```

avec pour l'amincissement dans les six directions:

```
REPETER
AMINCISSEMENT ELEMENTAIRE
ROTATION DE L'ELEMENT STRUCTURANT
JUSQU'A AVOIR TRAITE SIX DIRECTIONS
```

L'écriture de la procédure est la suivante:

```
Transf-im(i1,i2)      (* transfert d'image *)
REPETE                (* pour le test d'idempotence *)
  Transf-im(i2,i1)    (* nécessaire pour l'itération *)
  AmincRot(i2,i3)     (* résultat dans i2 *)
JUSQ-Im(i1,i2)
Fin
```

La transformée AmincRot est définie par:

```
Affect(n3,#6)        (* affecte la valeur 6 à la variable n3 *)
REPETE
```

```
Aminc-Im(i1,i2)      (* utilise l'élément structurant courant *)
RotStruct            (* rotation de l'élément *)
Transf-Im(i2,i1)    (* sauvegarde pour l'itération *)
Soust-N(n3,#1)      (* n3 = n3 - 1 *)
JUSQ-N(n3,#0)       (* boucle effectuée six fois *)
Fin
```

CONDITIONS D'UTILISATION

Ce logiciel fonctionne sur toute machine IBM PC ou compatible (l'installation sur un AT est préférable pour des raisons de vitesse d'exécution).

Les images traitées sont au format 256x256x6 bits.

La visualisation des images peut se faire sous deux formes:

- Avec une carte spécialisée 256x256x6 bits, visualisation en gris ou en fausses couleurs.

- En utilisant la visualisation standard du micro-ordinateur. La résolution et le nombre de couleurs ou de niveaux de gris dépendent du type de visualisation adopté:

Carte HERCULES: Images binaires.

Carte EGA: Images 16 fausses couleurs.

BIBLIOGRAPHIE

- [1] ABRAMCZUK T.A. : TV detection for monitoring traffic flow, final report on feasibility studies. Royal Institute of Technology, Stockholm, TRITA-ILA, Janvier 1982.
- [2] BEUCHER S., BLOSSEVILLE J.M., LENOIR F. : Traffic spatial measurements using video image processing, Application of mathematical morphology to vehicles detection. Congrès SPIE, Cambridge, Mass., Novembre 1987.
- [3] BEUCHER S., LANTUEJOUL C. : On the use of geodesic metric in image analysis. Journal of Microscopy, Vol 121, part 1, Janvier 1981.
- [4] BEUCHER S., SERRA J. : Shapes and patterns of microstructures considered as grey-tone functions. Stereol. Jugosl. 1981, proc. of the 3rd European symp. on stereology, Ljubljana, Juin 1981.
- [5] BLOSSEVILLE J.M. : Traitement d'image appliqué à la mesure du trafic. Journée spécialisée INRETS, Arcueil, Septembre 1986.
- [6] BLOSSEVILLE J.M., LENOIR F., BEUCHER S. : Mesure spatiale de trafic par analyse d'image. XXIIème congrès FISITA, Dearborn, USA, Septembre 1988 (à paraître).
- [7] COSTER M., CHERMANT J.L. : Précis d'analyse d'images. Editions du CNRS, 1985.
- [8] DICKINSON K.W., ASWORTH R. : Traffic data collection and analysis using video processing. T.E.C., Novembre 1983.
- [9] DODS J.S. : The Australian Road Research Board video based vehicle presence detector. IEE Conference publication no 242, Londres, Décembre 1984.
- [10] HOUKES A. : Measurement of speed and time headway of motor vehicles with video camera and computer. Proc. Int. Conf. on digital computer application to process control, Dusseldörf, Octobre 1980.
- [11] MAES W., BARBE A., VITS A. : Development of a detection system based on the use of CCD camera. Seminar on micro. electr., OCDE, Tokyo, Novembre 1984.
- [12] ONOE M., NOBUO M., OHBA K. : Computer analysis of traffic flow observed by subtractive television. Computer Graphics and Image Processing, pages 377-392, 1973.

- [13] ONOE M., OHBA K. : Digital image analysis of traffic flow. Proc. Int. joint conf. on pattern recognition, Coronada, Novembre 1976.
- [14] POSTAIRE J.G., STELMASZYK P., BONNET P. : A visual surveillance system for traffic collision avoidance control. IFAC/Transportation symposium, Vienne, 1986.
- [15] SCHULTSMEYER A.P. : Wide area detection system (WADS). Report no FHWA/RD 82/144, Federal Highway Administration, Washington DC, Septembre 1982.
- [16] SERRA J. : Image analysis and mathematical morphology. Academic Press, London, 1982.
- [17] TAKABA S. : Measurement of traffic flow using real time processing of moving pictures. Proc. 32nd IEE Conf. on vehicle technology, San Diego, Mai 1982.
- [18] TAKABA S., SEKINE T., HNANG B.W. : A traffic flow measuring system using a solid-state image sensor. IEE Conference publication no 242, Londres, Décembre 1984.
- [19] VU H. : Analyse par caméra vidéo d'un trafic dense. Rapport d'option, Ecole des Mines de Paris, Juin 1987.