

INSTITUT NATIONAL POLYTECHNIQUE
DE LORRAINE

Jbl

Laboratoire d'Electronique
d'Electrotechnique et d'Automatique
E. N. S. E. M.
2, Rue de la Citadelle E.N.S.E.M.
B. P. 850 - Tél. 52.68.32
54011 NANCY CEDEX

THESE DE DOCTORAT I.N.P.L.

Section : Genie Electrique
Option : Traitement d'image

présentée par

Odile HORN

sujet :

ETUDE ET MISE EN ŒUVRE D'UN ALGORITHME
DE POURSUITE DE CIBLE PAR
ANALYSE D'IMAGE

Soutenue publiquement le 29 juin 1989 devant la commission d'examen .

Rapporteurs

J.P. HATON
P.L. WENDEL

Membres du jury

E. YVROUD (*Président*)
J. CICCOTELLI
P. SIENER
R. HUSSON



D 136 019314 6

136019 3146

UD 1989 HORN, G.

INSTITUT NATIONAL POLYTECHNIQUE
DE LORRAINE

Ensemble d'Electrotechnique et d'Automatique
E. N. S. E. M.
2, Rue de la Citadelle
B. P. 850 - Tél. 52.68.32
54011 NANCY CEDEX

THESE DE DOCTORAT I.N.P.L.

*Section : Genie Electrique
Option : Traitement d'image*

présentée par

Odile HORN

sujet :

**ETUDE ET MISE EN ŒUVRE D'UN ALGORITHME
DE POURSUITE DE CIBLE PAR
ANALYSE D'IMAGE**



Soutenue publiquement le 29 juin 1989 devant la commission d'examen .

Rapporteurs J.P. HATON
 P.L. WENDEL

Membres du jury E. YVROUD (*Président*)
 J. CICCOTELLI
 P. SIENER
 R. HUSSON

Avant propos

Le travail présenté dans cette thèse a été effectué au Centre de Recherche en Automatique de Nancy (CNRS - UA 821) au sein de l'équipe Vision - Robotique de l'E.N.S.E.M.⁽¹⁾.

Que Monsieur R.HUSSON, Professeur à l'E.N.S.E.M. et Directeur du C.R.A.N.⁽²⁾ soit assuré de ma profonde gratitude pour son accueil et pour l'aide qu'il m'a apporté tout au long de cette étude.

Je remercie vivement Monsieur E. YVROUD, Directeur de Recherche au C.N.R.S. pour l'honneur qu'il me fait en acceptant de présider ce jury.

J'exprime ma sincère reconnaissance à Monsieur J.P. HATON, Professeur d'Informatique à l'Université de Nancy I, à Monsieur P.L. WENDEL, Professeur d'Electronique à l'E.N.S.P.S.⁽³⁾, à Monsieur J. CICCOTELLI, Ingénieur de Recherche en Traitement d'Image au C.T.B.A.⁽⁴⁾, et à Monsieur P. SIENER, Ingénieur de Recherche à l'I.R.S.I.D.⁽⁵⁾, qui ont accepté de s'intéresser à ce travail et de participer au jury.

Que tous mes collègues du laboratoire et de l'E.N.S.E.M. et, en particulier, Madame F. SIMONOT, et Messieurs D. WOLF, M. DUFAUT, J.M. MUNIER, B. MARCHAL, J.P. THOMESSE et P. LETERRIER soient ici vivement remerciés pour leur accueil chaleureux et leur contribution à ce travail.

Enfin, que Mademoiselle C. FLEURY qui a assuré la réalisation de ce mémoire trouve ici l'expression de ma gratitude pour sa patience et son efficacité.

(1) E.N.S.E.M. : Ecole Nationale Supérieure d'Electricité et de Mécanique

(2) C.R.A.N. : Centre de Recherche en Automatique de Nancy

(3) E.N.S.P.S. : Ecole Nationale Supérieure de Physique de Strasbourg

(4) C.T.B.A. : Centre Technique du Bois et de l'Ameublement

(5) I.R.S.I.D. : Institut de Recherche de la Sidérurgie

TABLE DES MATIERES

	Page
INTRODUCTION	
I - Présentation de l'analyse d'image appliquée au mouvement	6
I - 1 - Méthodes par analyse globale des images successives	8
I - 1 - 1 - Analyse différentielle des séquences d'image	8
I - 1 - 2 - Soustraction d'images successives	13
I - 2 - Méthodes par mise en correspondance	17
I - 2 - 1 - Transposition dans un autre espace	17
I - 2 - 1 - 1 - Transformée de Fourier	18
I - 2 - 1 - 2 - Transformée de Hough	19
I - 2 - 1 - 3 - Transformation pour l'interprétation	21
I - 2 - 2 - Association modèle - image	22
I - 2 - 2 - 1 - La Corrélation	22
I - 2 - 2 - 2 - La recherche séquentielle	24
I - 2 - 2 - 3 - Les recherches orientées	31
I - 3 - Le contexte de l'analyse	36
I - 3 - 1 - Les prétraitements	37
I - 3 - 2 - Les architectures spécialisées	38
 II - Corrélation par restriction	 41
II - 1 - Situation du problème - présentation du système	41
II - 2 - Choix établis pour l'écartométrie	42
II - 2 - 1 - Principe	43
II - 2 - 2 - La caméra "Net - Flou"	43
II - 2 - 2 - 1 - Maximisation du pic de corrélation	43
II - 2 - 2 - 2 - Différents extracteurs de contours	47
II - 2 - 2 - 3 - Principe et réalisation du capteur "N-F"	48
II - 2 - 2 - 4 - Exemples d'images traitées - Effet sur la fonction de corrélation	50
II - 2 - 3 - Mise en oeuvre de la corrélation	54
II - 2 - 3 - 1 - Principe de l'algorithme	54
II - 2 - 3 - 2 - Problème de seuillage - critère de décision	59
II - 2 - 3 - 3 - La Pondération	63
II - 2 - 3 - 4 - Réalisation - Exemple	67

III - Application à la poursuite de cible : étude quantitative	74
III - 1 - Définition statistique des images	74
III - 1 - 1 - Modélisation des données issues du capteur "N-F"	74
III - 1 - 2 - Evaluation du nombre de points dépassant un niveau donné dans l'image	78
III - 1 - 2 - 1 - Aspect théorique	78
III - 1 - 2 - 2 - Résultats expérimentaux	81
III - 1 - 2 - 3 - Conclusions	88
III - 2 - Estimation des temps de traitement	90
III - 2 - 1 - Influence du nombre de points traités sur le temps d'exécution	90
III - 2 - 2 - Application aux choix des seuils	92
III - 2 - 3 - Recherche de la taille optimale de la fenêtre	95
III - 2 - 4 - Modèle de corrélation	99
 IV - Mise en oeuvre de la poursuite - filtrage de Kalman	 101
IV - 1 - Dispositif expérimental	101
IV - 1 - 1 - Description des différents éléments	101
IV - 1 - 2 - Premiers essais de poursuite	103
IV - 2 - Filtrage de Kalman	105
IV - 2 - 1 - Notations générales	105
IV - 2 - 2 - Application à la poursuite de cible	107
IV - 2 - 3 - Simulation de la poursuite	110
IV - 2 - 4 - Mise en oeuvre sur le dispositif - Résultats	113
 CONCLUSION	 125
 BIBLIOGRAPHIE	 126

INTRODUCTION

Cette thèse a été préparée dans le cadre de l'équipe vision - robotique du CRAN (Centre de Recherche en Automatique de Nancy) et financée par une BDI du CNRS. Ce mémoire traite de l'analyse de scène dynamique et de son application au suivi d'objet mobile. L'étude développée est fondée sur la corrélation monodimensionnelle entre images successives, elle mène à la réalisation d'un système expérimental de poursuite de cible, validant les choix établis pour les différentes phases de l'analyse. Les applications envisagées concernent à priori la robotique industrielle.

Le système développé se compose d'un capteur optoélectronique délivrant, en temps réel, les contours de la scène, d'un algorithme de corrélation rapide donnant la position de l'objet pour chaque acquisition d'image, et d'une procédure de filtrage calculant la trajectoire du mobile.

Dans le premier chapitre, nous faisons état des différentes méthodes employées pour l'analyse des mouvements par traitement d'image, et des champs d'exploitation de ces techniques.

Le second chapitre illustre les principes retenus pour le capteur et la corrélation. Nous montrons tout d'abord le caractère optimal des représentations haute fréquence des images pour les fonctions de corrélation, puis nous exposons la technique optoélectronique choisie pour extraire instantanément des contours isotropes. Ensuite, nous établissons un algorithme de corrélation binaire multiniveaux pour localiser un modèle dans une image. Enfin des exemples de scènes, de leurs contours, et des fonctions de corrélation calculées sont donnés.

Dans le chapitre 3, le problème de l'évaluation des temps d'exécution nécessaires à chaque étape est abordé. La modélisation des images issues du capteur est effectuée en utilisant des études menées en traitement du signal, puis le modèle choisi est testé. Les résultats permettent l'établissement d'une relation liant les niveaux de seuils choisis pour la corrélation aux temps de traitement requis pour cette dernière.

Le chapitre 4 décrit le dispositif expérimental, et la procédure de filtrage adaptatif mise en oeuvre, en aval de la corrélation, pour l'interprétation des résultats. Cette dernière opération permet de faire une prédiction à chaque acquisition concernant la position prochaine du mobile, elle augmente ainsi la vitesse limite envisageable pour celui-ci. De plus, ce filtre autorise certaines carances de perception, par exemple le masquage du mobile durant plusieurs étapes, sans altérer la poursuite. Des essais sont réalisés à partir d'une cible animée de différents mouvements connus et programmés, ils mettent en évidence l'effet du niveau de filtrage sur les résultats délivrés par le système. Cette étude débouche sur une procédure de réglage des paramètres du système en fonction des conditions d'observation et du type de trajectoire envisagée.

CHAPITRE I

I - PRESENTATION DE L'ANALYSE D'IMAGE APPLIQUEE AU MOUVEMENT

Durant les 15 dernières années, le traitement d'image par ordinateur a connu un essor notable grâce aux progrès technologiques en matière de composants électroniques. Ce domaine nécessite l'exploitation de nombreuses mémoires et le traitement de multiples instructions ; ces deux fonctions ne sont devenues envisageables qu'à partir du moment où de nouveaux circuits sont apparus sur le marché.

Pour la vision artificielle, l'image numérisée est considérée comme un ensemble de points : les pixels dont la valeur est le niveau de gris de l'image réelle à l'emplacement considéré. Les traitements développés dans ce domaine peuvent être regroupés en trois phases principales [ROS - 82] :

- L'acquisition (numérisation - compression)

Le signal analogique représentant l'intensité lumineuse est discrétisé et éventuellement codé, pour son stockage en mémoire.

- La transformation (réhaussement, restauration, reconstruction) :

L'image numérique est rendue plus exploitable par filtrage du bruit, reconstruction des zones altérées, ou estimation de données non perçues.

- L'interprétation (association, description, reconnaissance)

Les données sont segmentées et classées pour extraire les propriétés de la scène considérée, ou les relations entre différentes parties de celle-ci. Les résultats décrivent des modèles définissant des classes d'image.

A partir de ces trois types d'opération, de multiples domaines d'application ont pu être couverts. Citons la reconnaissance des formes, la vision en trois dimensions, l'étude de texture, l'interprétation d'image obtenue par satellite et l'évaluation des mouvements.

L'analyse de scène dynamique s'est récemment développée. Les travaux de recherche, s'y rapportant, reposent sur différentes méthodes et le sujet est si nouveau qu'aucune théorie générale ne peut prétendre unifier les approches existantes. Les méthodes employées dépendent de la nature des évolutions attendues entre les images et du domaine d'application prévu.

Il peut être :

- l'interprétation d'image aérienne en météorologie,
- le suivi de cellule biologique en médecine,
- la localisation d'un mobile en balistique,
- l'analyse du trafic routier,
- la commande de véhicule autonome,
- le codage intertrame en TV.

La perception du mouvement diffère des autres informations déduites du traitement d'image par le fait que la caractéristique recherchée n'est pas inhérente aux mesures à l'instant considéré, mais dépend de l'histoire du système.

Cette dépendance nécessite l'exploitation de modèles mathématiques : pour le choix des trajectoires possibles et pour le mode d'interprétation des variations survenues dans la scène. Nous donnons succinctement les différentes stratégies employées dans ce domaine en faisant ressortir quelques lignes directrices actuelles de l'analyse du mouvement.

Notons que les nombreuses publications donnant une vue d'ensemble des techniques existantes : [LET - 83], [PAV - 86], [NAG - 86], [ZAV - 88], [BOU - 88] montrent par la diversité de leurs critères de classement l'hétérogénéité du domaine. Nous distinguerons pour notre part deux tendances :

- le traitement global des séquences d'image par comparaison pixel par pixel des niveaux de gris à chaque instant d'acquisition,
- l'analyse indépendante des trames successives pour l'extraction de formes élémentaires porteuses d'information, puis la mise en correspondance des données extraites.

A partir de ces deux axes, nous proposons une partition des différentes méthodes développées. Notre classement reposant sur la manière de concevoir le mouvement et non sur l'application envisagée, certains travaux peuvent être décrits dans plusieurs paragraphes ; par exemple une initialisation par soustraction d'images successives, et une poursuite par mise en correspondance.

I - 1 - Méthodes par analyse globale des images successives

L'hypothèse de base, de ces méthodes, est que les changements apparaissant entre les images successives sont dus aux seuls mouvements des éléments de la scène. Elles ne nécessitent donc aucune connaissance préalable des objets mobiles mais imposent des contraintes de qualité sur les signaux analysés ; tout bruit intempestif étant susceptible de fausser les résultats.

I - 1 - 1 - Analyse différentielle des séquences d'image

Les niveaux de gris de l'image sont considérés comme une fonction du temps et de l'espace : $I(x,y,t)$ dont les variations temporelles révèlent le mouvement.

Représentons un exemple monodimensionnel :

$$I_1(x) = I_1(x, t_1) \dots$$

$$I_2(x) = I_2(x, t_2) \dots$$

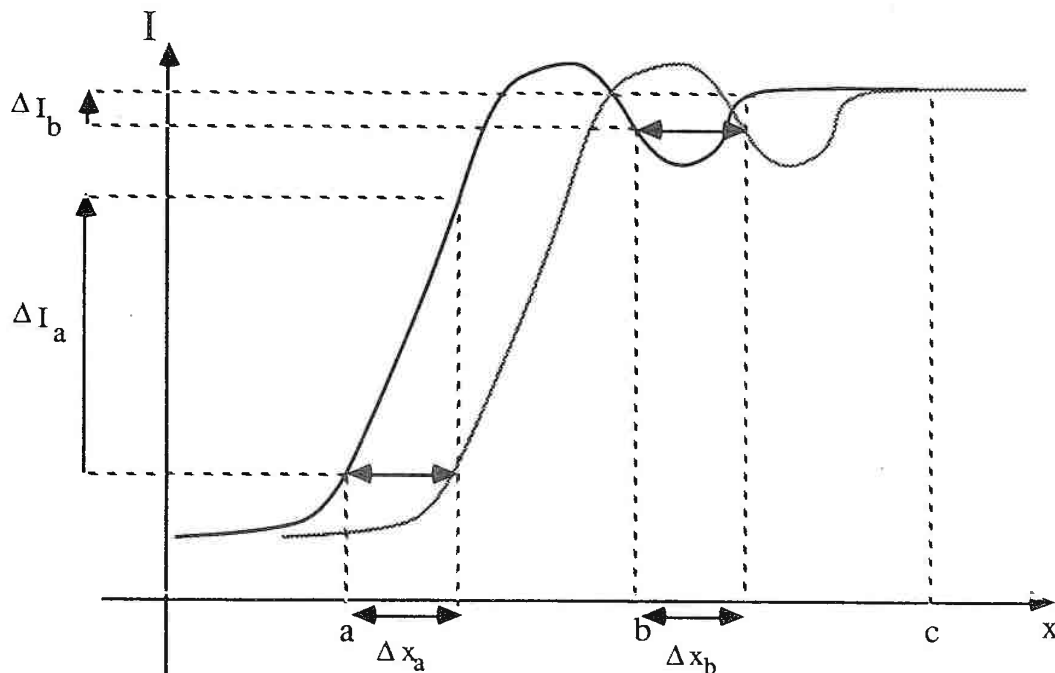


Figure 1 : Signaux image en deux instants successifs

La mesure des ΔI doit permettre d'évaluer le déplacement Δx entre les instants t_1 et t_2 .

$$[1.1] \quad \Delta I(x) = I_1(x) - I_2(x) = I(x, t) - I(x, t + \Delta t) \text{ avec } \Delta t = t_2 - t_1$$

Les approches exposées dans ce qui suit diffèrent par le choix du modèle considéré pour déduire Δx de $\Delta I(x)$.

Si I est une fonction spatialement dérivable dont on approxime le développement de Taylor au premier ordre [GUI - 84] :

$$[1.2] \quad \Delta I = \frac{\delta I(x)}{\delta x} dx$$

soit dans le cas bidimensionnel

$$[1.3] \quad \Delta I = \text{grad } I(x, y) \cdot D$$

où D est le vecteur déplacement entre deux instants.

L'équation [1.3] ne permet pas à elle seule de calculer D , cependant elle le limite à une droite affine. Ainsi, en connaissant ΔI en plusieurs points ayant subis le même mouvement (par exemple a et b sur le dessin), l'intersection des droites affines définies par [1.3] délivrera les composantes de D . Le résultat issu de ce calcul est obtenu au pixel près, une telle stratégie n'est donc intéressante que dans la mesure où les images traitées ont une bonne définition. Par ailleurs $\Delta I(a) \gg \Delta I(b)$, l'incertitude relative de la mesure en a sera donc intérieure à celle en b , il est donc intéressant de choisir, dans l'image, les points à fort gradient. Ce choix est conforté par une remarque sur la vision humaine : le déplacement d'un objet gris sur un fond gris sera indiscernable, par contre, la présence de contraste (donc de fort gradient) entre l'objet et le fond, par exemple l'un noir et l'autre blanc, donnera une impression immédiate de mouvement.

Pour une étude analytique de cette approche différentielle, les travaux cités dans [ZAV - 88] exploitent une modélisation mathématique de $I(x)$ quand il y a absence d'écart entre les niveaux de gris à deux instants d'observation, deux cas sont envisagés :

- l'image est à luminance constante au voisinage du point considéré (point c de la figure 1), le mouvement est ici indécélable,
- l'image comporte un extréma local ; l'approximation de la formule de Taylor au premier ordre engendre une perte d'information, il faut considérer les termes dérivés d'ordre supérieur pour évaluer le mouvement.

Plus concrètement, Rodrigues [ROD - 85] propose un modèle local pour la détection des mouvements perpendiculaires aux contours : l'unité de covariance temporelle.

Elle comprend deux unités exploitant les données images issues de deux points spatialement voisins, suivant le schéma ci-dessous.

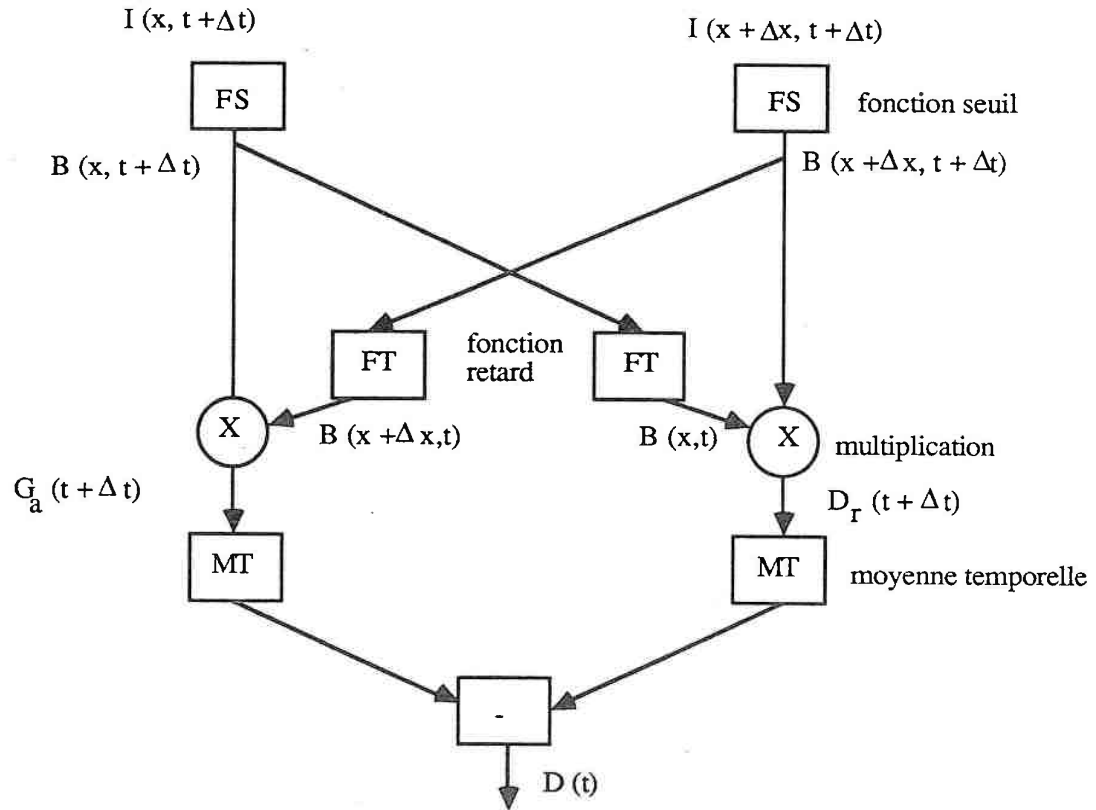


Figure 2 : Schéma de l'unité de covariance temporelle

Les signaux $I(x, t)$ sont binarisés par seuillage en $B(x, t)$. La détection du déplacement est fondée sur la comparaison de similarité pour les couples $(B(x, t + \Delta t), B(x + \Delta x, t))$ et $(B(x + \Delta x, t + \Delta t), B(x, t))$; représentant respectivement des translations à gauche et à droite. La différence $D(t)$ n'est remarquable que sur les contours de l'image (où $B(x) \neq B(x + \Delta x)$), et révèle la composante du déplacement qui leurs est orthogonale. Ainsi, les déplacements d'un objet peuvent être détectés en satisfaisant aux hypothèses de rigidité, continuité et translation.

Le cas d'éléments de formes variables est traité par ce type d'approche pour un codage intertrame en TV [CAR - 87] ; les contours sont divisés en petits segments indépendants affectés d'un vecteur vitesse calculé par la formule :

$$[1.4] \quad \frac{\delta I}{\delta t} + \frac{\delta I}{\delta x} v_x + \frac{\delta I}{\delta y} v_y - \frac{\delta I}{\delta x} \omega (y - y_0) + \frac{\delta I}{\delta y} \omega (x - x_0) = 0$$

avec

v_x, v_y : vitesse en x et y ($\frac{dx}{dt}$ et $\frac{dy}{dt}$)

(x_0, y_0) : centre de rotation pour le segment considéré

ω : vitesse de rotation ($\frac{d\theta}{dt}$)

En notant $\frac{\delta I}{\delta \theta} = \frac{\delta I}{\delta y}(x - x_0) - \frac{\delta I}{\delta x}(y - y_0)$.

Le critère des moindres carrés le long du contour s'écrit :

$$[1.5] \quad \sum_i \left(\frac{\delta I}{\delta x_i} v_x + \frac{\delta I}{\delta y_i} v_y + \frac{\delta I}{\delta \theta_i} \omega + \frac{\delta I}{\delta t_i} \right)^2$$

Cette mesure doit être minimisée car le vecteur déplacement ne varie pas brutalement entre deux segments. Ceci constitue une information complémentaire nécessaire au calcul du champ de vitesses le long du contour. L'étape suivante est la détermination de ce champ en tout point de l'image ; la minimisation d'un critère est à nouveau nécessaire, il s'agit d'une discrétisation de la technique développée dans [HOR - 81], la limitation du gradient spatial des vitesses est mesurée par

$$[1.6] \quad \iint_x \left(\left(\frac{\delta v_x}{\delta x} \right)^2 + \left(\frac{\delta v_x}{\delta y} \right)^2 + \left(\frac{\delta v_y}{\delta x} \right)^2 + \left(\frac{\delta v_y}{\delta y} \right)^2 \right) dx dy$$

dans le domaine discret [2.5] est transposée par

$$[1.7] \quad \sum_m \sum_n [(d_x(m,n) - d_x(m-1,n))^2 + (d_x(m,n) - d_x(m,n-1))^2 + (d_y(m,n) - d_y(m-1,n))^2 + (d_y(m,n) - d_y(m,n-1))^2]$$

où $d_x(m,n)$ et $d_y(m,n)$ sont les déplacements en x et en y du point de coordonnées (m,n) .

[1.7] est rendue minimum séparément selon les deux directions par une méthode de relaxation ; les itérations sont appliquées en tous points exceptés ceux des contours, et donnent la carte des vitesses de l'image.

Ce modèle de déplacement composé d'une translation plus une rotation est aussi exploité pour connaître la trajectoire d'objets rigides [DIE - 86].

Soient deux images successives :

$$[1.8] \quad I_1(x,y) = S(x,y)$$

$$[1.9] \quad I_2(x,y) = S(x \cos \varphi - y \sin \varphi - d_1, x \sin \varphi + y \cos \varphi - d_2).$$

$I_2(x,y)$ est issue de $I_1(x,y)$ par une rotation (φ) et une translation (d_1, d_2). Ces paramètres du mouvement sont ici recherchés en tous points par un algorithme itératif de type Newton - Raphson.

En notant l'estimé de I_2 :

$$[1.10] \quad I_m(x,y) = S(x \cos \hat{\varphi} - y \sin \hat{\varphi} - \hat{d}_1, x \sin \hat{\varphi} + y \cos \hat{\varphi} - \hat{d}_2)$$

cette technique dirige les variations de $\hat{\varphi}$, $\hat{d}_1$ et $\hat{d}_2$ jusqu'à rendre minimum le critère d'erreur.

$$[1.11] \quad E\{e^2\} = E\{(I_m(x,y) - I_2(x,y))^2\}$$

Une itération est exécutée en 320 ms et 4 à 5 itérations sont nécessaires pour déterminer d_1 , d_2 et φ .

Une approche voisine est développée dans [LEG - 82] où la minimisation de l'erreur d'estimation est dirigée par la méthode du gradient, pour un système de poursuite de cible simulé.

Citons enfin les stratégies "mixtes" exploitant le formalisme différentiel comme une source d'information complémentaire ou optionnelle mise en place dans un système d'un autre type.

Ainai Samy [SAM - 86] traite la reconnaissance dynamique d'objet en mouvement par mise en correspondance des "points coins" entre deux images ; en se conformant à la vitesse calculée de manière différentielle pour ces mêmes points.

Tziritas [TZI - 88] estime le mouvement de manière récursive : la vitesse est prédite à partir des résultats des instants précédents ; la validité des hypothèses est testée par un détecteur de discontinuité fondé sur une analyse différentielle de l'image ; cette dernière fournit en outre une donnée pour le calcul de la vitesse estimée à posteriori.

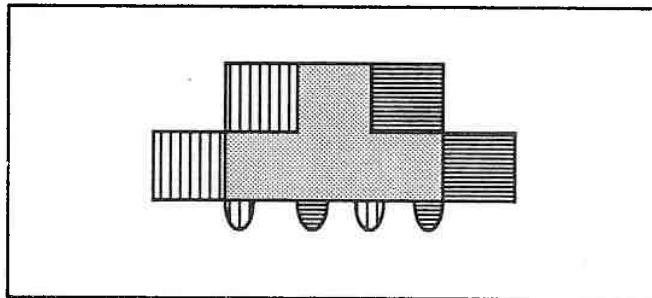
Enfin Millour [MIL - 88] utilise des mécanismes pré-attentifs pour trouver les zones d'intérêt d'une scène, puis propose en choix d'opérateurs de traitement d'image en fonction des cas à traiter. Le flot optique sert à initialiser une poursuite ou à analyser une fenêtre de grande taille.

En conclusion, nous soulignerons que ces méthodes sont intéressantes par la précision des résultats qu'elles délivrent, mais elles sont très sensibles au bruit, coûteuses en temps d'exécution, et nécessitent des images à grande définition. Notons sur ce point les travaux de Narathong [NAR - 87] visant à rapprocher le champ de réception de l'image, d'une rétine humaine. Dans ce but, la taille des pixels augmente en proportion avec la distance au centre du champ. Par ailleurs, ce mode d'analyse s'accommode bien des variations de forme des objets, c'est pourquoi il est approprié au codage d'images télévisuelles, ou à la mesure de mouvements d'un capteur.

I - 1 - 2 - Soustraction d'images successives

La même hypothèse de base que pour les méthodes différentielles est ici exploitée pour décrire des régions de l'image dont la nature est déterminée par les positions successives de l'objet.

Exemple



- (1) région fond 
- (2) région objet avant mouvement 
- (3) région objet après mouvement 
- (4) région objet avant et après mouvement 

Figure 3 : Véhicule dans deux positions successives

Ces régions sont extraites par différences entre images. Les résultats révèlent la forme des objets et leur trajectoire. Un algorithme typique est développé par Jain [JAI - 84 a] :

Pour deux images I_1 , I_2 représentant la même scène à deux instants différents, il distingue les régions différences positives (PDP) et les régions différences négatives (NDP) par le critère suivant.

Soit un seuil T_1

si $I_1(x,y) - I_2(x,y) > T_1$ alors PDP = 1
sinon PDP = 0

si $I_1(x,y) - I_2(x,y) < -T_1$ alors NDP = 1
sinon NDP = 0

La prise en compte du signe de la différence permet de distinguer le cas où l'objet couvre le fond de celui où le fond réapparaît derrière l'objet. Pour l'exemple de la figure 3, si la voiture est claire, PDP = 1 dans (2) et NDP = 1 dans (3).

Pour une séquence, la première image est prise comme référence (I_1) et les régions sont calculées à chaque étape par comparaison de I_1 à l'image courante ; puis accumulées suivant les équations.

$$[1.12] \quad \text{PADP}(x,y,t) = \text{PADP}(x,y,t-1) + \text{PDP}(x,y,t)$$

$$[1.13] \quad \text{NADP}(x,y,t) = \text{NADP}(x,y,t-1) + \text{NDP}(x,y,t)$$

La croissance respective des régions cumulées indique la forme de l'objet et le sens du mouvement.

PADP ne s'étend que sur le domaine couvert par le mobile à l'instant initial, et NADP croît dans le sens du mouvement. Lorsque l'objet a totalement quitté sa position de départ, la surface de PADP cesse d'augmenter, la région est dite "mature" et détermine la forme de l'objet. Nous visualiserons (figure 4) une séquence d'image et la valeur des régions qui lui sont associée à chaque étape.

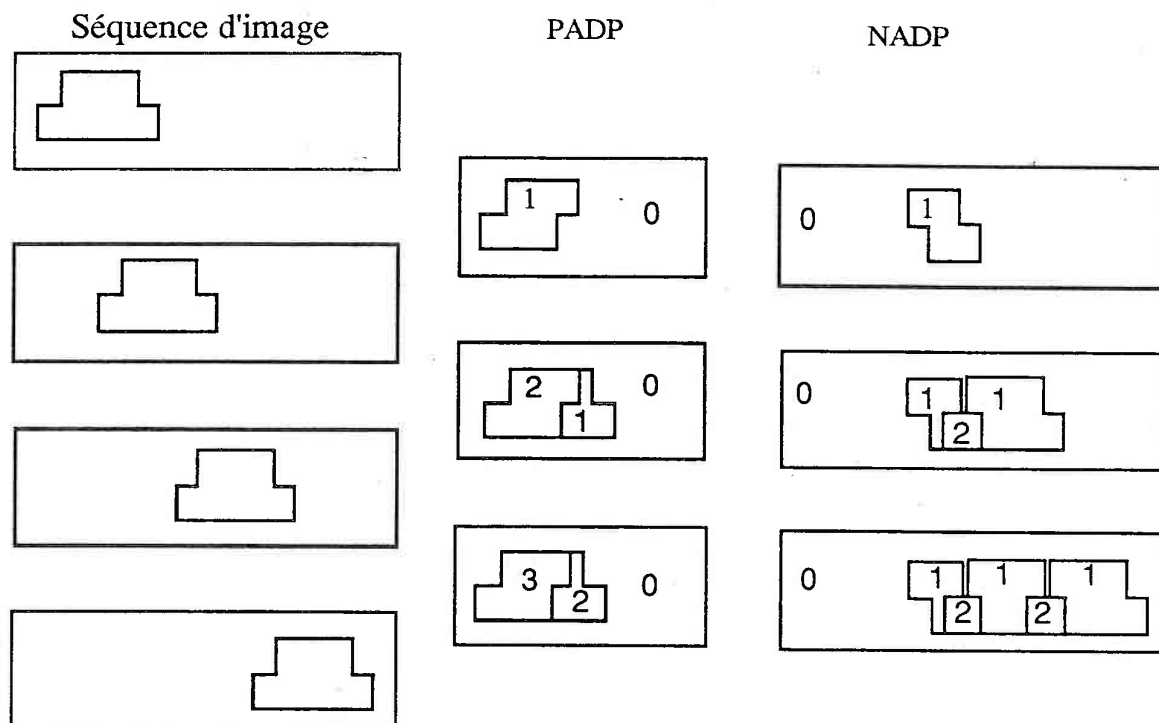


Figure 4 : Représentation des zones d'accumulation

Cet algorithme est applicable en présence de plusieurs mobiles, à condition d'introduire un test de non superposition mutuelle (test lié au taux d'accroissement mutuel de PADP et NADP). Le défaut de cette technique est la nécessité d'attendre que l'objet ait quitté sa position initiale pour en déterminer la forme ; ce délai peut être long. Mais une fois le modèle établi une poursuite temps réel est exécutable sur un processeur tel que le Motorola 68000.

Cette approche rappelle les méthodes de gradient en détection de contours, elle utilise une comparaison pixel à pixel entre images successives, elle nécessite donc soit une caméra fixe, soit un recalage préalable dans le cas d'un observateur mobile. Si la caméra est animée par un mouvement connu [JAI - 84 b], chaque trame est transposée dans l'espace d'EMP (égo motion parameters). Cet espace admet comme origine le point focal d'expansion de la caméra (cf figure 5 bis et définition) et varie en fonction du mouvement du capteur ; il permet l'application de la méthode de croissance des régions.

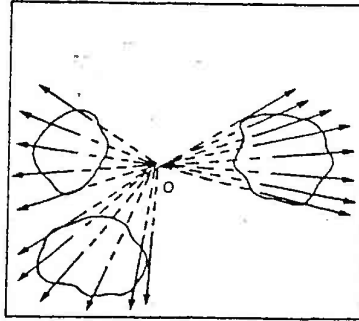


Figure 5 : Point focal d'expansion de la caméra (O)

Les vecteurs issus des représentations successives des objets immobiles convergent vers le point focal d'expansion de la caméra, il se trouve dans l'axe du mouvement de celle-ci.

L'exploitation des contours est aussi envisagée dans ce type de méthode. Stelmaszyk [STE - 85] extrait par un opérateur de Sobel les contours de l'image différence et ceux de l'image courante ; puis par multiplication pixels par pixels des résultats obtenus établit l'image des contours des objets en mouvement. Cette opération permet de dissocier les objets mobiles, des formes statiques présentes dans la scène ; elle est testée par simulation et donne avec précision la position des objets dans la scène. Son implantation sur site réel n'est toutefois pas réalisée. La détection de contours peut aussi être abordée [SUG - 86] par calcul du Laplacien temporel suivant la formule :

$$[1.14] \quad L(t_n) = I(t_{n+1}) - 2I(t_n) + I(t_{n-1})$$

avec

$I(t_n)$: image à l'instant t_n

$L(t_n)$ donne une représentation des bords de l'objet mobile, en temps réel au moyen d'un système numérique à architecture spécialisée MFIP (multimémoire - multiprocesseur). Les données sont exploitées par un algorithme de poursuite en trois étapes :

- (1) détection de l'objet mobile par Laplacien temporel
- (2) binarisation des contours par seuillage
- (3) localisation du centre de gravité de l'objet.

Cette technique est testée en temps différé sur des images de circulation automobile comportant plusieurs mobiles ; un traitement de prédiction vérification est établi pour les cas d'occlusion. La poursuite est donc simulée. Des variations brutales de vitesses sont possibles si elles ne surviennent à un instant où le mobile est masqué.

Un système similaire est développé par Wiklund [WIK - 87] qui soustrait trois images successives pour localiser un objet. Le suivi du mobile est traité par prédiction - vérification, la donnée nouvelle à chaque étape étant délivrée par différence. Cette approche appliquée à une scène de trafic automobile paraît fiable pour une séquence de courte durée (50 s). Ce type de poursuite est aussi retenu par Coulon [COU - 85] pour la mesure de trajectoire d'une petite particule (spermatozoïde ou rotoïde) ; l'initialisation étant faite manuellement. Ces dernières techniques font appel à des informations complémentaires de celles issues des soustractions. C'est aussi le cas du système décrit dans [MIL - 88] pour lequel les zones d'intérêt de la scène sont définies par quatre soustractions entre images successives ; la recherche des objets fait ensuite appel à d'autres techniques.

Pour conclure, ces méthodes permettent de mesurer les variations, de l'image dans le temps, qui révèlent sous certaines conditions les zones de mouvement dans la scène. Il est donc important pour envisager leur application de disposer d'images à fort contraste, d'une caméra fixe et d'objets mobiles rigides. Leur domaine d'opération privilégié est donc le trafic routier, d'autant qu'elles représentent l'avantage de rendre possible la poursuite multicible.

Les deux types d'approches exposées dans cette partie considèrent l'image comme une fonction variable du temps et de l'espace. Les premières méthodes, par leur analyse fine du mouvement en chaque point permettent d'envisager des objets de forme variable ou une caméra mobile ; les secondes par leur vision globalisante de la scène nécessitent des mobiles de forme fixe, mais admettent qu'ils soient multiples.

I - 2 - Méthodes par mise en correspondance

Nous abordons dans cette seconde partie les diverses techniques de mise en correspondance entre images. La relation établie à cette occasion sera ensuite utilisée comme base de l'analyse du mouvement. Les modes d'association sont variables : passage dans un espace transposé (Hough, Fourier, ...), corrélation, ou recherches séquentielles. Ils nécessitent en général la connaissance à priori de l'objet recherché.

I - 2 - 1 - Transposition dans un autre espace

Les travaux décrits dans cette partie utilisent un espace permettant la représentation des séquences d'image et la description des mouvements dont elles sont issues.

I - 2 - 1 - 1 - Transformée de Fourier

L'exploitation de la transformée de Fourier repose sur le fait qu'une translation se lit comme un déphasage dans l'espace transposé [ZAV - 88].

Soient :

$I_1(x,y)$: l'intensité des points image d'un objet à l'instant t_1 .

$I_2(x,y)$: l'intensité des points image du même objet à l'instant t_2 .

Si le mobile a subi un décalage dx, dy entre t_1 et t_2

$$[1.15] \quad I_2(x,y) = I_1(x + dx, y + dy)$$

dans l'espace de Fourier

$$[1.16] \quad I_2(u,v) = \iint I_2(x,y) \exp(-2j\pi(ux + vy)) dx dy$$

et l'équation [1.15] devient

$$[1.17] \quad I_2(u,v) = I_1(u,v) \exp(2j\pi(udx + vdy))$$

le déphasage entre I_2 et I_1 est $d\phi$:

$$[1.18] \quad d\phi = udx + vdy$$

la connaissance de $d\phi$ pour plusieurs couples (u,v) permet de calculer le mouvement d'origine (dx,dy) ; l'intérêt principal du passage à une représentation en fréquence est la réduction du nombre des données à traiter, en particulier pour les images de grande taille. Cependant, bien que cette technique soit souvent citée comme référence, elle se heurte à de multiples obstacles pour sa mise en oeuvre éventuelle, citons :

- l'indétermination de la phase à $2k\pi$ près,
- la nécessité d'un grand nombre de calcul,
- la limitation de la validité des résultats au cas d'un objet se déplaçant sur un fond uniforme.

Bien que des solutions puissent être envisagées pour chacun de ces problèmes, cette méthode est rarement utilisée sous sa forme originale. Mais elle constitue le fondement de plusieurs techniques. Par exemple dans [WAL - 80] on utilise la transformée de Fourier monodimensionnelle des contours d'avions comme codage de ceux-ci. La position des mobiles est ensuite repérée par comparaison du codage courant à un dictionnaire représentant les avions en différents emplacements ; puis leur mouvement est déduit par un critère de continuité.

I - 2 - 1 - 2 - Transformée de Hough

Ce modèle [SKL - 78] [MOH - 88] est issu de la reconnaissance des formes. Il a été initialement mis au point pour détecter des figures, décrites par un nombre limité de paramètres, dans des images bruitées. Par exemple, une droite d'équation $y = ax + b$ est représentée par le couple (a, b) , un cercle a pour image son centre et son rayon, etc ...

La généralisation de cette approche aux séquences d'images repose donc sur une paramétrisation des formes contenues dans la scène. Aggarwal [AGG - 81] présente un système de ce type pour l'analyse des mouvements des nuages. Ceux-ci sont décrits par les points de leur contour : $B = \{(X_i, Y_i)\}$ et par un point donné $p = (X_p, Y_p)$ qui peut être le barycentre des (X_i, Y_i) ; la représentation de Hough de B à partir de p est donnée par :

$$[1.19] \quad H(B, p) = \{d_i\}_{i=0, n}$$

avec

$$[1.20] \quad d_i = (dX_i, dY_i) \quad dX_i = X_p - X_i \quad dY_i = Y_p - Y_i$$

Si $H(B, p)$ représente l'objet dans l'image à l'instant t ; à partir d'une acquisition à $t + dt$ le nouveau point p' , barycentre des contours, va être déterminé en exploitant des accumulateurs h , suivant la procédure ci-dessous :

Initialisation

A chaque point (X, Y) de l'image associer un accumulateur vide
 $h(X, Y) = 0$

Localisation

Pour chaque point des contours $e_j = (X_j, Y_j)$ et pour chaque élément $d_i = (dX_i, dY_i)$ de $H(B, p)$, incrémenter l'accumulateur du point $(X_j + dX_i, Y_j + dY_i)$

$$[1.21] \quad h(X_j + dX_i, Y_j + dY_i) = h(X_j, Y_j) + 1$$

Le point p' représentatif du nouvel emplacement du mobile sera celui dont l'accumulateur est le plus rempli.

$$p' = (X_{p'}, Y_{p'}) / \forall (X, Y) \quad h(X_{p'}, Y_{p'}) > h(X, Y)$$

Exemple :

(1) Définition

La figure ci-dessous illustre la procédure de modélisation, en ne considérant que trois points des contours de l'objet (e_1, e_2, e_3).

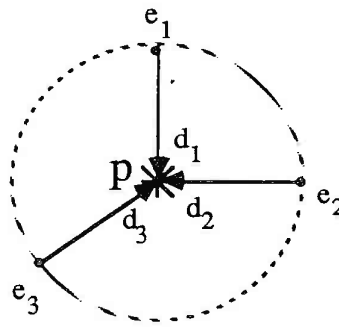


Figure 6 : Calcul de $H(B,p)$ sur trois points

L'objet est donc défini à partir des points $\{e_1, e_2, e_3\}$ et du point p par :

$$H(B,p) = \{d_1, d_2, d_3\}$$

(2) Localisation

Une image contenant cet objet est ensuite captée, une technique classique d'extraction de contours permet de détecter 4 points (c_1, c_2, c_3, c_4)

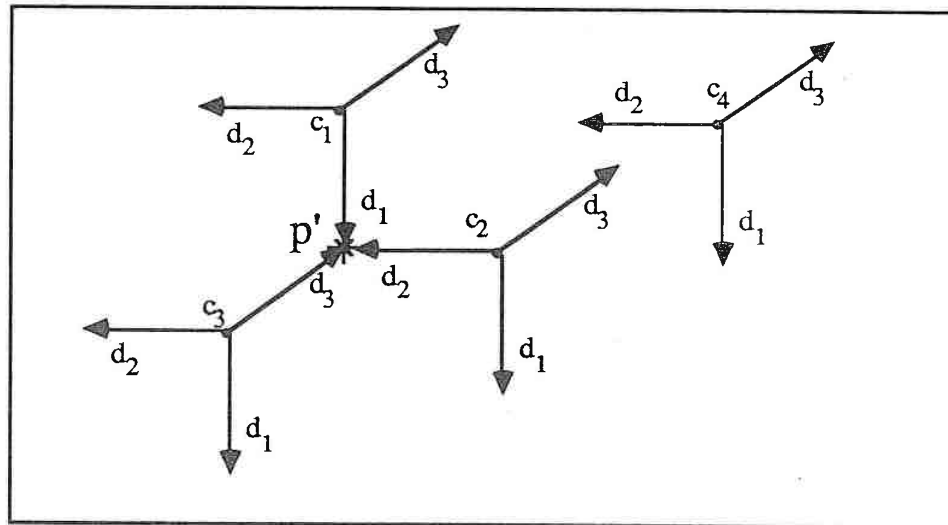


Figure 7 : Recherche du point p'

La procédure d'incrémentation des accumulateurs est :

pour $i = 1$ à 4

et pour $j = 1$ à 3

$$h(c_i + d_j) = h(c_i) + 1$$

Le seul point de l'image ayant son accumulateur à 3 sera celui correspondant à

$$c_1 + d_1 = c_2 + d_2 = c_3 + d_3$$

ainsi p' est donné par $h(p') = 3$ et les points représentatifs de l'objet sont (c_1, c_2, c_3) .

Les résultats délivrés par cet algorithme sont sensiblement les mêmes que ceux d'une corrélation. Il peut être modifié pour prendre en compte une rotation dans le mouvement, les temps de traitement et les emplacements mémoire nécessaires sont alors beaucoup plus importants.

L'intérêt d'une telle méthode est qu'elle offre une grande immunité au bruit ou aux occlusions, elle peut être améliorée [SAM - 84] par des informations complémentaires sur le mouvement ; toutefois, les temps de traitement requis sont souvent incompatibles avec une poursuite de cible en temps réel.

I - 2 - 1 - 3 - Transformation pour l'interprétation

Dans cette dernière partie, nous évoquons le problème de la vision 3D où une image en 2D constitue une transposition (projection) de l'espace observé. Une opération inverse est donc nécessaire pour replacer les données dans l'espace (reconstruction). Faugeras dans [FAU - 88] donne une étude critique des recherches développées et des axes de recherches prometteurs.

Comme, nous nous intéressons au mouvement nous mentionnerons l'approche élaborée dans [TSA - 81] - [TSA - 83] où les auteurs recherchent les caractéristiques du déplacement d'un solide dans l'espace. Dans ce but, ils calculent les "purs paramètres" $(a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8)$ définis par :

$$[1.22] \quad X' = \frac{a_1 X + a_2 Y + a_3}{a_7 X + a_8 Y + 1} \quad Y' = \frac{a_4 X + a_5 Y + a_6}{a_7 X + a_8 Y + 1}$$

où (X, Y) représente un point du solide dans le plan image avant le mouvement et (X', Y') est ce même point après le mouvement.

Si le solide considéré est un plan, ces paramètres sont connus par l'emplacement de quatre points du mobile avant et après le déplacement ; à condition qu'aucun groupe de trois points ne soit aligné et que le plan ne soit pas dans l'axe du point focal de la caméra.

Le calcul des "purs paramètres" est expliqué ainsi que leurs liens avec les translations et rotations représentant le mouvement.

Le cas de formes courbes est aussi envisagé. Il s'agit ici d'un formalisme complexe reposant sur des connaissances mathématiques spécialisées (telles les groupes de Lie). Les calculs sont donc assez lourds mais les résultats obtenus montrent une grande fiabilité de l'approche à condition que les points de correspondance entre images soient connus avec une grande précision. La technique d'obtention de ces points n'est pas évoquée, une mise en oeuvre réelle n'est donc pas envisageable sans un approfondissement préalable.

En conclusion, nous remarquons que ces méthodes par transposition sont en général attrayantes par les modélisations théoriques qu'elles proposent. Mais elles requièrent des temps de traitement ou des emplacements mémoire très importants ; elles sont donc exploitables pour des analyses en temps différé, mais mal adaptées aux problèmes tels que la poursuite ou la commande de véhicule.

I - 2 - 2 - Association modèle - image

La méthode est fondée sur l'analyse statique de la scène : la localisation d'un modèle prédéterminé ; elle est établie en plusieurs instants, puis les informations extraites sont liées pour déterminer la trajectoire de l'objet.

I - 2 - 2 - 1 - La Corrélation

La technique de base pour la mise en correspondance entre deux images est le calcul de leur fonction d'intercorrélation et la recherche du maximum de cette dernière.

Soient :

$S(x,y)$: image de $N \times N$ pixels représentant une scène

$R(x,y)$: image de $M \times M$ pixels représentant un objet.

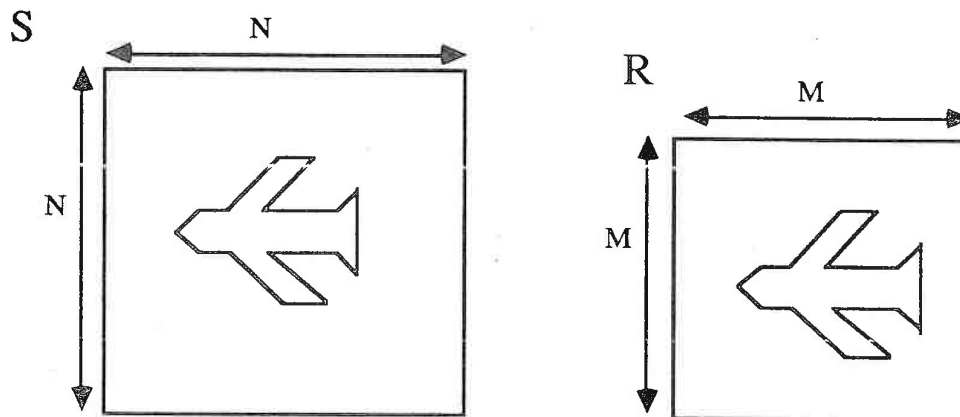


Figure 8 : Image de recherche et modèle de l'objet

Si $C(u,v)$ est la valeur de la corrélation normée entre S et R pour R débutant au point (u,v) de S .

$$[1.23] \quad C(u,v) = \frac{\sum_{i=1}^M \sum_{j=1}^M R(i,j) S(i+u, j+v)}{\left[\sum_{i=1}^M \sum_{j=1}^M R^2(i,j) \right]^{1/2} \cdot \left[\sum_{i=1}^M \sum_{j=1}^M S^2(i+u, j+v) \right]^{1/2}}$$

le point (u_c, v_c) de meilleure correspondance entre S et R est donné par le calcul de $C(u,v)$ aux $(N - M + 1)^2$ positions possibles pour R ; et par la relation

$$[1.24] \quad \forall (u,v) \neq (u_c, v_c) \quad C(u_c, v_c) > C(u,v)$$

En présence de bruit ou de mauvais contraste dans les images, le maximum peut être faiblement marqué et sa détection est difficile. Les incertitudes engendrées par ce type de problème sont formalisées dans [OLI - 84] : une étude analytique est menée pour établir le lien entre la surface de corrélation et les caractéristiques de l'image. A partir de formalismes obtenus, la probabilité de détection du pic est calculée en fonction des paramètres définissant les bruits.

Afin d'augmenter la dynamique de la fonction de corrélation et donc de faciliter la détection de son pic, Pratt [PRA - 78] établit un filtrage variable avec le taux d'autocorrélation des images. Quand le taux est élevé, le filtre est proche d'un extracteur de contour, sinon il est sans effet. La qualité des fonctions de corrélation issues des hautes fréquences spatiales de l'image justifie l'approche que nous développerons par la suite.

Dans le but d'augmenter le rapport signal sur bruit de la fonction de corrélation, Garret [GAR - 76] utilise trois fonctions de synthèse procédant à l'appariement des points en prenant en compte l'image dans son ensemble. Les résultats obtenus sont intéressants pour ce qui concerne le rapport signal sur bruit mais les calculs requis sont lourds.

D'une manière générale les méthodes corrélatives font appel à un grand nombre d'opérations, les moyens mis en oeuvre pour limiter les temps de traitement peuvent être de deux ordres :

- utilisation d'une architecture informatique appropriée permettant la parallélisation des tâches,
- calcul de l'appariement dans des zones privilégiées, déterminées par une analyse des mouvements antérieurs ou par une approche séquentielle de la corrélation.

I - 2 - 2 - 2 - La recherche séquentielle

La publication référence dans le domaine de la conception d'algorithmes séquentiels connus sous l'abréviation SSDA (Sequential Similarity Detection Algorithm) est la communication de Barnéa et Silverman [BAR - 72]. Il s'agit d'associer un modèle à une partie de l'image sans calculer la distance objet - image en tout point de cette dernière.

Soient

S : la scène de $N \times N$ points

R : le modèle de $M \times M$ points

$P_{u,v}$: la partie de la scène de M^2 points débutant au point (u,v) .

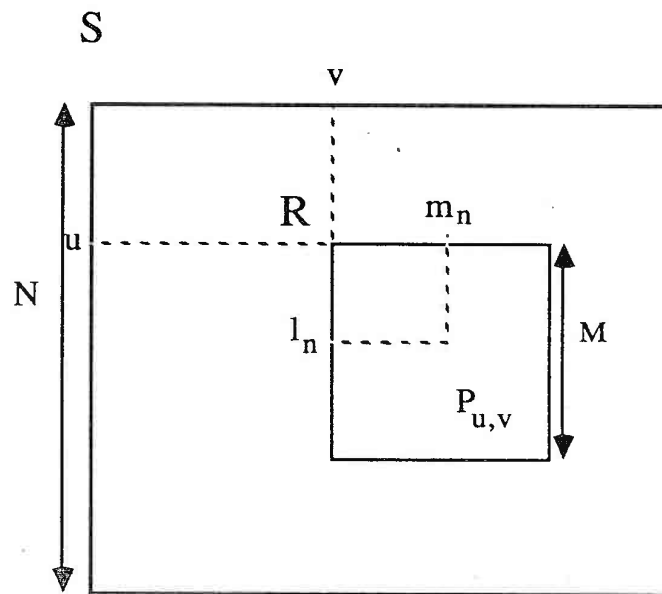


Figure 9 : Modèle débutant au point (u,v) de l'image

Pour $n = 1, 2 \dots M^2$ l'erreur e_n à l'étape n est définie par :

$$[1.25] \quad e_n(u,v) = |R(l_n, m_n) - P_{u,v}(l_n, m_n)|$$

l'écart $E(u,v)$ entre S et $P_{u,v}$ est alors :

$$[1.26] \quad E(u,v) = \sum_{n=1}^{M^2} e_n(u,v)$$

L'originalité de l'algorithme réside dans la limitation du nombre de terme e_n à calculer. Dans ce but, un seuil T est défini, et lorsque l'erreur cumulée ($\sum e_n$) atteint ce niveau, les opérations cessent pour le point (u,v) de recherche considéré. Elles impliquent un nombre de point $r < M^2$. Le module du SSDA au point (u,v) est noté $I(u,v)$ défini par :

$$[1.27] \quad I(u,v) = \left\{ r / \min_{1 \leq r \leq M^2} \left\{ \sum_{n=1}^r e_n(u,v) \geq T \right\} \right\}$$

Exemple :

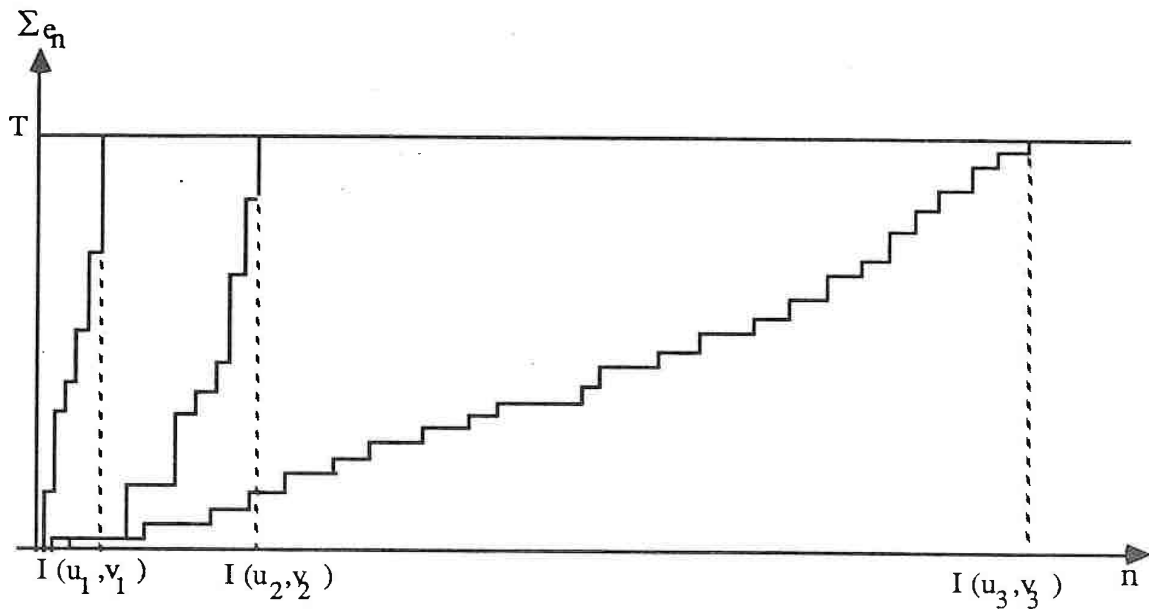


Figure 10 : Représentation du SSDA

$$I(u,v) = 9 \quad I(u_2, v_2) = 33 \quad I(u_3, v_3) = 137$$

La localisation du modèle en (u_3, v_3) présente une meilleure similitude qu'en (u_1, v_1) ou (u_2, v_2) . Le nombre de comparaisons requises pour l'évaluation de I est égal au nombre d'itérations opérées avant que l'erreur cumulée n'atteigne T (9 pour (u_1, v_1) , 137 pour (u_3, v_3)). En conséquence, plus la similitude est grande, plus les calculs nécessaires à son évaluation sont nombreux. Mais ce traitement est toujours moins coûteux qu'une corrélation classique.

Des études ont été menées à partir de la technique de base décrite ci-dessus pour déterminer une optimalité des paramètres suivant :

- le seuil T
- la loi de variation de l_n, m_n
- la procédure de recherche du maximum de I

Pour le seuil, des essais effectués pour différentes valeurs de T montrent que les valeurs moyennes et maximales de I (sur l'image) varient proportionnellement à T . Pour avoir une bonne précision, il faut que I soit grand afin de bien dissocier les points similaires, mais alors le coût en temps de la méthode est important. Ces constatations ont porté Barnéa et Silverman à mettre au point des séquences de seuils $T(n)$ croissants avec le nombre d'itération comme l'indique la figure ci-dessous.

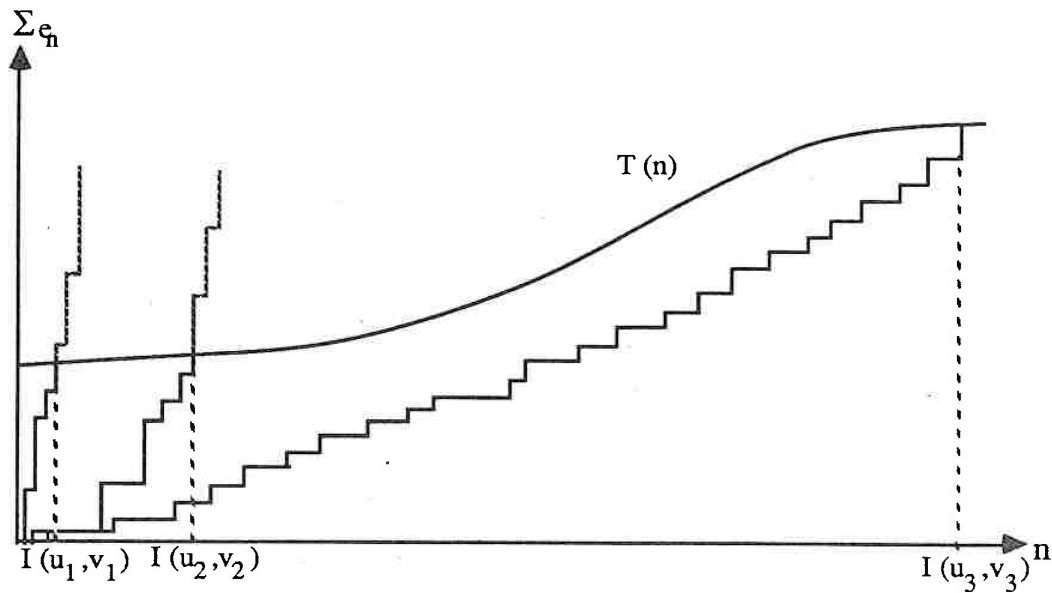


Figure 11 : SSDA à seuil croissant

$$I(u_1, v_1) = 6 \quad I(u_2, v_2) = 26 \quad I(u_3, v_3) = 178$$

Ainsi, le nombre de calcul est restreint aux points de mauvaise correspondance $((u_1, v_1), (u_2, v_2))$, et dans le cas de bonne similitude (u_3, v_3) , la valeur de I est amplifiée. Deux procédures sont envisagées pour modéliser les variations de $T(n)$.

- la prise en compte de la moyenne a et de l'écart type σ du bruit présent dans l'image (moyenne - écart type) :

$$[1.28] \quad T(n) = an + \sigma \sqrt{n}$$

- l'affectation à chaque étape de la même probabilité de dépassement du seuil en un point donné (seuil équiprobable).

Les deux séquences obtenues diffèrent fondamentalement par les stratégies qui leurs sont sous-jacentes. Mais Munteanu [MUN - 81] montre que les valeurs de seuil engendrées sont proches. Toutefois lorsque le bruit est fort, une tendance à la saturation est observée, pour l'approche "moyenne - écart type", aux points de bonnes correspondances. Il est donc préférable de choisir le "seuil équiprobable" dans ce cas. Cependant, son étude analytique sur des images binaires montre qu'aucune de ces deux séquences n'est optimale concernant le rapport signal sur bruit. Mais leurs performances sont très proches et laissent penser que si une suite de seuils optimale existe, elle engendrera des résultats du même ordre.

L'objectif du SSDA étant de réduire les coûts de traitement, nous mesurons son apport en comparant le nombre d'équivalent addition requis pour plusieurs méthodes [BAR - 72] :

- corrélation directe $\rightarrow 4,5 M^2 (N - M + 1)^2$
- FFT $\rightarrow 200 N^2 \log_2 N$
- SSDA $\rightarrow (4 + 4 \times \frac{10}{(M/32)^{1/2}}) (N - M + 1)^2$ (si $M > 32$)

Soit pour une aire de recherche de 256×256 pixels et un modèle de 64×64 , le SSDA est 50 fois plus rapide que la FFT et 300 fois plus rapide que la corrélation directe. Il s'agit ici d'une évaluation moyenne, la technique du SSDA n'étant pas systématique le nombre d'opération nécessaire dépend du contexte dans lequel elle est utilisée.

Jusqu'à présent, nous avons considéré que l_n et m_n varient continuellement de 0 à M ; il est possible de choisir des paliers, c'est le cas de la technique "Two stage template matching" [AGG - 81] qui comme son nom l'indique opère en deux étapes :

Etape 1 : recherche d'un sous modèle de K points dans l'image, et sélection d'un certain nombre d'emplacements pour lesquels la recherche est fructueuse.

Etape 2 : recherche parmi les emplacements sélectionnés de celui présentant la meilleure correspondance avec les M^2 points du modèle.

Pour la première étape, les points sont sélectionnés si la similarité entre le sous-modèle et l'image dépasse un seuil t prédéterminé ; en supposant qu'il existe une probabilité P pour que cet événement soit réalisé, le nombre de comparaisons nécessaires pour l'application de cette méthode est :

$$[1.29] \quad N_c = (N - M + 1)^2 \times K + P \times (N - M + 1)^2 \times M^2$$

Le coût de la réalisation et la validité des résultats dépendent donc des valeurs de K et t .

La probabilité de fausse alarme P_{rF} est la probabilité que l'appariement entre le sous modèle et l'image soit supérieure à t en un point erroné. Afin de choisir les valeurs de K et t , Vanderbrug [VAN - 77] définit comme fonction coût, la somme de la taille du sous modèle (K) et de la probabilité de fausse alarme (P_{rF}). Il montre analytiquement que sur des images binaires cette fonction possède un minimum ; soit qu'à partir d'une certaine taille de la sous-fenêtre, l'accroissement de celle-ci n'entraîne pas une notable diminution de P_{rF} .

Il existe donc pour chaque valeur de t une taille optimale K . La généralisation de ce calcul aux images à niveaux de gris n'est pas réalisable, la probabilité de fausse alarme dépendant autant des luminances présentes dans le sous modèle que de la taille de ce dernier. Cependant à partir d'une certaine connaissance des caractéristiques statistiques de l'image, il est possible de se référer au cas binaire.

Dans une optique plus qualitative Li [LI - 85] teste en préliminaire du "Two stage template matching" si l'objet est présent dans l'image ; il propose dans ce but deux statistiques basées sur des critères de vraisemblance.

Une autre approche issue du SSDA est établie par Wong [WON - 78], c'est un algorithme séquentiel hiérarchique. La première étape de la séquence de recherche est une comparaison entre un modèle de bas niveau de résolution et l'image de même nature, après avoir effectué un filtrage passe bas.

Cette première comparaison ne nécessite que peu d'opérations et permet la sélection des points de bonne correspondance. Pour ceux-ci, l'appariement est à nouveau exécuté avec un niveau de résolution plus élevé, (pulsation de coupure du filtre plus grande) engendrant un second niveau de choix. Le processus est ainsi renouvelé jusqu'à traiter l'image à son niveau de définition nominal, et en déduire la position de l'objet. Pour chaque fréquence de coupure, les associations sont réalisées suivant les modalités décrites au début du paragraphe. Le seuil de décision T_n^k de la $n^{\text{ième}}$ paire de point au $k^{\text{ième}}$ niveau de recherche est calculé par :

$$[1.30] \quad T_n^k = \gamma_m \sqrt{2^{(n-k)}} (n + g_k \sqrt{n})$$

avec γ_m : erreur moyenne de mesure

g_k : densité de probabilité de correspondance au niveau k .

Cet algorithme repose donc sur l'interaction entre deux facteurs de réduction du nombre d'information à traiter : niveau fréquentiel (filtrage passe bas) puis niveau spatial (considération des points du modèle par ordre croissant de l_n et m_n).

L'évaluation de l'économie en temps de calcul η est calculée comme suit :

Soient

C_1 : nombre de comparaisons pour une corrélation directe

$$[1.31] \quad C_1 = (N - M + 1)^2 (M^2)$$

C_2 : nombre de comparaison par niveau de fréquence

$$[1.32] \quad C_2 = K_1 M^2 \log (N - M + 1)^2$$

où K_1 est une constante dépendant de la nature de l'image

C_3 : nombre d'opération pour créer les tableaux filtrés

$$[1.33] \quad C_3 = 2 \sum_{i=1}^{L-1} (N/2^i)^2$$

où L est le nombre de niveau de fréquence

alors

$$[1.34] \quad \eta = \frac{C_1}{C_2 + C_3} = \frac{(N - M + 1)^2}{K_1 \log (N - M + 1)^2 + 2 \sum_{i=1}^{L-1} (N/2^i)/M^2}$$

A titre d'exemple, pour une image de 256 x 256 points, un modèle de 64 x 64 points et 4 niveaux de recherche, le gain en calcul varie de 1 000 à 3 000 suivant les valeurs de K_1 .

Nous avons décrit dans cette partie le mode de recherche séquentiel du maximum de la fonction de corrélation, et donné les axes d'amélioration de la méthode de base en évaluant pour chacun d'entre eux le bénéfice apporté.

Une autre approche séquentielle de l'analyse du mouvement a été mise au point par Le Guilloux [GUI - 84] pour un système de poursuite de chars. Cette technique diffère radicalement des précédentes car elle ne nécessite aucune connaissance à priori de l'objet recherché. Inspirée du traitement du signal, elle repose sur l'établissement du chemin optimal liant une à une les lignes respectives de deux images successives. Ceci étant réalisé par programmation dynamique à travers le tableau des écarts pixels à pixels entre les deux trames considérées.

La figure ci-dessous illustre le résultat obtenu pour une ligne i entre les images n et $n + 1$.

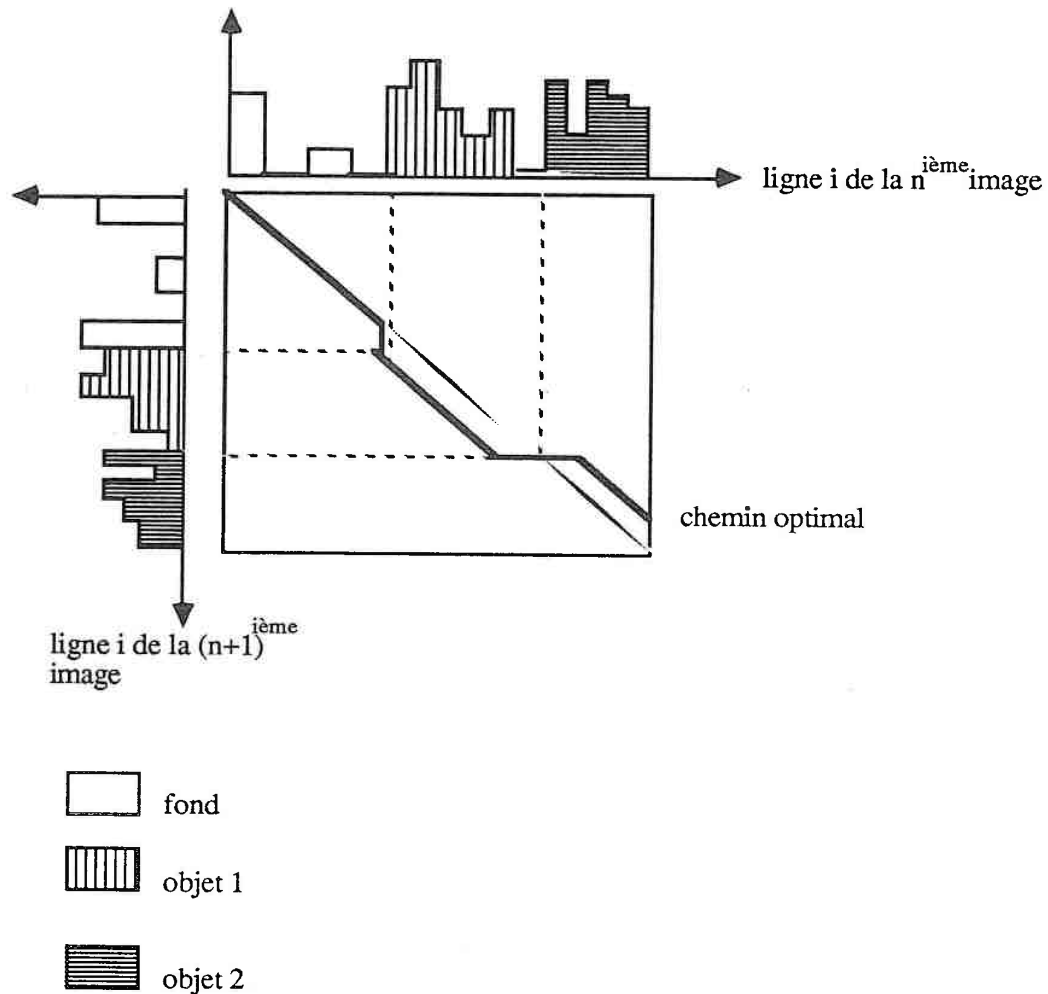


Figure 12 : tableau de mise en correspondance

Dans cet exemple, l'objet 1 s'est déplacé vers le droite et l'objet 2 vers la gauche, ces deux mouvements sont lisibles sur le tableau (fig. 12) par la distance du chemin optimal à la diagonale, aux emplacements des objets considérés. Les déplacements sont ainsi mesurés pour chaque ligne, les formes des mobiles sont fixées à postériori par les ensembles de points animés par un même mouvement. Cet algorithme s'applique à des translations horizontales de faible amplitude, mais il ne requiert aucune autre hypothèse ou connaissance à priori sur les scènes analysées.

I - 2 - 2 - 3 - Les recherches orientées

La mise en évidence de direction dans lesquelles l'objet est probablement présent précède la recherche de ce dernier. Ces orientations privilégiées peuvent être détectées par plusieurs types d'analyse ; par exemple :

- détermination grossière des zones variables dans l'image,
- interprétation des caractéristiques des déplacements aux instants précédents,
- hypothèse sur les déplacements possibles, sur la nature des images ...

Ainsi, la recherche du point de maximum de corrélation se fait par un parcours non aléatoire au sein de l'image.

Jain [JAI - 81] développe un algorithme typique pour coder des séquences représentatives de scènes animées ; il s'agit ici de connaître les mouvements de différents sous blocs de l'image entre deux acquisitions. Soit U un sous bloc d'une trame, de taille $M \times N$, et U_R la partie de la trame suivante dans laquelle le sous bloc U est recherché. U_R est de taille $(M + 2p) \times (N + 2p)$, où p est le nombre de pixels représentant le déplacement maximum possible de U . La distortion qui est représentative de l'écart entre U_R et U positionné en (i,j) est mesurée par $D(i,j)$:

$$[1.35] \quad D(i,j) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (U(m,n) - U_R(m+i, n+j))^2$$

La recherche directe du point (i,j) , pour lequel D est minimum, nécessiterait sa mesure en $(2p + 1) \times (2p + 1)$ emplacements (121 si $p = 5$). L'hypothèse complémentaire exploitée ici est la décroissance de $D(i,j)$ dans la direction menant à son minimum, appelée direction du minimum de distortion (DMD).

La procédure de localisation du bloc U dans U_R est donc la suivante (illustrée fig. 11) :

Algorithme du DMD

- Initialisation

* calcul du premier pas de recherche

$n = \text{partie entière de } p/2$

* définition du premier point origine de la procédure

$(i,j) = (0,0)$

au centre de U_R

- Recherche

- (1) * définition des 5 points de mesure de $D(i,j)$ représentant l'immobilité $(0,0)$ et les déplacements suivant les 4 directions. Nord Sud Est et Ouest, pour le pas n .

$$M(n) = \{(0,0), (-n,0), (+n,0), (0,n), (0,-n)\}$$

- (2) * calcul de l'écart entre U_R et U
pour $(q,l) \in M(n)$ évaluer $D(i + q, j + l)$

* détermination de la direction du minimum de distortion représentée par l'axe du point où D est minimum

$$(q_m, l_m) = \underset{M(n)}{\text{Min}} D(i + q, j + l)$$

* si la DMD n'est pas l'immobilité $((q_m, l_m) \neq (0,0))$

Alors recommencer la recherche avec le même pas, en partant du point d'écart minimum, et en éliminant le point origine de l'étape précédente.

$$+ i = i + q_m \quad j = j + l_m$$

$$+ M(n) = M(n) - (-q_m, -l_m)$$

Sinon réduire le pas de recherche

$$n = n/2$$

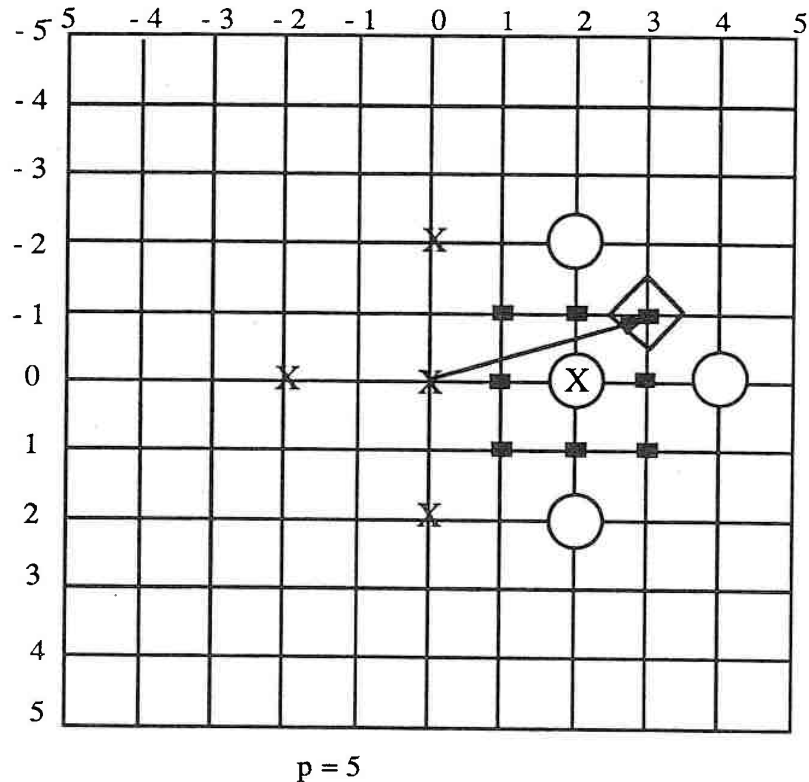
si $n > 1$ alors recommencer la procédure aller en (1)

sinon donner le point précis pour lequel D est minimum parmi les 8 voisins de (i,j) et lui-même.

$$\text{calculer } D(i + u, j + v)$$

$$\text{pour } u = \{-1, 0, 1\} \quad v = \{-1, 0, 1\}$$

la position du bloc U est donnée par $\text{Min}_{(u,v)} D(i + u, j + v)$.



1ère itération : points X $n = 2$ $(q_m, l_m) = (0, 2)$

2ème itération : points O $n = 2$ $(q_m, l_m) = (0, 0)$

3ème itération : points ■ $n = 1$ $(q, l) = (-1, 1)$

le maximum est trouvé après 17 comparaisons

Figure 13 : Localisation du bloc U par l'algorithme du DMD

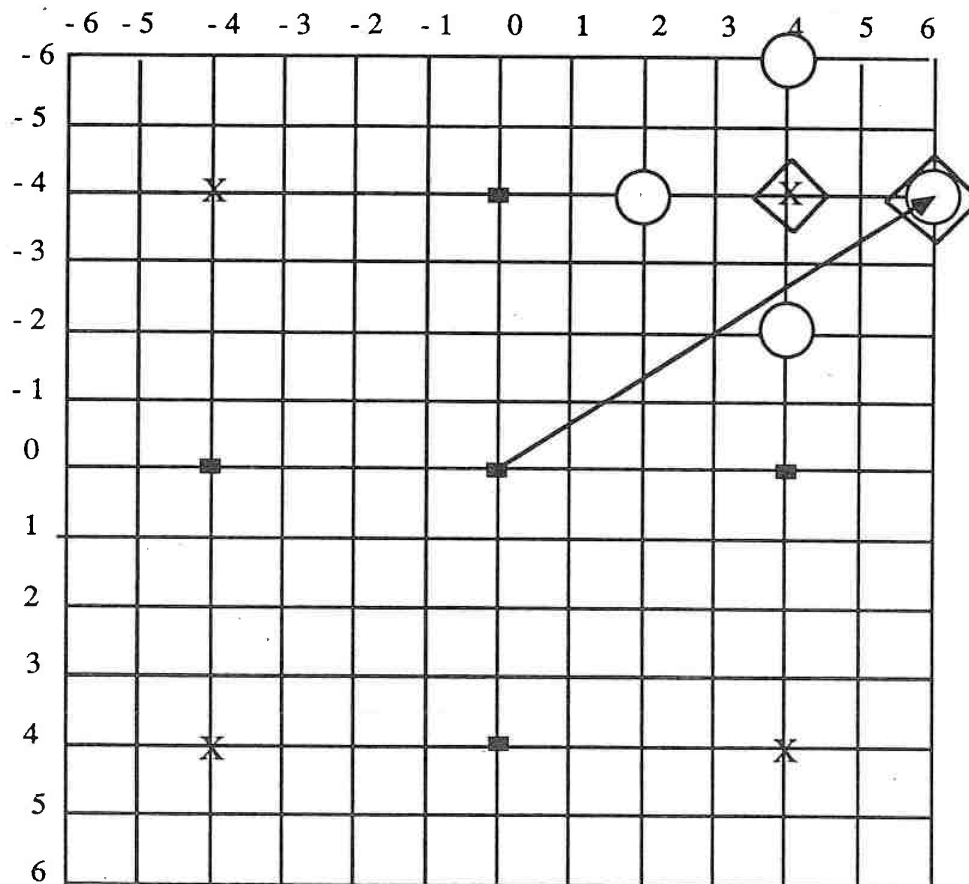
Le nombre d'emplacement de recherche est de 13 à 21 selon les cas (contre 121), soit une réduction de 80 à 90 % par rapport à la corrélation directe. Cette procédure est appliquée pour des codages d'image TV de 256 x 256 pixels divisées en blocs de 16 x 16 dont on cherche les mouvements entre chaque acquisition (tous les 1/24^{ème} de seconde). Ceci suppose que les déplacements sont de faible amplitude, condition indispensable pour que l'hypothèse de monotonie de $\bar{D}(i,j)$ soit vérifiée.

Une méthode voisine est exploitée par Koga [KOG - 83] pour un même type d'application ; à chaque étape, le bloc est recherché parmi les huit points voisins du point considéré ; le pas de recherche étant décrémenté à chaque niveau. Pour un déplacement maximal de 6 pixels, cet algorithme donne la position d'un bloc à partir des mesures en 25 emplacements.

Enfin Kappagantala [KAP - 85] combine les deux techniques précédemment citées en ajoutant un test sur la mesure de la distortion.

Ainsi, à l'étape (2) de l'algorithme du DMD, D est comparé à une limite D_{lim} ; si D est inférieur à D_{lim} , le sous bloc est considéré comme se trouvant à la position (i,j) et la recherche s'arrête. De la même manière, les investigations sont d'abord faites dans les directions horizontales et verticales, si en chacun des 5 points testés la distortion est supérieure à D_{lim} , les 4 points des diagonales sont testés, sinon la recherche est affinée pour l'étape suivante.

Un exemple de séquence est donné dans la figure ci-dessous pour un déplacement maximal de 6 pixels.



- | | |
|----------------|---|
| 1ère itération | points ■ : $D > D_{lim}$ pour les 5 points |
| 2ème itération | points X : le point $(-4,4)$ est choisi |
| 3ème itération | points ○ : $D(0,2) < D_{lim}$ |
| | le bloc est positionné en $(-4,6)$ par un calcul en 13 places |

Figure 14 : Algorithme de Kappagantala

L'exemple ci dessus montre la possibilité pour cette procédure d'être stoppée, avant une recherche systématique du meilleur point parmi les neuf voisins. Cependant les résultats obtenus sont très proches de ceux du DMD, et le nombre d'opération est ramené de 1 à 19 pour un déplacement maximal de 6 pixels ; contre 169 pour une corrélation directe.

Plus récemment Bonnin [BON - 88] a développé un système de poursuite de cible par imagerie infra - rouge. Pour le problème traité les variations du contenu de l'image peuvent être importantes, l'hypothèse de continuité temporelle est alors hasardeuse ; une méthode de recherche des régions d'intérêt est alors entreprise. Les points à variance élevée sont extraits, puis une méthode de croissance guidée par les mesures des luminances des points du voisinage, permet la délimitation de régions. Le modèle initial des objets est établi automatiquement au cours d'une phase d'apprentissage. Après quoi la poursuite est exécutée en deux étapes :

- association entre le modèle de l'objet et la région de l'image la plus ressemblante
- réactualisation du modèle.

Il est ainsi envisageable de poursuivre simultanément plusieurs objets.

Citons enfin Millour [MIL - 88] qui utilise la différence entre images successives pour déterminer les fenêtres d'intérêt dans lesquelles les objets en mouvement seront recherchés ; et Coulon [COU - 85] qui prévoit par la vitesse mesurée à l'étape précédente, la zone dans laquelle une particule mobile sera supposée apparaître.

Pour conclure sur ces méthodes par association entre modèle et image, présentons une étude comparée des différentes fonctions utilisées pour l'appariement publiée, par Sécilla [SEC - 87]. Elle concerne :

- 1 - somme des valeurs absolues des différences
- 2 - somme des carrés des différences
- 3 - fonction de corrélation
- 4 - coefficient de corrélation
- 5 - différence de phase de la transformée de Fourier
- 6 - nombre de changement de signe dans l'image différence
- 7 - estimation de la position par le maximum à postériori.

Les tests sont exécutés sur des images de 256 x 256 points contenant un objet de 32 x 32 points en leur centre. Ces images sont artificiellement bruitées, la localisation de l'objet est exécutée plusieurs fois, et pour différentes valeurs du rapport signal sur bruit.

La distance moyenne, entre la position estimée et le point de coordonnées (16,16) du modèle, est enregistrée. Les résultats pour un bruit additif (A) et multiplicatif (B) sont visualisés sur les courbes ci-dessous.

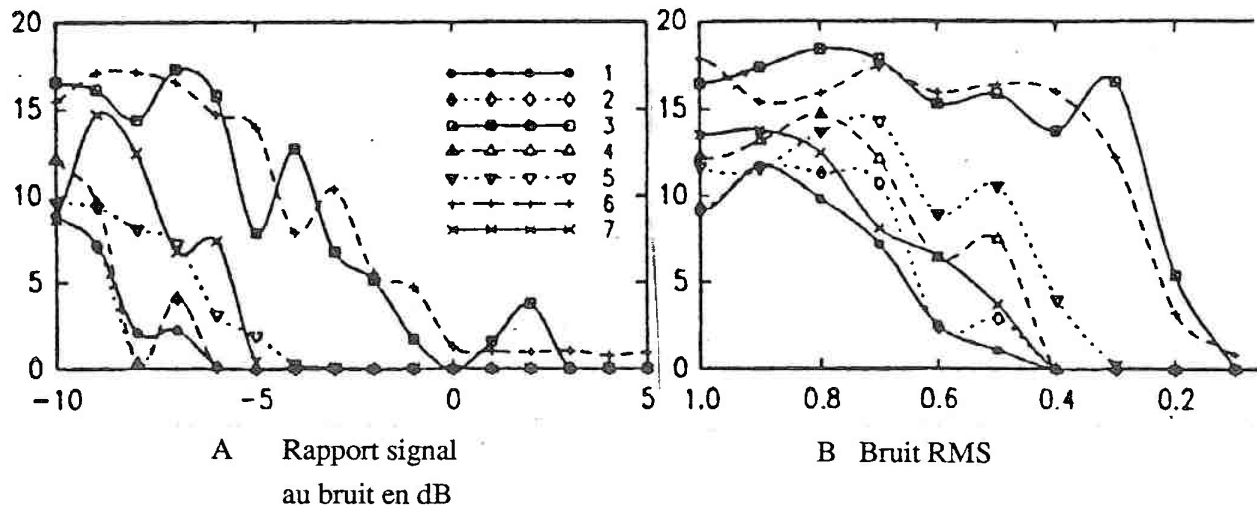


Figure 15 : Performances des fonctions d'appariement [SEC - 87]

Les performances les plus encourageantes sont obtenues par le coefficient de corrélation et le maximum à postériori. Un test sur des images réelles confirme la qualité de ces méthodes. Le seul critère considéré est ici le résultat, mais il est aussi essentiel de prendre en compte le temps de traitement, la place mémoire, et la possibilité de mettre en oeuvre des algorithmes sophistiqués. Il apparaît alors que la mesure la plus favorable est celle du coefficient de corrélation. Nous exploiterons ce résultat pour la réalisation de notre système de poursuite.

I - 3 - Le contexte de l'analyse

Nous avons décrit dans ce qui précède différentes techniques pour l'analyse de scènes variables dans le temps ; le choix peut s'opérer dans cet ensemble en fonction des performances souhaitées et des domaines d'application envisagés.

Des recherches complémentaires ont été menées sur les moyens d'améliorer les conditions de l'analyse :

- par prétraitement des images afin de les rendre plus exploitables (élimination du bruit, réhaussement des contours ...)
- par utilisation d'architectures informatiques spécialisées permettant un calcul plus rapide des fonctions.

Nous décrivons dans cette partie les tendances actuelles concernant ces deux points.

I - 3 - 1 - Les prétraitements

Nous appelons prétraitements les opérations intervenant préalablement à la détection du mobile. Ils sont indépendants de toute hypothèse concernant la nature des mouvements ou les types d'objet présent. Ils ont pour but d'obtenir la représentation optimale de l'image en vue de son traitement, leur apport pour une corrélation est mis en évidence dans [HAL - 80] comme suit :

Notons

$f(x)$: la fonction image originale

$s(x)$: sa représentation

$$[1.36] \quad s(x) = f(x) \cdot h(x)$$

où h est la fonction fenêtre recherchée, permettant d'obtenir une fonction de corrélation r idéale entre f et s ; cela peut s'écrire :

$$[1.37] \quad r(x) = f(x) \odot s(x) = \delta(x)$$

où $\odot$ est le produit de convolution, soit en passant aux transformées de Fourier

$$[1.38] \quad R(u) = \mathcal{F}(r(x)) = 1$$

or

$$[1.39] \quad R(u) = F^*(u) S(u) = F^*(u) \{H(u) * F(u)\}$$

d'où

$$[1.40] \quad h(x) = \frac{1}{f(x)} \times \mathcal{F}^{-1} \left\{ \frac{1}{F^*(u)} \right\}$$

La fenêtre théorique permettant d'obtenir la meilleure corrélation dépend donc explicitement de la fonction image, elle n'est donc pas réalisable de manière systématique. Cependant, il est remarquable que le terme $\mathcal{F}^{-1} \left\{ \frac{1}{F^*(u)} \right\}$ ait ses valeurs les plus grandes aux points de haute fréquence spatiale, donc pour les images, au niveau des contours. Il apparaît donc que la meilleure représentation d'une image, est celle de ses contours ; ceux-ci n'étant pas nécessairement connexes. Le problème est abordé sous un autre angle par Burt [BUR - 82], il compare les résultats de 5 types de corrélation dont une précédée d'un filtrage passe haut de type Laplacien ; et montre que cette dernière est la plus performante lorsque les images comportent du bruit.

Les deux études mettent en évidence l'intérêt d'une extraction de contours précédant la mise en correspondance par corrélation ; de plus une telle opération engendre une réduction du nombre de données à traiter. Le filtre passe haut requis peut être réalisé par de multiples moyens : optiques, électroniques, ou informatiques. Nous décrivons par la suite une technique originale, choisie pour notre système, elle utilise un capteur optoélectronique qui délivre les contours en temps réel.

Pour la corrélation, le prétraitement vise à améliorer une approche pour laquelle il n'est pas indispensable. D'autres algorithmes ne sont exécutables qu'à partir d'images traitées ; c'est le cas pour celui décrit dans [SAM - 84] fondé sur la transformée de Hough, qui nécessite l'établissement de l'image des contours. C'est aussi le cas pour les techniques reposant sur une mise en correspondance entre points d'intérêt. Citons dans ce domaine le système PRIAM [DUB - 84] qui réalise une association arborescente entre points représentatifs des angles de l'image courante et du modèle. Cette approche est inconcevable sans extraction préalable des points anguleux.

I - 3 - 2 - Les architectures spécialisées

L'analyse d'image en général est un grand demandeur de puissance de calcul, la poursuite de cible nécessite de plus que les traitements soient exécutés en un temps minimum ; c'est pourquoi la mise au point d'architectures informatiques appropriées est apparue comme un axe de recherche intéressant.

En premier lieu, le requête de nombreuses opérations mutuellement indépendantes à induit la mise au point de structures parallèles. Celles-ci sont de différents types :

- * structure à matrice de processeurs pour une parallélisation des tâches.

Une architecture de ce type est décrite dans [BRE - 86], elle complète une méthode d'association modèle image en permettant le traitement simultané de plusieurs zones de recherche ; un mobile de 64 x 64 pixels est ainsi trouvé dans une image de 256 x 256 en un temps variant de 0,5 à 1 s.

- * structure de type pipeline pour une parallélisation du flux de données.

Elles sont surtout exploitées pour des opérations de filtrage - détection de contours [SIL - 87], ou pour le calcul d'histogramme et de transformée de Hough [OLS - 87].

Afin d'unifier ces deux approches, les auteurs de [PIE - 88] ébauchent une machine offrant un choix entre ces deux configurations suivant les traitements envisagés. Par ailleurs, une architecture à haute rapidité de calcul est développée [SUG - 86] comme support à un algorithme par soustraction (décrit au paragraphe I.1.2). Elle est constituée d'un bus orienté appelé MFIP, capable de transférer des mémoires multitrames vers des modules multiprocesseurs, elle permet ainsi de traiter simultanément quatre trames d'images de 512 x 512 pixels ; en utilisant une unité rapide de traitement d'image IPU - A.

Une autre tendance est l'implantation matérielle sous forme compacte des principaux traitements de base rendant leur temps d'exécution négligeable. La création la plus spectaculaire dans ce domaine est le système APA 512 (de Vision Security) qui s'applique sur une interface Max Vidéo. L'APA 512 est composé de deux cartes VME utilisant la technologie VLSI, et la technologie porte. Les traitements sont implémentés sur 6 composants de 50 000 portes à base de puces HMOS VLSI. Il est ainsi possible d'opérer des algorithmes de vision, à la vitesse d'acquisition des images. Par exemple, pour une scène, 255 objets peuvent être identifiés par leur position, leur surface, leur périmètre et leurs moments d'ordre 1 et 2. Ceci n'est réalisable que sur des images binaires, le seuillage est assuré par le système Max Vidéo ; un objet est défini par un ensemble de points connexes de même valeur. Un algorithme de poursuite établi sur ce système devrait donc tester les cas de superpositions éventuelles entre mobiles ; au risque d'augmenter les temps d'exécution. Ce système étant récent, il n'est pas encore exploité pour des applications.

Conclusion

Le classement établi dans ce chapitre pour les approches de détection du mouvement par vision artificielle, a permis de mettre en relief l'originalité de chacune d'entre elles. Cependant, il présente un caractère légèrement arbitraire ; en effet certaines applications nécessitent l'exploitation de plusieurs techniques. Citons par exemple celle décrite dans [MIL - 88] qui exploite une initialisation différentielle suivie d'une poursuite par association modèle-image, tout ceci étant exécuté sur une architecture spécialisée.

Nous avons remarqué au cours de notre classement la constance de divers résultats tels l'intérêt des contours en analyse de scène, et l'avantage de la fonction de corrélation en présence de bruit. Ces deux constatations serviront donc de base au développement de notre système de poursuite de cible.

CHAPITRE II

II - CORRELATION PAR RESTRICTION

Cette partie concerne les fondements de notre approche pour l'analyse de scène dynamique. Nous décrivons tout d'abord le système de poursuite de cible dans son ensemble. Puis nous expliquons les choix établis concernant la phase d'écartométrie.

II - 1 - Situation du problème - présentation du système

Dans le cadre des recherches effectuées au Centre de Recherche en Automatique de Nancy sur la commande adaptative de manipulateurs articulés, en vue de la manutention, de l'assemblage ou de l'usinage automatique de pièces, nous participons au développement d'un capteur d'image associant la vision et la commande en robotique. L'objectif poursuivi est la réalisation des opérations de l'analyse de scène statique (reconnaissance des formes, détection de défauts ...) sur des scènes animées. Dans ce but une caméra fixée au bras du robot devra être recalée en permanence sur un mobile cible, grâce aux informations caractérisant le mouvement de ce dernier. Le schéma du système est le suivant :

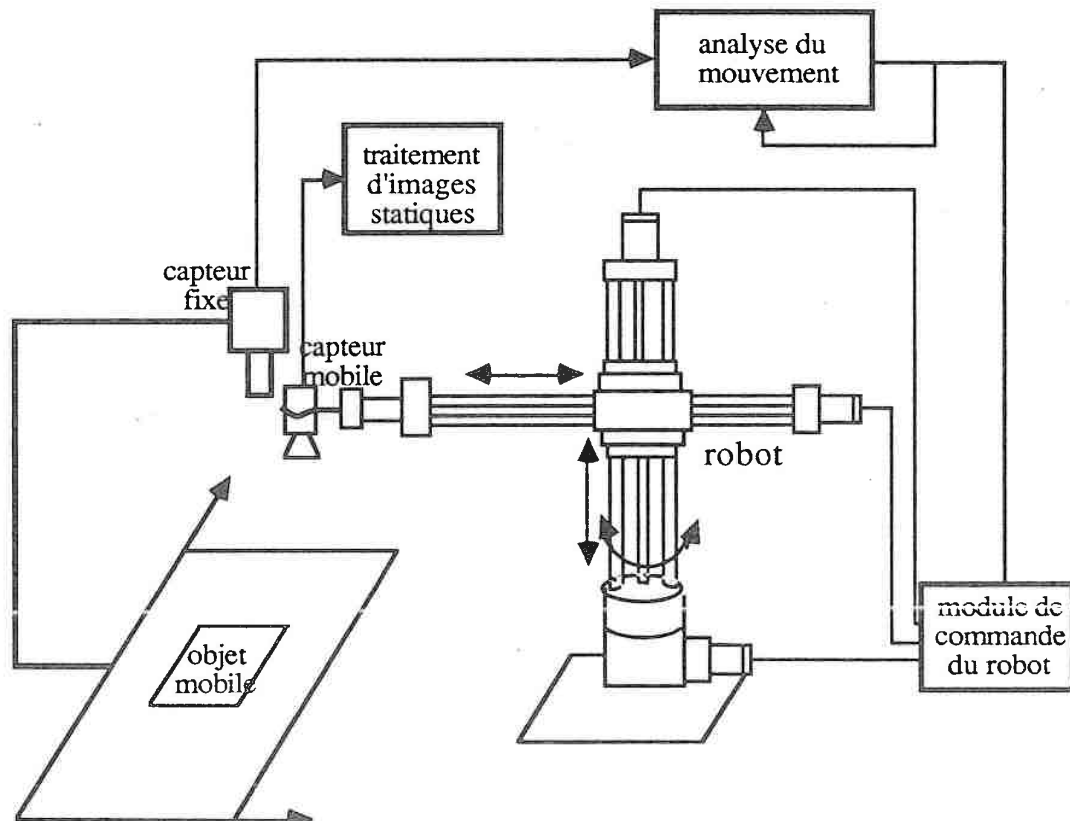


Figure 16: Schéma de principe du système, capteur fixe plus mobile

Nos travaux d'attachent à la perception et à l'interprétation du mouvement en vue de délivrer les données nécessaires à la commande du robot qui fait l'objet d'une étude parallèle. Nous considérons donc uniquement le capteur fixe et la chaîne de traitement qui lui fait suite :

Le système d'analyse de mouvement, objet de notre travail, pourrait trouver d'autres domaines d'application comme par exemple la poursuite de véhicule aérien ; il relève de la vision artificielle et se compose de deux phases :

- (1) : Reconnaissance et localisation de l'objet dans l'image
- (2) : Interprétation des résultats pour une modélisation du mouvement.

Le premier module est exécuté par une méthode correlative entre le modèle de l'objet et une partie de l'image appelée aire de recherche. Le second fait appel aux théories d'estimation statistiques appliquées pour un filtrage adaptatif.

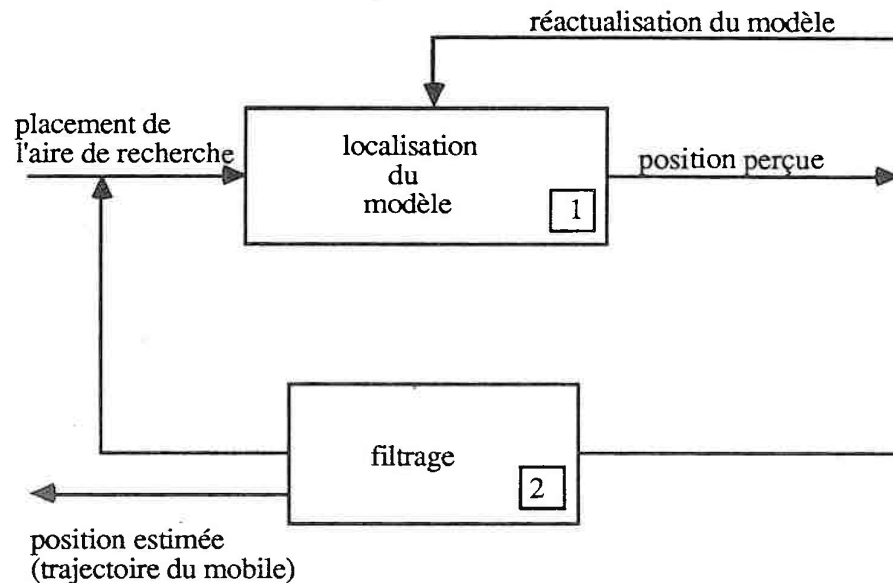


Figure 17 : Description globale de la procédure de détection du mouvement

II - 2 - Choix établis pour l'écartométrie

Nous appelons écartométrie la recherche d'un signal déterministe (modèle) au sein d'un ensemble de données éventuellement bruitées (image).

II - 2 - 1 - Principe

L'étude bibliographique (§ I.2.2) fait ressortir l'attrait des fonctions de corrélation en présence de bruit ; mais ce type de traitement peut être coûteux en temps d'exécution. C'est pourquoi nous proposons une corrélation "restreinte" s'exécutant sur un nombre de point limité. La restriction s'opère sur deux niveaux : tout d'abord une extraction de contours permet la prise en compte des seules zones de l'image de haute fréquence spatiale, puis un seuillage multi-niveaux délivre une représentation quantifiée des images. Ces images discrétisées peuvent être considérées comme une succession d'images binaires pour lesquelles l'opération de corrélation sera remplacée par une association entre point de même niveau.

Le choix des contours est ici effectué d'une part pour améliorer la qualité de la fonction de corrélation et, d'autre part, pour limiter le temps d'exécution. C'est pourquoi l'exploitation d'un extracteur de contours "temps réel" est indispensable ; nous employons dans ce but un capteur original que nous présentons dans le paragraphe suivant.

II - 2 - 2 - La caméra "Net - Flou"

En premier lieu, l'emploi d'un tel capteur est justifié par la recherche théorique du type de filtre à insérer dans une chaîne d'acquisition pour améliorer la dynamique du pic de corrélation. Puis les modalités de la réalisation pratique sont décrites ; enfin des exemples d'images prétraitées sont fournis.

II - 2 - 2 - 1 - Maximisation du pic de corrélation

Considérons le cas d'une image composée d'une tache d'intensité uniforme se détachant sur un fond homogène [WOL - 89]. Supposons qu'entre deux instants t_1 et t_2 la tache a subi une translation de module T_0 . Notons f_1 la trame acquise à l'instant t_1 et f_2 celle de l'instant t_2 .

La fonction de corrélation entre f_1 et f_2 : $\phi_{f_1 f_2}$ aura une allure variable avec la forme de la tache et les niveaux d'intensité respectifs des deux parties de l'image, mais peut être modélisée par une exponentielle [CIC - 87]. Choisissons ce type de représentation :

$$[2.1] \quad \phi_{f_1 f_2}(T) = C e^{-\alpha [T - T_0]^2}$$

T représente le décalage entre f_1 et f_2 .

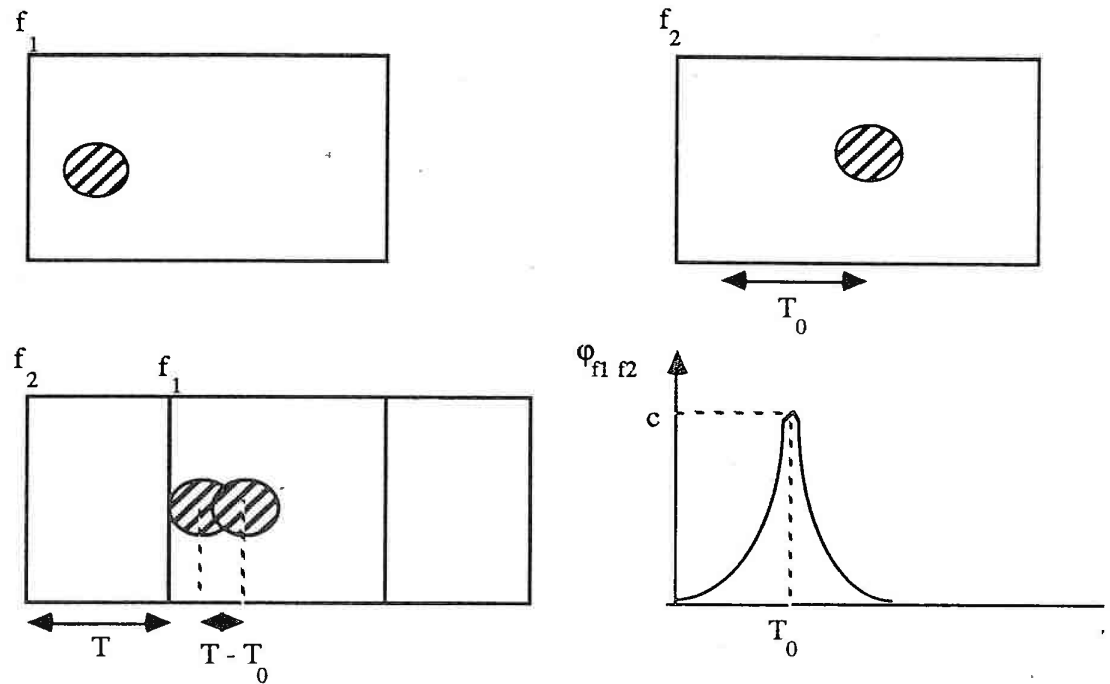


Figure 18 : Exemple de fonction de corrélation

C et α sont deux constantes, elles dépendent de la nature des images et influent sur la dynamique de la fonction de corrélation ; le cas idéal de l'impulsion de Dirac en T_0 résulterait de valeurs infiniment grandes pour C et α .

Le filtre recherché engendrera des images f'_1 et f'_2 telles que leur fonction de corrélation soit :

$$[2.2] \quad \phi_{f'_1 f'_2}(T) = C' e^{-\alpha' (T - T_0)^2}$$

avec $C' \gg C$ et $\alpha' \gg \alpha$

Notons H la fonction de transfert d'un tel filtre ;

$$[2.3] \quad f'_1 = H * f_1 \quad \text{et} \quad f'_2 = H * f_2$$

Soit en fréquence

$$[2.4] \quad f'_1(\omega) = H(\omega) f_1(\omega) \quad f'_2(\omega) = H(\omega) f_2(\omega)$$

donc

$$[2.5] \quad \phi_{f'_1 f'_2}(\omega) = H^2(\omega) \phi_{f_1 f_2}(\omega)$$

en prenant le modèle exponentiel de la formule [2.3]

$$[2.6] \quad \phi_{f1 f2}(\omega) = \sqrt{\frac{\pi}{\alpha}} C e^{-j\omega T_0} e^{-\omega^2/4\alpha}$$

$$[2.7] \quad \phi_{f1 f2}(\omega) = \sqrt{\frac{\pi}{\alpha'}} C' e^{-j\omega T_0} e^{-\omega^2/4\alpha'}$$

ainsi

$$[2.8] \quad H^2(\omega) = \frac{C' \sqrt{\alpha}}{C \sqrt{\alpha'}} e^{\frac{\omega^2}{4 \left(\frac{1}{\alpha} - \frac{1}{\alpha'} \right)}}$$

En imposant la même coénergie à $\phi_{f1 f2}$ et $\phi_{f1 f2}$, les rapports $\frac{c}{\sqrt{\alpha}}$ et $\frac{c'}{\sqrt{\alpha'}}$ sont égaux.

il vient.

$$[2.9] \quad H(\omega) = e^{\frac{\omega^2}{2\omega_1^2}} \quad \text{où} \quad \omega_1 = \frac{8 \alpha \alpha'}{\alpha - \alpha'}$$

H est un filtre passe haut non réalisable physiquement car il nécessiterait un gain infini pour les grandes valeurs de ω . De plus un tel phénomène ne serait pas souhaitable dans la mesure où il amplifierait démesurément toute forme de parasites hautes fréquences, noyant le signal déterministe.

Le problème est ramené à la synthèse d'un filtre dont l'effet se rapproche de celui de H pour les basses fréquences. Utilisons pour cela les notations générales de filtrage d'image [GOU - 72]. Soient z_0 et z_1 les amplitudes complexes représentant les champs entrée et sortie d'un système décrit par sa réponse impulsionnelle h , ces grandeurs sont liées par :

$$z_1(x_i) = h * z_0(x_i)$$

et dans le domaine fréquentiel

$$z_1(\omega) = H(\omega) z_0(\omega)$$

Si H est le filtre décrit précédemment, z_1 sera l'image représentant les hautes fréquences spatiales de z_0 . Il est alors possible de définir K par :

$$z_0(\omega) = K(\omega) z_1(\omega)$$

avec

$$K(\omega) = \frac{1}{H(\omega)} = e^{-\frac{\omega^2}{\omega_1^2}}$$

K est le filtre passe bas inverse de H , pratiquement il se rapproche d'une dégradation appliquée à l'image z_1 par la défocalisation de l'objectif de réception (effet de flou).

Représentons les fonctions H , K et $1 - K$:

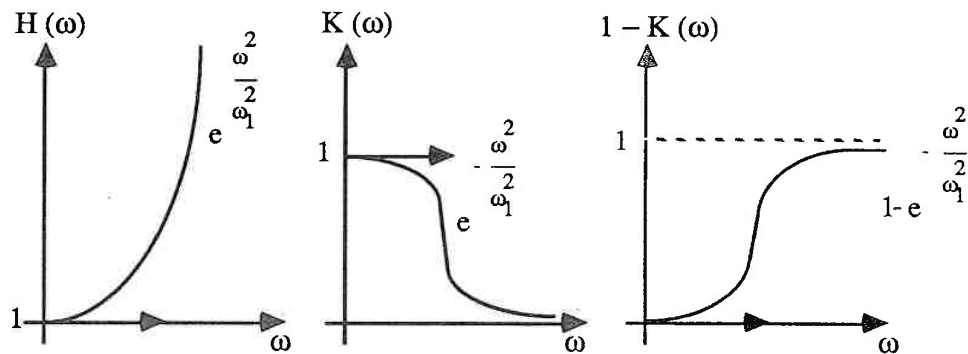


Figure 19 : représentations de $H(\omega)$, $K(\omega)$ et $1 - K(\omega)$

La fonction $(1 - K)$ qui pour les basses fréquences se rapproche de H peut donc être réalisée par la différence entre une image nette et cette même image légèrement défocalisée ; cet ensemble d'opération constituant un prétraitement réalisable pour l'optimisation de la fonction de corrélation.

Le filtrage ainsi décrit est un cas particulier de la technique du masque flou [PRA - 78] exploitée pour le réhaussement d'images photographiques à faible contraste. Il s'agit donc d'un extracteur de contours ; ce résultat est conforme aux hypothèses [HAL - 80] décrites au paragraphe I.3.1.

II - 2 - 2 - 2 - Différents extracteurs de contours

La recherche des contours peut être un préalable à de nombreuses opérations d'analyse d'image (notamment en reconnaissance des formes) ; elle constitue donc un pôle d'étude au développement précoce (Roberts 1965) et vaste [ROS - 82] [HAL - 79]. Les opérateurs classiques de détection reposent sur une approximation discrète des variations spatiales de la luminance. Ainsi les dérivées premières sont considérées, pour les méthodes de gradient mettant en oeuvre plusieurs masques de convolution, par exemple citons : Roberts - Kirsh - Sobel - Prewitt ... ; ou les dérivées secondes sont approximées par les filtres laplaciens.

Nous ne rappellerons pas le détail de ces techniques [SAM - 84], elles sont très répandues, et bien qu'elles nécessitent beaucoup de calcul, leur temps d'exécution peut être infime grâce au cablage de certaines opérations. Cependant, un filtrage complémentaire est souvent mis en place afin d'éliminer les points non significatifs, d'affiner les contours, et de les rendre continus. C'est le cas pour les deux méthodes comparées dans [BOL - 88]. La première (opérateur de Castan - Shen) est constituée d'un filtrage passe bas, d'un opérateur Laplacien puis d'une vérification par gradient adaptatif, son temps d'exécution peut être long mais les contours délivrés sont fins. La seconde combine un filtre médian et un opérateur de Roberts, elle engendre des contours plus épais mais continus. Ces deux techniques comparées à un opérateur de Sobel montrent une plus grande immunité au bruit.

Remarque :

Ces opérateurs par masquage ne peuvent jamais être totalement isotropes ; en effet, qu'ils considèrent 4 ou 8 directions pour les proches voisins d'un pixel, il s'agit toujours d'une approximation des variations réelles dans l'image analogique. Les résultats sont de ce fait plus ou moins fortement biaisés.

Une autre tendance, en matière de détection, est représentée par la technique d'accroissement de région. Un point donné est associé à ses voisins par un critère de similarité et le domaine couvert s'étend jusqu'à la rencontre d'un point dissemblable ; les transitions entre les parties ainsi déterminées constituent les contours. Ce type de méthode peut être amélioré par un test de vraisemblance ; par exemple dans [COO - 79] l'auteur modélise l'objet par un processus de Markov, et les parasites par un bruit blanc gaussien, ainsi chaque hypothèse de contours est validée ou non par un critère probabiliste fonction de toute l'image. Si ce type de procédure offre des résultats intéressants, sa complexité n'autorise pas une application dans le domaine qui nous occupe.

Nous nous sommes donc attachés à utiliser un capteur alliant des fonctions optiques et électroniques afin de délivrer très rapidement une représentation des contours. L'avantage d'une technique optique pour rechercher des transitions dans l'image est qu'elle présente d'une part une totale isotropie, et d'autre part un temps d'exécution négligeable.

II - 2 - 2 - 3 - Principe et réalisation du capteur "N-F"

La mise en oeuvre d'un tel système requiert l'élaboration de deux images, l'une nette, l'autre floue ; toutes deux issues du même objet, la seconde subissant un passage par un filtre passe bas (défocalisation de caméra).

Ensuite une soustraction électronique entre ces deux images génère le tracé des contours.

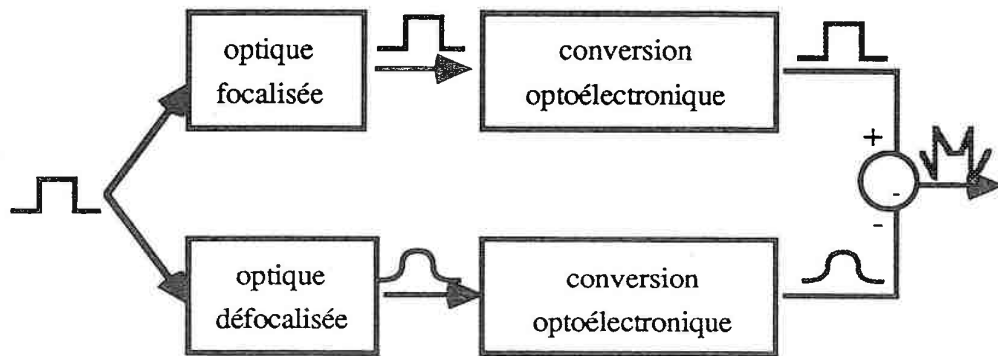


Figure 20 : Principe du filtre passe haut optique

Le diagramme représentant la réalisation pratique d'un tel capteur est donc le suivant :

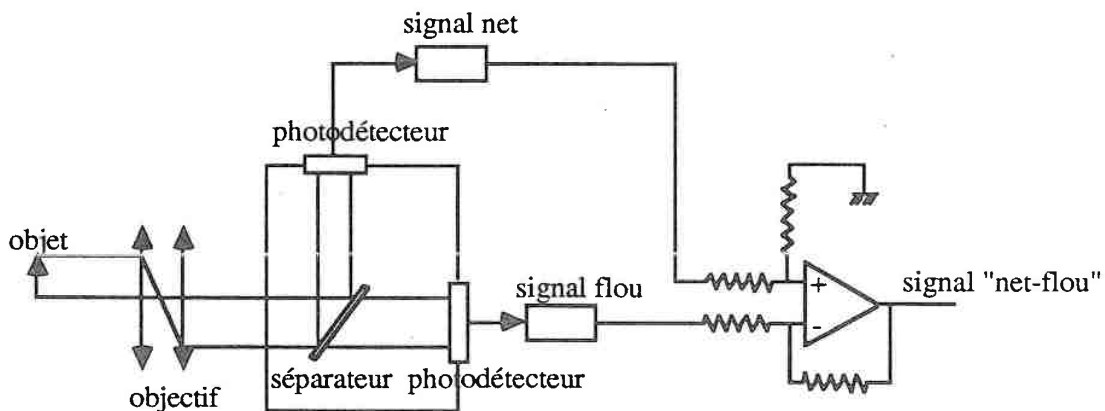


Figure 21 : Développement du filtre passe haut optique

Le système est construit à partir de capteurs classiques d'analyse d'image. Le signal lumineux provenant de l'objet passe à travers un séparateur afin d'être envoyé simultanément sur deux photodétecteurs ; l'un recevant l'image nette, l'autre l'image floue. Deux matrices photosensibles assurent la conversion optique-électrique ; puis une soustraction électronique délivre le signal numérisé représentant l'image des contours.

Les données issues de ce capteur sont donc prêtes à être stockées et traitées par un ordinateur. Les éléments optiques du système sont fixés sur support mécanique de position ajustable comme l'indique la figure ci-dessous.

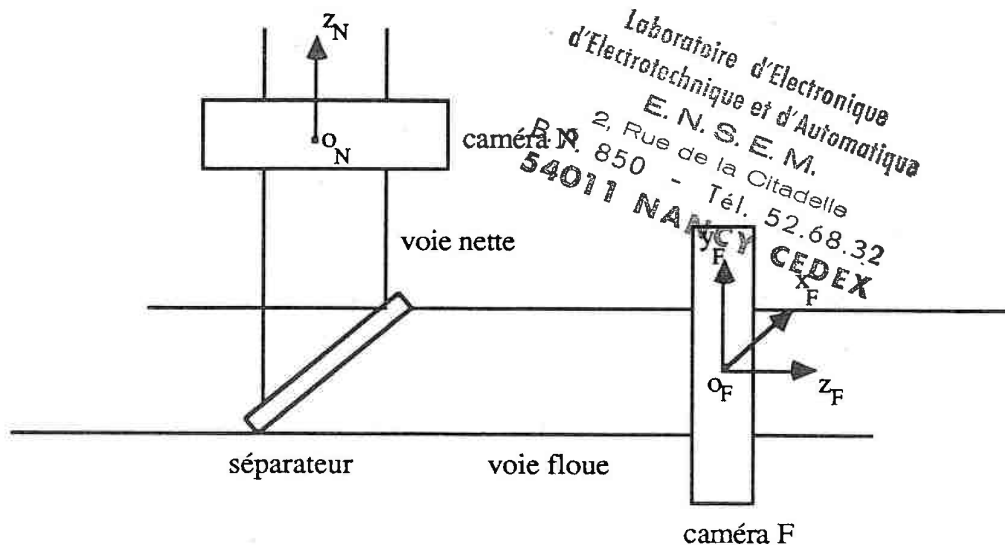


Figure 22 : Axes de translation des capteurs N et F

Les réglages sont assurés par translation des caméras de la manière suivante :

- mise au point de l'image nette : translation de N suivant $\vec{O_N z_N}$.
- défocalisation de l'image floue : translation de F suivant $\vec{O_F z_F}$
- calage relatif des images nettes et floues : translation de F suivant $\vec{O_F x_F}$ et $\vec{O_F y_F}$

Le circuit électronique mis en oeuvre en aval de ce dispositif permet de visualiser successivement les images, nette, floue ou nette moins floue afin de faciliter la mise au point.

Le principe de ce capteur, appliqué à l'automatisation du contrôle magnétoscopique, a fait l'objet d'un travail antérieur [CIC - 84] développé au sein de la société SOPELEM. La réalisation du prototype est donc le fruit d'une collaboration entre le CRAN et cette entreprise.

Le dispositif optoélectronique réalisé est le suivant :

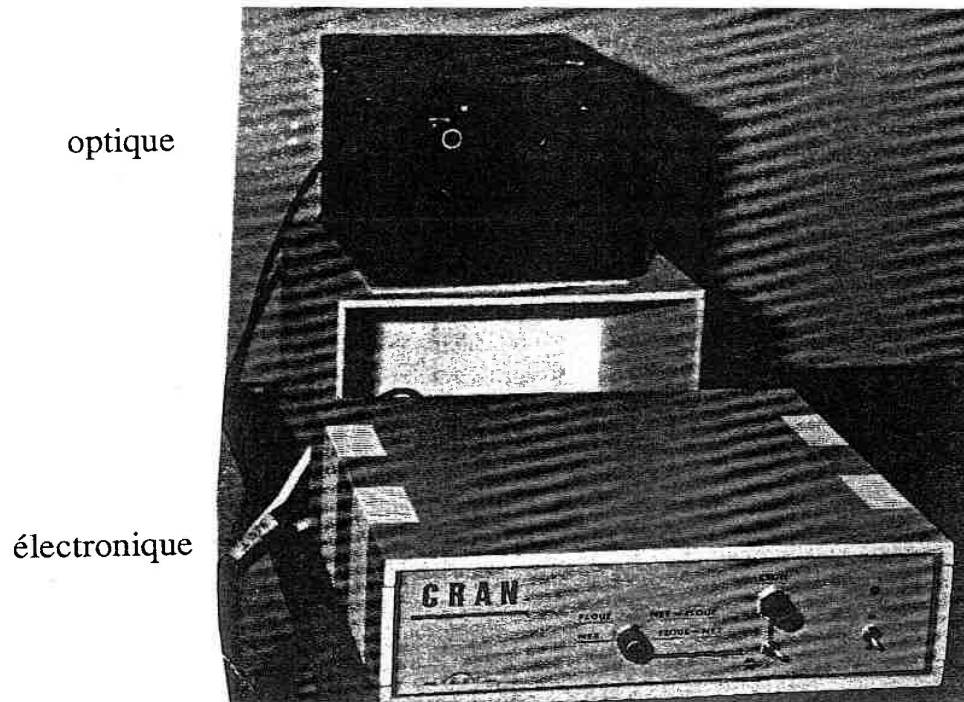


Figure 23 : Photo du prototype

II - 2 - 2 - 4 - Exemples d'images traitées - Effet sur la fonction de corrélation

Le rôle du filtre, ainsi réalisé, est la sélection des hautes fréquences de l'image. la dynamique des fonctions de corrélation doit par conséquent être augmentée ; nous visualiserons leurs allures avec et sans prétraitement.

Nous expliciterons par la suite l'algorithme d'appariement mis en oeuvre pour le système de poursuite de cible. La simulation est réalisée sur des exemples à partir d'un modèle d'objet R (de 50 sur 50 pixels) recherchée dans une image I (de 100 sur 100 pixels). Les courbes représentent les fonctions monodimensionnelles, pour $\tau = 0$ à 50 répondant à l'équation :

$$[2.10] \quad \phi(\tau) = \sum_{i=1}^{50} \sum_{j=1}^{50} I(i + \tau, j) \cdot R(i, j)$$

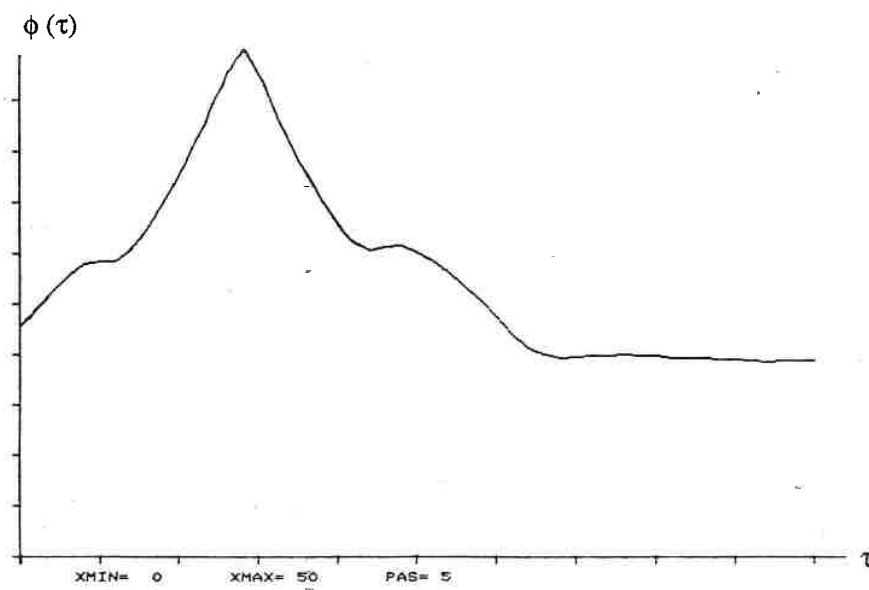
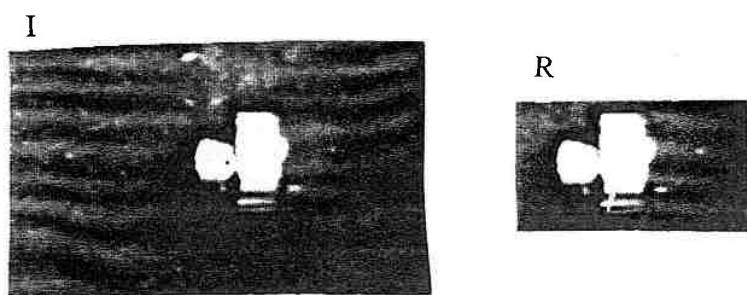


Figure 24 : pièce
mécanique sans
prétraitement

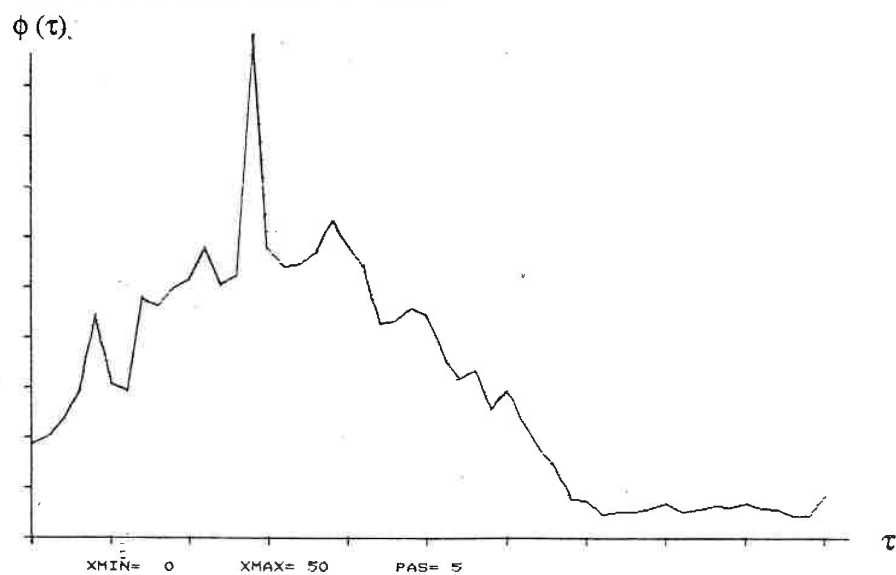
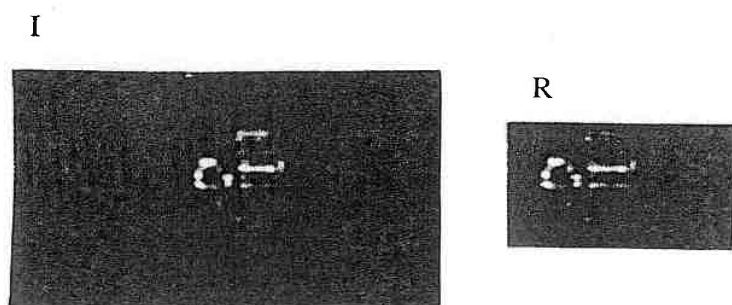
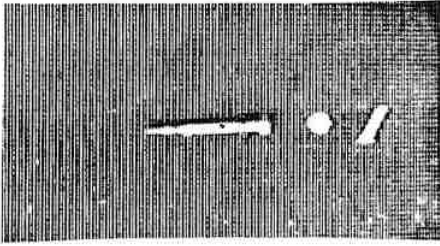


Figure 25 : pièce
mécanique avec
prétraitement

I



R

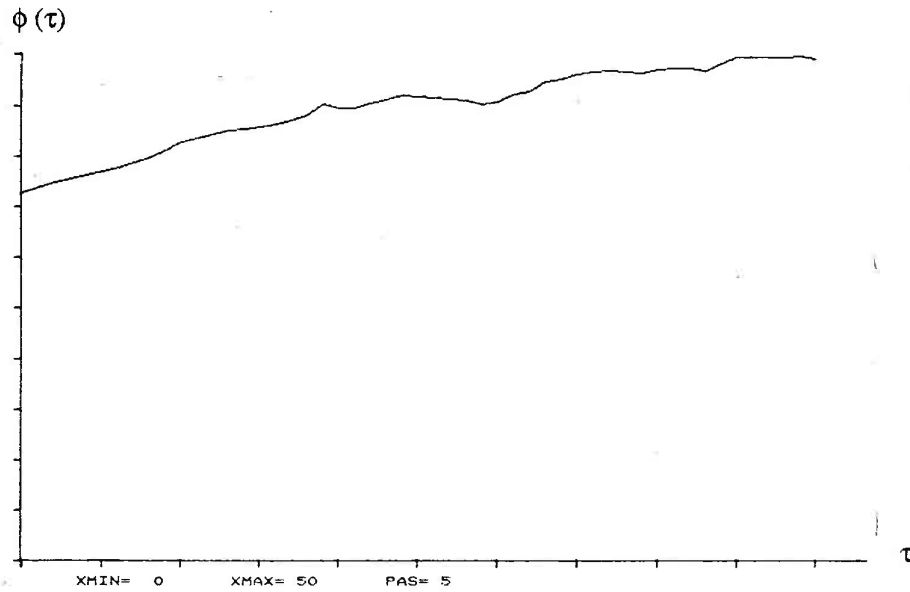
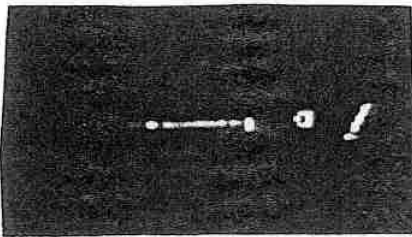


Figure 26 : vis sur
plan de travail sans
prétraitement

I



R

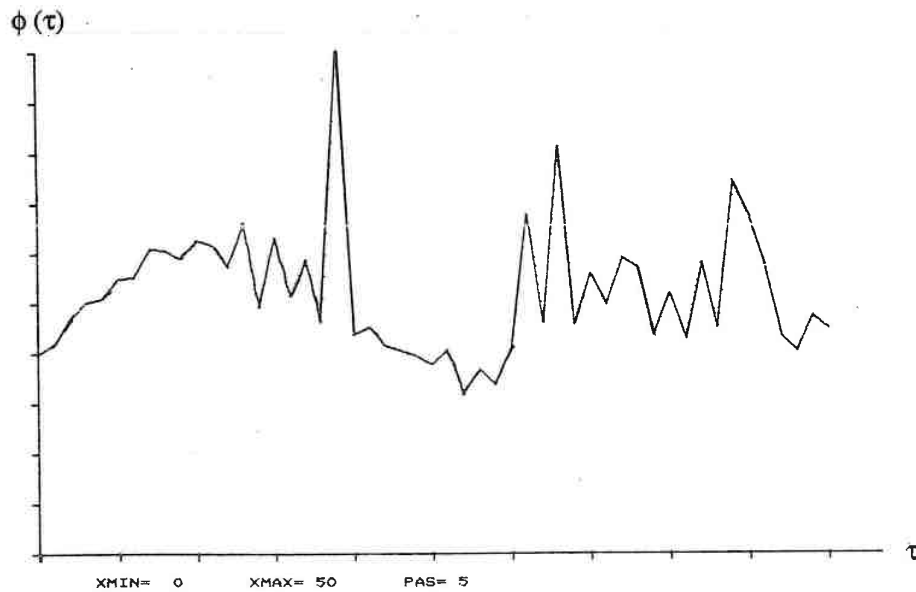
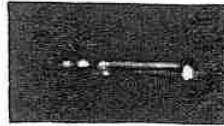
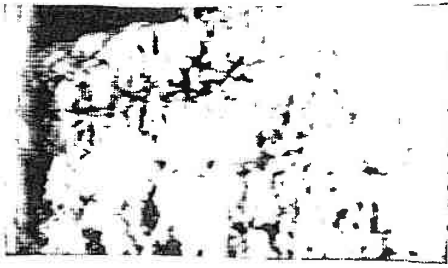


Figure 27 : vis sur
plan de travail avec
prétraitement

I

53



R

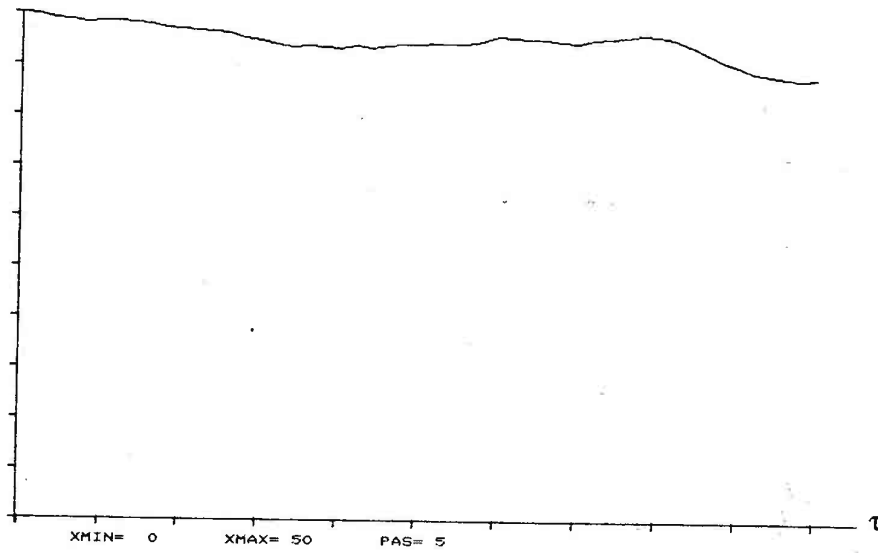
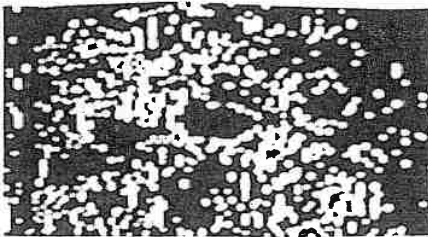
 $\phi(\tau)$ 

Figure 28 : véhicule
dans un paysage sans
prétraitement

I



R

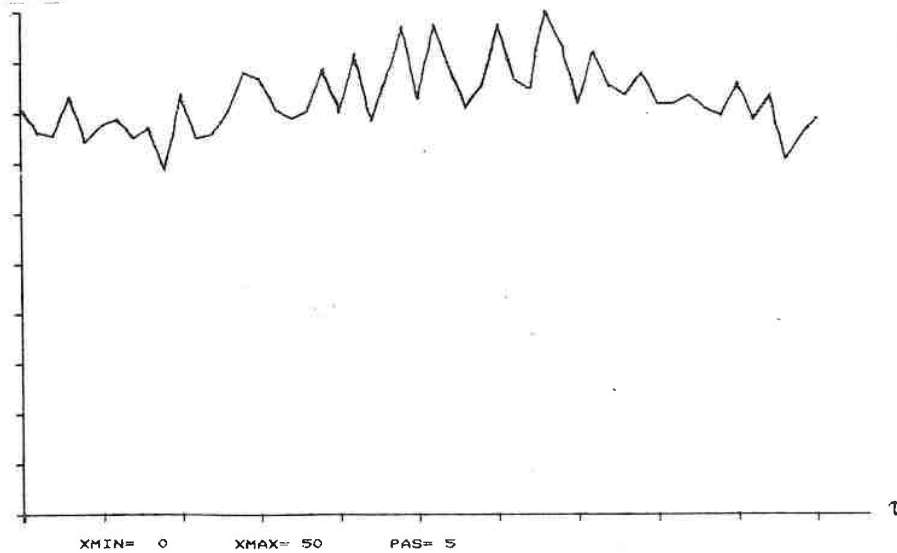
 $\phi(\tau)$ 

Figure 29 : véhicule
dans un paysage avec
prétraitement

Le premier exemple est un cas simple où le filtrage donne les résultats escomptés lors de sa mise au point : diminution du nombre de point révélant l'objet et accentuation du pic de corrélation. Notons que ce dernier est bien localisé, dans le cas filtré, pour toutes les images considérées. Pour le deuxième exemple, les objets secondaires et la différence d'éclairement empêchent la détection du pic qui devient possible après l'extraction de contours. Enfin dans le troisième exemple la fonction de corrélation ne représente rien dans le cas non filtré. La seconde courbe montre une dynamique plus marquée et permet la localisation du véhicule dans le paysage.

Toutefois les deux derniers exemples révèlent un inconvénient du filtre passe haut : il augmente la dynamique du bruit. Pour les situations considérées, ce phénomène n'est pas rédibitoire, mais nous remarquons sur la dernière courbe un écart presque nul entre la valeur du pic réel et celles des pics secondaires dus au bruit. Nous tenterons, de réduire les inconvénients de ce phénomène par le mode de corrélation mis en place.

II - 2 - 3 - Mise en oeuvre de la corrélation

L'objectif est la localisation en un minimum de temps d'un modèle, représentant un objet, dans une aire de recherche, le contenant. Les opérations que nous allons décrire sont réalisées à partir d'images des contours, nous considérerons donc implicitement dans ce qui suit que le prétraitement est exécuté.

II - 2 - 3 - 1 - Principe de l'algorithme

Dans un premier temps nous décrivons notre approche sur des signaux monodimensionnels continus.

Soient $w(x)$ et $f(x)$ les réalisations des processus aléatoires représentant respectivement le modèle et le domaine de recherche ; f et w sont entachés de bruits limitant leur intelligibilité. Notre technique repose sur une succession de corrélations entre les signaux binaires obtenus par seuillage à différents niveaux (figure 30).

Une première sélection au niveau s_1 permet de déterminer les fonctions binaires f_{s_1} et w_{s_1} en affectant la valeur 1 aux points dépassant le niveau s_1 et 0 aux autres.

La multiplication de ces deux représentations pour différents décalages de w détermine la première fonction de corrélation :

$$[2.11] \quad \phi_{s_1}(\tau) = \sum_x f_{s_1}(x) w_{s_1}(x + \tau)$$

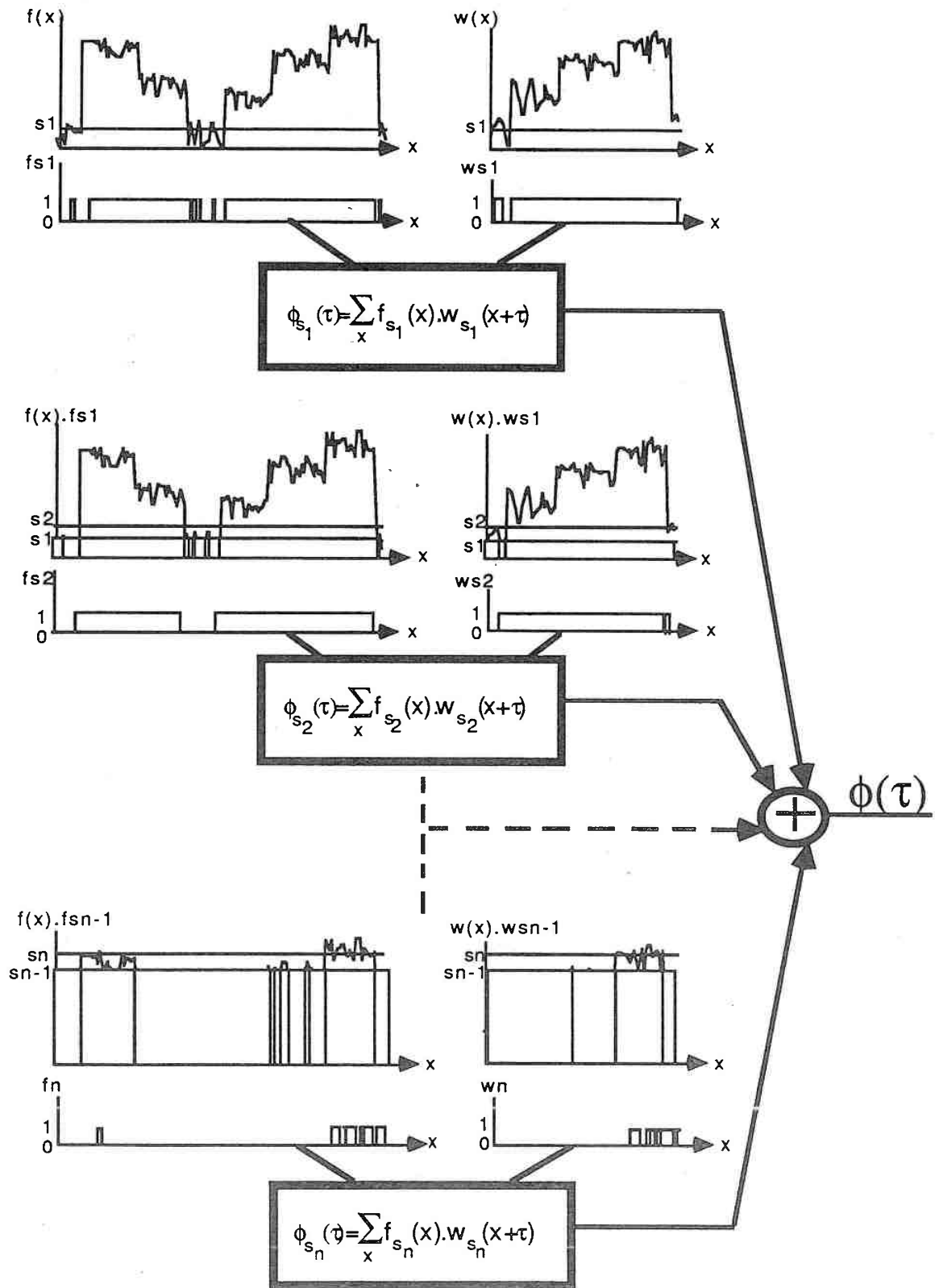


Figure 30 : Principe de la corrélation restreinte

Dans le cas simple de signaux fortement contrastés non bruités, la seule recherche du maximum de ϕ_{s1} , indiquerait l'emplacement de w dans f .

En réalité, il paraît indispensable de reproduire cette opération, tant pour limiter l'arbitraire des résultats obtenus au premier niveau, que pour tenir compte des informations déterministes apparaissant aux niveaux supérieurs.

La procédure décrite est donc appliquée de nouveau pour $s_2 > s_1$, à partir des seules zones de f et w ayant satisfait à la condition précédente ; elles sont notées $f \cdot f_{s1}$ et $w \cdot w_{s1}$ car issues de f et w masquées par f_{s1} et w_{s1} .

Une seconde fonction de corrélation est alors obtenue :

$$[2.12] \quad \phi_{s2}(\tau) = \sum_x f_{s2}(x) \cdot w_{s2}(x + \tau)$$

Cette succession de traitements est répétée jusqu'au $n^{\text{ième}}$ seuillage de niveau s_n , sur les signaux $f \cdot f_{sn-1}$, engendrant la fonction de corrélation $\phi_{sn}(\tau)$.

Puis par addition de chacun des résultats précédents nous calculons :

$$[2.13] \quad \phi(\tau) = \sum_{i=1}^n \phi_{si}(\tau)$$

Le cumul des fonctions de corrélation binaires successives, permet de déterminer le maximum d'appariement entre modèle et image à partir des informations déduites de chaque seuillage. Si les niveaux sont choisis de manière appropriée en fonction des caractéristiques de l'image, les erreurs de quantification pourront être minimisées. Nous détaillerons ce point par la suite, illustrons tout d'abord la technique sur un exemple simple.

Les images étant en fait représentées par des signaux discrets, la mise en oeuvre du principe décrit nécessite quelques transformations. Présentons le mode opératoire sur un exemple élémentaire. Considérons f et w composés respectivement de 9 et 4 points.

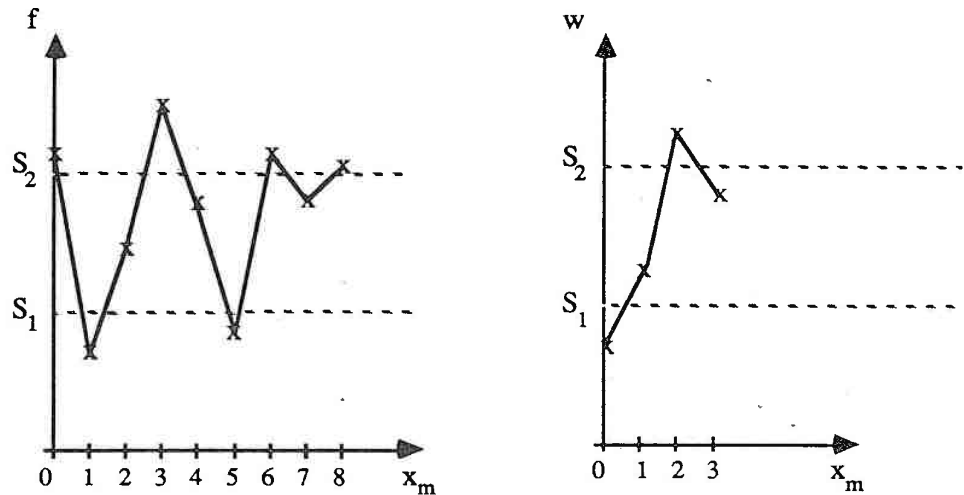


Figure 31 :

Les fonctions binaires définies précédemment sont maintenant remplacées par des tableaux ne contenant que les points affectés à 1.

Le premier seuillage permet d'établir :

$$f_{s1} = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 4 \\ 6 \\ 7 \\ 8 \end{bmatrix} \quad w_{s1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

La corrélation entre f_{s1} et w_{s1} pour les 5 emplacements possibles de w est mesurée par $\phi_{s1}(d)$ où d représente le décalage entre f et w .

Calcul de $\phi_{s1}(d)$:

Soient $f_{s1}(i)$: le i ème élément de f_{s1}

$w_{s1}(i)$: le i ème élément de w_{s1}

n : le nombre de points dans f_{s1}

Alors

$$[2.14] \quad \phi_{s1}(d) = \sum_{i=1}^m a_i(d)$$

$$\text{avec} \begin{cases} a_i(d) = 1 \text{ si } \exists k \in \mathbb{Z} / f_{s1}(i) = w_{s1}(i+k) + d \\ a_i(d) = 0 \text{ sinon} \end{cases}$$

Ainsi, la multiplication du cas continu est remplacée par la mise en correspondance des adresses des points dépassant le seuil.

Pour l'exemple ci dessus, nous obtenons :

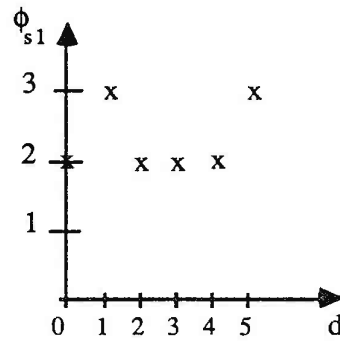


Figure 32 : Valeurs de $\phi(d)$ au niveau s_1

L'information obtenue ne permet pas de choisir entre les points 1 et 5 pour la meilleure localisation de w .

Un deuxième seuillage est effectué au niveau s_2 sur les seuls points de f_{s1} et w_{s1} .

Il vient

$$f_{s2} = \begin{bmatrix} 0 \\ 3 \\ 6 \\ 8 \end{bmatrix} \quad w_{s2} = [2]$$

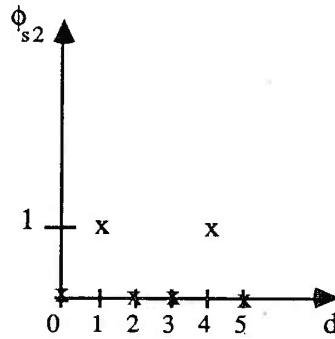


Figure 33 : Valeurs de $\phi(d)$ au niveau s_2

Pour ce second seuillage deux positions sont envisageables $\{1,4\}$: Mais en considérant ce résultat comme un complément d'information au précédent, nous obtenons :

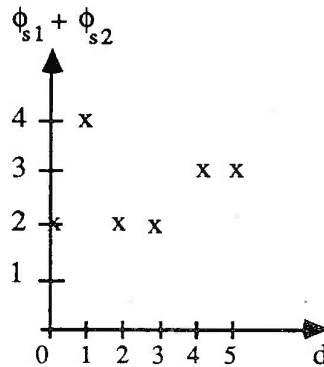


Figure 34 : valeurs de $\phi_{s1} + \phi_{s2}(d)$

A partir de ce tableau, la seule place possible pour w est : $\{1\}$

Cet exemple met en valeur l'importance du choix des seuils : un niveau trop bas induira des informations parasites, un niveau trop haut exclura le signal recherché. Afin de rendre nos choix les moins arbitraires possibles, nous nous sommes reportés aux techniques utilisées en traitement du signal pour la détection d'un signal déterministe dans un univers bruité.

II - 2 - 3 - 2 - Problème de seuillage - critère de décision

Le problème de la décision binaire peut être formulé de la manière suivante : une source d'information produit deux signaux possibles s_0 et s_1 , incompatibles, et possédant des probabilités d'apparition respectives q et p . L'observation est exécutée par la mesure d'une variable x composée du signal et d'un facteur de bruit.

Une valeur observée x_0 appartient à l'une des deux distributions incompatibles : $W(x/s_0)$ ou $W(x/s_1)$, il s'agit alors d'opérer un choix entre les deux hypothèses :

γ_0 : s_0 présent

γ_1 : s_1 présent

Les stratégies de décision reposent sur un critère permettant de limiter les domaines respectifs de γ_0 et γ_1 .

Dans le cas monodimensionnel, ces deux domaines sont définis par une valeur limite s .

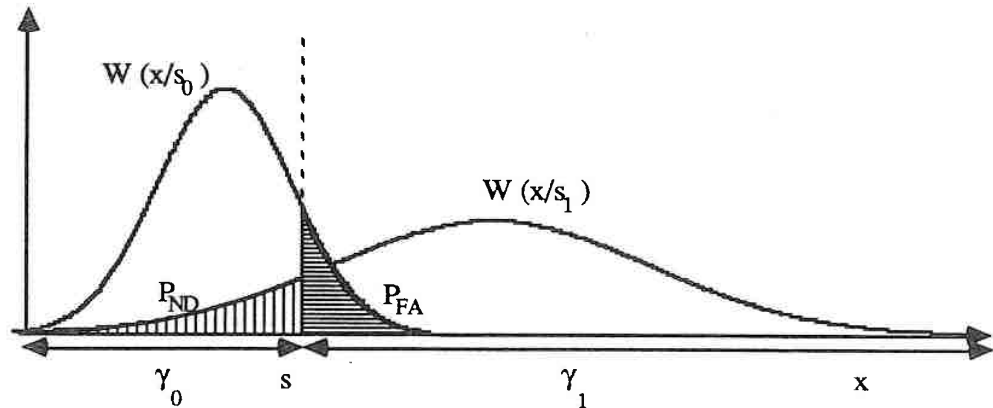


Figure 35 : distributions de probabilités $W(x/s_0)$ et $W(x/s_1)$

Les écueils de la décision sont représentés par la fausse alarme (γ_1 choisie et s_0 vraie) et la non détection (γ_0 choisie et s_1 vraie) dont les probabilités sont :

$$[2.15] \quad P_{FA} = P(\gamma_1/s_0) = \int_s^{\infty} W(x/s_0) dx$$

$$[2.16] \quad P_{ND}(\gamma_0/s_1) = \int_{-\infty}^s W(x/s_1) dx$$

A titre d'exemple, énonçons quelques critères :

- critère du maximum de probabilité à postériori. Le seuil s est choisi de manière à satisfaire l'équation :

$$[2.17] \quad \frac{p}{q} = \frac{W(s/s_0)}{W(s/s_1)}$$

et le critère de décision s'exprime par :

$$[2.18] \quad \gamma_1 = \left\{ x / \frac{p W(x/s_1)}{q W(x/s_0)} > 1 \right\} \text{ et } \gamma_0 = \left\{ x / \frac{p W(x/s_1)}{q W(x/s_0)} < 1 \right\}$$

- critère du maximum de vraisemblance. C'est un cas particulier du précédent en supposant implicitement que $p = q = 1/2$ (même probabilité d'apparition de s_0 et s_1).

$$[2.19] \quad \gamma_1 = \left\{ x / \frac{W(x/s_1)}{W(x/s_0)} > 1 \right\} \text{ et } \gamma_0 = \left\{ x / \frac{W(x/s_1)}{W(x/s_0)} < 1 \right\}$$

le seuil s est défini par

$$[2.20] \quad \frac{W(s/s_0)}{W(s/s_1)} = 1$$

dans le cas du schéma (fig. 35) s se situera au niveau où les courbes $W(x/s_0)$ et $W(x/s_1)$ se coupent.

Il existe d'autres critères (Bayes - minimax ...) qui nécessitent aussi les connaissances des probabilités p et q [COU - 84] ; dans le cas où celles-ci sont indéterminées à priori, il est intéressant d'utiliser le critère de Neyman Pearson.

L'objectif suivi pour la mise en place de ce critère est la minimisation simultanée de P_{FA} et P_{ND} . Malheureusement, la diminution de l'une provoque généralement l'augmentation de l'autre. La procédure retenue consiste à fixer arbitrairement la probabilité de fausse alarme (P_{FA}) à une valeur admissible α , puis à minimiser sous cette contrainte, et dans la mesure du possible, la probabilité de non détection P_{ND} .

Dans le cas général, la méthode utilisée est celle des multiplicateurs de Lagrange :

$$[2.21] \quad L(\mu) = P_{ND} + \mu (P_{FA} - \alpha)$$

$L(\mu)$ sera rendue minimale pour [LIF - 81] :

$$[2.22] \quad \mu = \frac{W(x/s_1)}{W(x/s_0)}$$

et la décision sera donc

$$[2.23] \quad \gamma_1 = \left\{ x / \frac{W(x/s_1)}{W(x/s_0)} > \mu \right\} \text{ et } \gamma_0 = \left\{ x / \frac{W(x/s_1)}{W(x/s_0)} < \mu \right\}$$

Dans le cas plus simple d'une observation monodimensionnelle et d'un seuil de décision unique, le choix de α détermine la valeur s par :

$$[2.24] \quad \alpha = \int_s^{\infty} W(x/s_0) dx$$

connaissant s , la probabilité de non détection est donnée par

$$[2.25] \quad P_{ND} = \int_{-\infty}^s W(x/s_1) dx$$

cette grandeur n'est pas toujours calculable car la loi de $W(x/s_1)$ peut être inconnue.

La connaissance de la loi $W(x/s_0)$ est la seule donnée nécessaire à l'application de ce critère ; une condition supplémentaire étant que le signal s_1 n'apparaisse pas souvent.

Pour les images que nous devons considérer, s_1 représentera la présence de contours et s_0 la réception du fond ; s_1 a donc une probabilité d'apparition beaucoup plus faible que s_0 . Nous supposons le signal s_0 (représentant le bruit) connu et régi par une loi normale de paramètres m et σ . Cette hypothèse est souvent vérifiée en pratique par application du théorème central limite ; si le bruit est composé de n variables aléatoires indépendantes de distributions quelconques $p_1, p_2, \dots, p_n$, la densité de probabilité totale de p tend asymptotiquement vers une loi normale à la seule condition que

$$\sigma_1^2, \dots, \sigma_n^2 \ll \sum_{i=1}^n \sigma_i^2$$

La modélisation des images permettant de déterminer m et σ sera traitée au Chapitre III (justification du modèle gaussien).

Sur ces bases et par l'application du critère :

$$[2.26] \quad P_{FA} = \int_s^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

Le choix de s sera donc fonction de la probabilité de fausse alarme admissible.

Afin que le seuil minimal soit en relation simple avec les coefficients caractéristiques du bruit, nous choisissons

$$s = m + \sigma$$

soit

$$P_{FA} = \int_{m+\sigma}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dX = 16 \%$$

Cette valeur est acceptable pour une première mesure, sachant que des estimations complémentaires confirmeront ou infirmeront les résultats obtenus à ce niveau.

II - 2 - 3 - 3 - La Pondération

Le type de corrélation exploitée (binaire multi-niveaux) est équivalent à une corrélation classique opérant sur des signaux quantifiés, les niveaux de quantification étant alors les seuils successifs. Cette technique est choisie en raison de l'économie en temps qu'elle prodigue, mais le nombre restreint de valeurs prises en compte engendre une perte d'information par rapport à leur représentation continue. Nous tentons de compenser cette perte en donnant plus d'importance aux parties susceptibles de contenir le signal déterministe.

Dans ce but [CIC - 87], un poids est affecté à chaque niveau de seuil, il est proportionnel à la probabilité de l'événement " γ_0/s_0 " (cf § II.2.3.2), soit à $1 - P_{FA}$. Ainsi plus la probabilité de fausse alarme est faible, plus l'information détectée est importante.

Le seuil minimal ayant été choisi à la valeur $m + \sigma$, $(1 - P_{FA})$ varie de 0,84 à 1, un tel facteur de pondération serait insignifiant. Pour donner une influence à cette opération sans éloigner trop fortement la courbe de corrélation de celle obtenue par le calcul direct, les poids sont pris entre 0,5 et 1. Les valeurs p_k associées aux niveaux s_k sont donc issues d'une transformation affine de $(1 - P_{FA})$.

Soit :

$$[2.27] \quad p_k = 2,9 (1 - P_{FA}) - 1,9$$

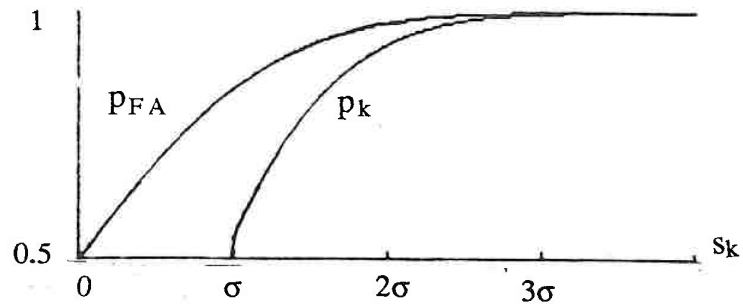


Figure 36 : Représentation de p_k

La fonction de pondération pour un bruit blanc gaussien s'écrit :

$$s_k \rightarrow p_k = 2,9 \left(1 - \int_{s_k}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \right) - 1,9$$

Remarque : Si les signaux aléatoires sont régis par une autre loi, il sera possible de calculer les facteurs de pondération sur le même modèle.

L'effet de la pondération est visualisé, sur des signaux monodimensionnels bruités :

Exemple 1 : Rapport signal sur bruit de f et $w = 4$

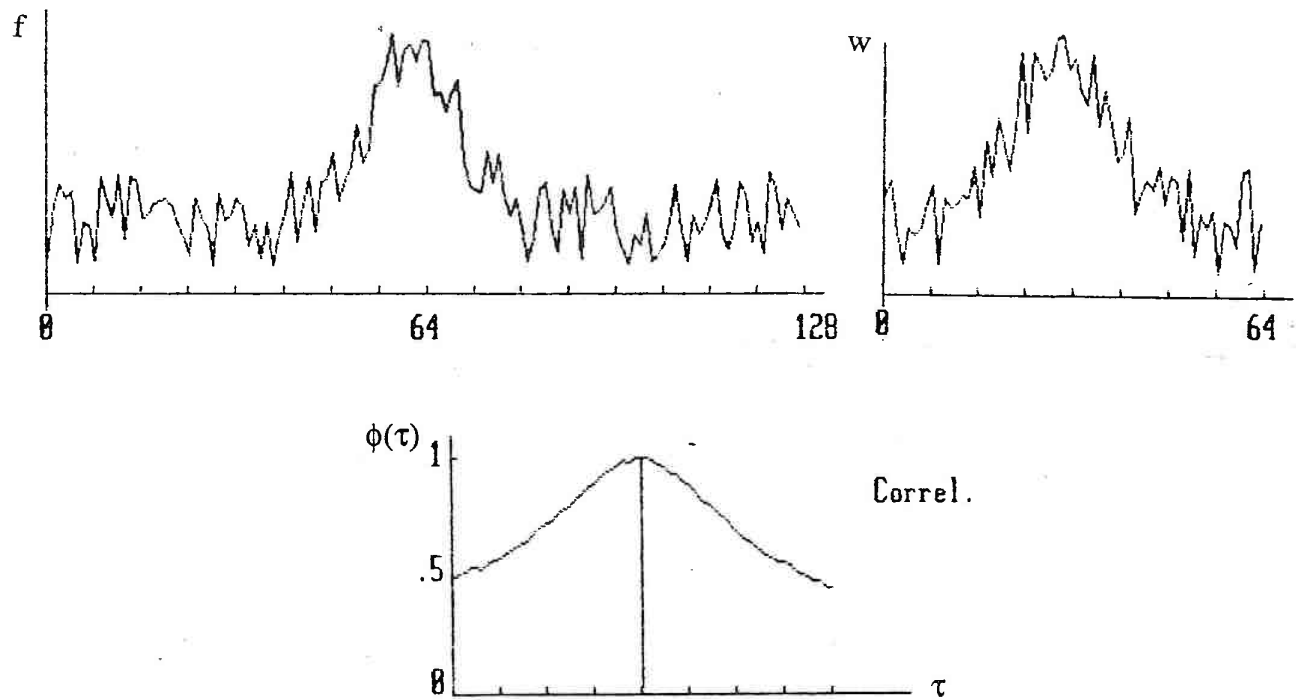


Figure 37 : Signaux continus

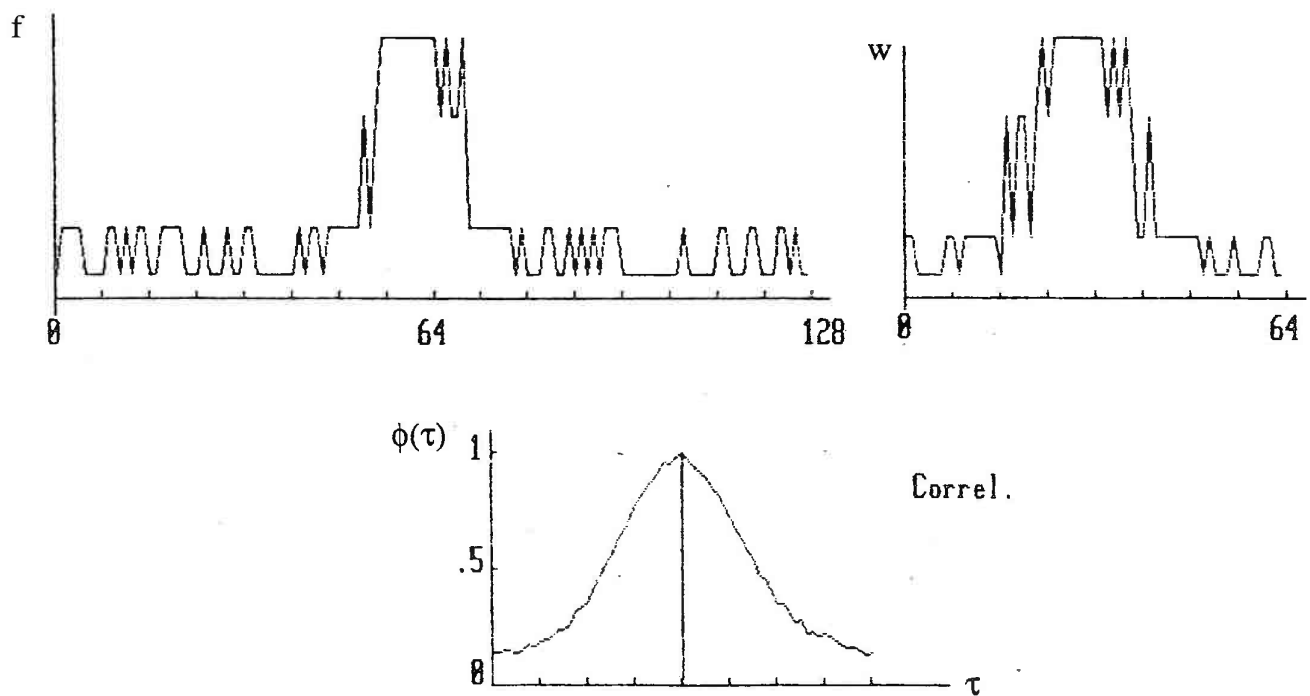


Figure 38 : Signaux quantifiés sur 4 niveaux puis pondérés

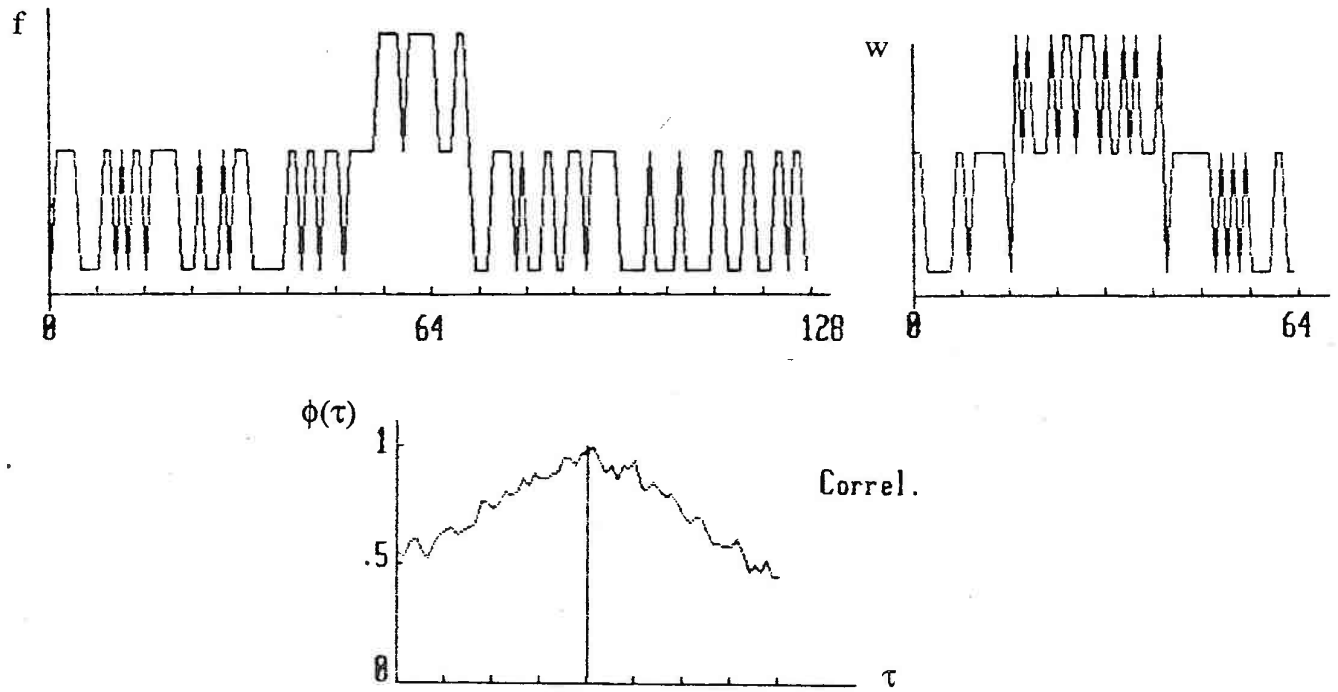
Exemple 2 : Rapport signal sur bruit de f et $w = 2$ 

Figure 39 : Signaux quantifiés sur 3 niveaux

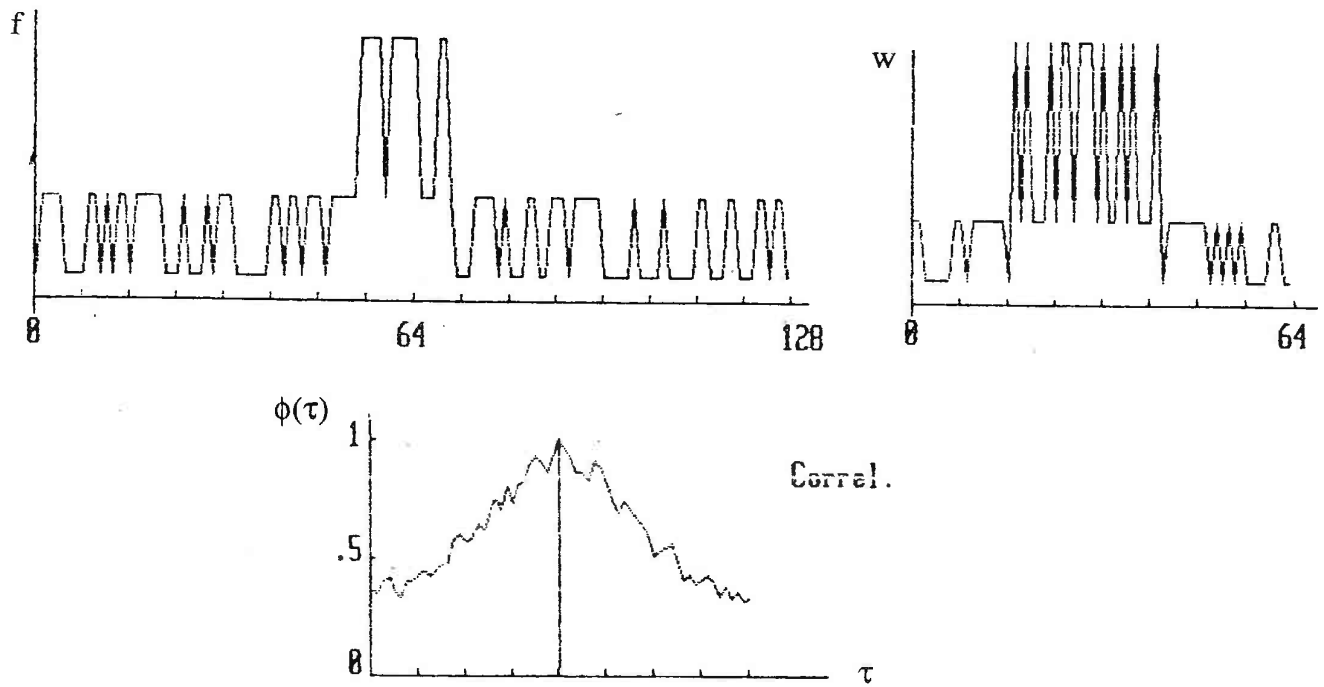


Figure 40 : signaux quantifiés sur 3 niveaux puis pondérés

Pour l'exemple 1, le bruit est faible, la quantification sur 4 niveaux n'altère donc pas sensiblement les signaux, toutefois, la pondération augmente le RSB de la fonction de corrélation. Dans l'exemple 2, la représentation sur trois paliers induit une perte d'information, il en résulte un décalage du maximum de corrélation. Cette erreur est compensée par la réduction de l'importance des données issues du premier seuillage.

L'apport de la pondération étant illustré, son principe est appliqué pour le calcul de la fonction de corrélation.

Le schéma de combinaison des résultats obtenus à chaque niveau (cf. fig. 30) devient donc le suivant :

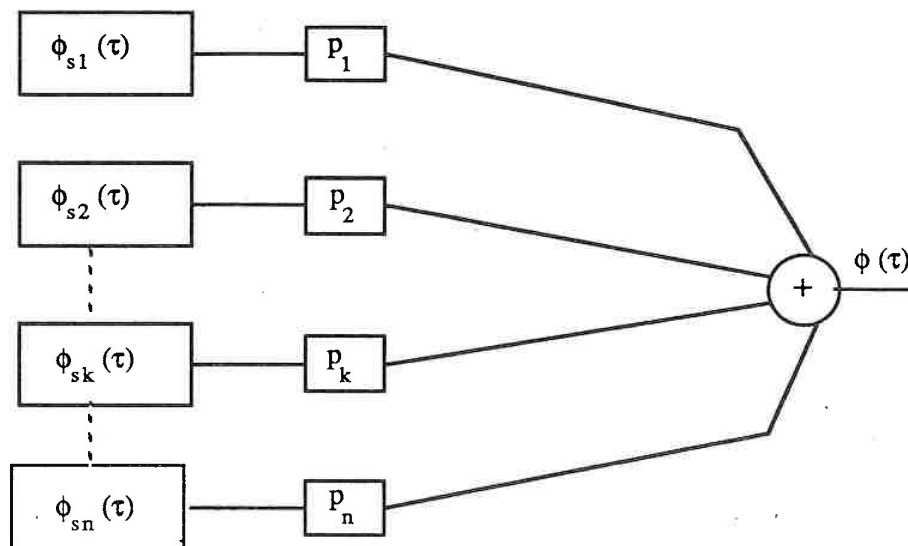


Figure 41 : sommation pondérée des fonctions de corrélation

II - 2 - 3 - 4 - Réalisation - Exemple

Après avoir défini la structure de l'algorithme, celui-ci est mis en oeuvre sur le système dédié à la poursuite de cible. Nous décrivons ici la partie correlative du système final.

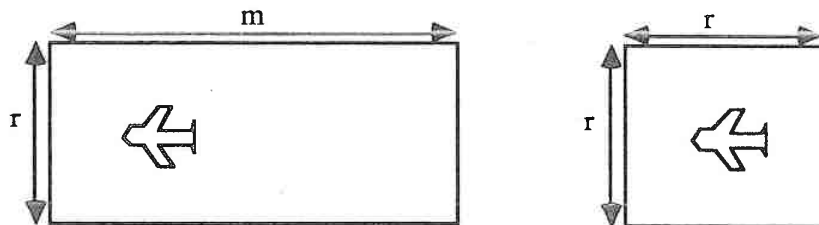


Figure 42 : Image de recherche (rec)

Image de référence (ref)

Le modèle (ref) représentant les contours de l'objet mobile est localisé dans l'aire de recherche (rec) par la procédure suivante :

Soient

$s_1 s_2 \dots s_n$: les n niveaux de seuils

$p_1 p_2 \dots p_n$: les n poids associés

$ref_{sk} [j]$: les éléments des tableaux des adresses relatives des points de ref dépassant le seuil s_k

$rec_{sk} [l]$: les éléments des tableaux des adresses relatives des points de rec dépassant le seuil s_k .

$cor [d]$: l'élément du tableau de dimension $m - r$ représentant la valeur de ϕ pour un décalage d entre ref et rec.

Initialisation

- seuillages de l'image de référence

. pour $k = 1$ à n

création du tableau ref_{sk}

Traitement

- seuillages de l'image de recherche

. pour $k = 1$ à n

création du tableau rec_{sk}

- corrélation

. pour $d = 1$ à $m - r$ $cor [d] = 0$

. pour $k = 1$ à n

$+j = 1$ $l = 1$ $d = 1$

boucle + tant que $ref_{sk} [j] > rec_{sk} [l]$

$[l = l + 1$

+ si $ref_{sk} [j] = rec_{sk} [l]$

[alors $cor [d] = cor [d] + p_k$

+ $ref_{sk} [j] = ref_{sk} [j] + 1$

+ $j = j + 1$

+ si $j \leq \dim \text{ref}_{sk}$

[	Alors aller en boucle					
	sinon $j = 1$ $l = 1$ $d = d + 1$					
	<table border="0"> <tr> <td style="padding-right: 10px;">si $d < m - r$</td> <td>[</td> <td>alors aller en boucle</td> </tr> <tr> <td></td> <td>]</td> <td>sinon rien</td> </tr> </table>	si $d < m - r$	[	alors aller en boucle		]
si $d < m - r$	[	alors aller en boucle				
	]	sinon rien				

. recherche $d_{\max} / \forall d \quad \text{cor}[d_{\max}] > \text{cor}[d]$

. pour $k = 1$ à n

+ réactualiser les tableaux ref_{sk}

. décaler l'aire de recherche.

. aller en Traitement.

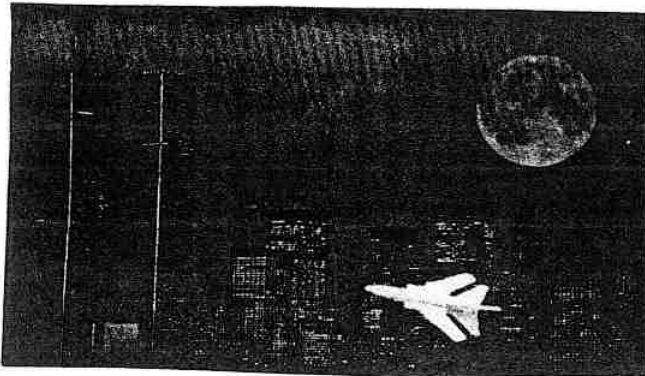
A chaque niveau de seuillage s_k et pour chaque décalage d entre ref et rec , on détermine les points présents à la fois dans ref_{sk} et dans rec_{sk} . Pour chacun de ces points la valeur du poids p_k est ajoutée à l'élément du tableau de corrélation ($\text{cor}[d]$), représentant l'appariement entre rec et ref lorsque le modèle est positionné au point d de l'aire de recherche. Les décalages sont représentés par une incrémentation des éléments de ref_{sk} . L'emplacement de l'objet est donné par le maximum du tableau de corrélation.

Puis l'image de référence est réactualisée en la remplaçant par la partie de l'aire de recherche englobant l'objet.

Ainsi le traitement peut être à nouveau exécuté à partir d'une acquisition à l'instant suivant. Il est alors nécessaire de repositionner l'aire de recherche. Nous aborderons au chapitre IV la technique de filtrage prédictif permettant d'exploiter les paramètres du mouvement à cette fin.

L'algorithme décrit est implémenté sur un système à base de MC-68020, les images issues de la caméra "NET-FLOU" sont converties par deux cartes de traitement d'image.

Donnons deux exemples de fonctions de corrélations engendrées par le programme (fig. 43 - fig. 44).



Scène 1

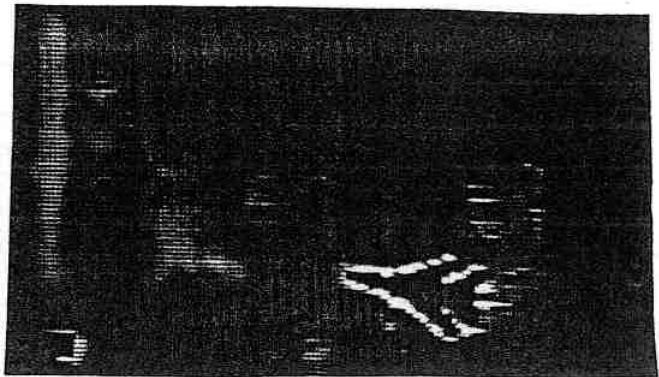


Image des contours 1

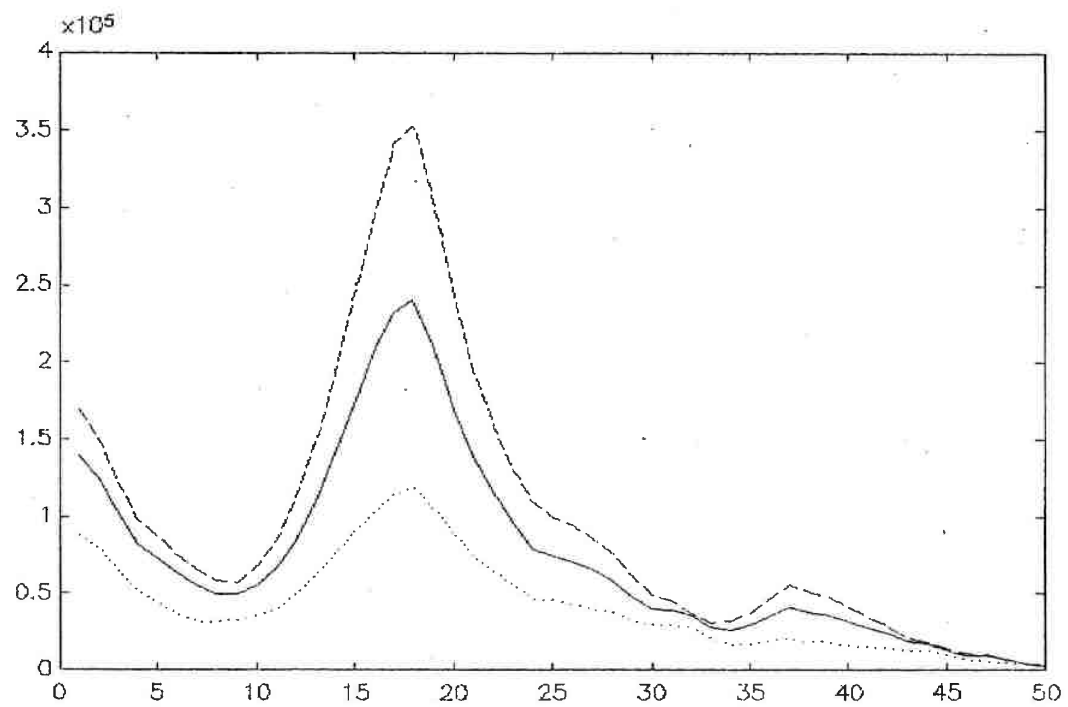
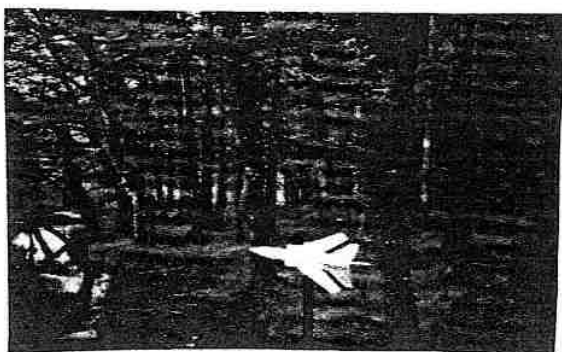


Figure 43 : Fonction de corrélation pour l'avion sur fond de ville

..... 1 seuil

———— 2 seuils

----- 3 seuils



Scène 2

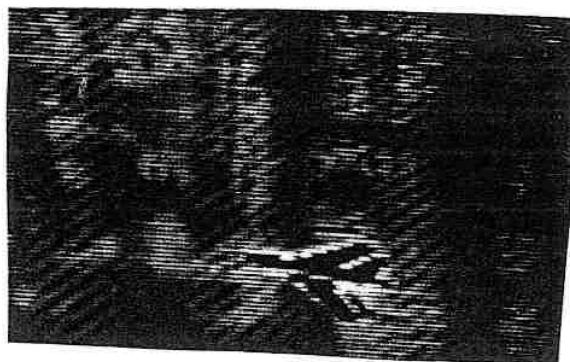


Image des contours 2

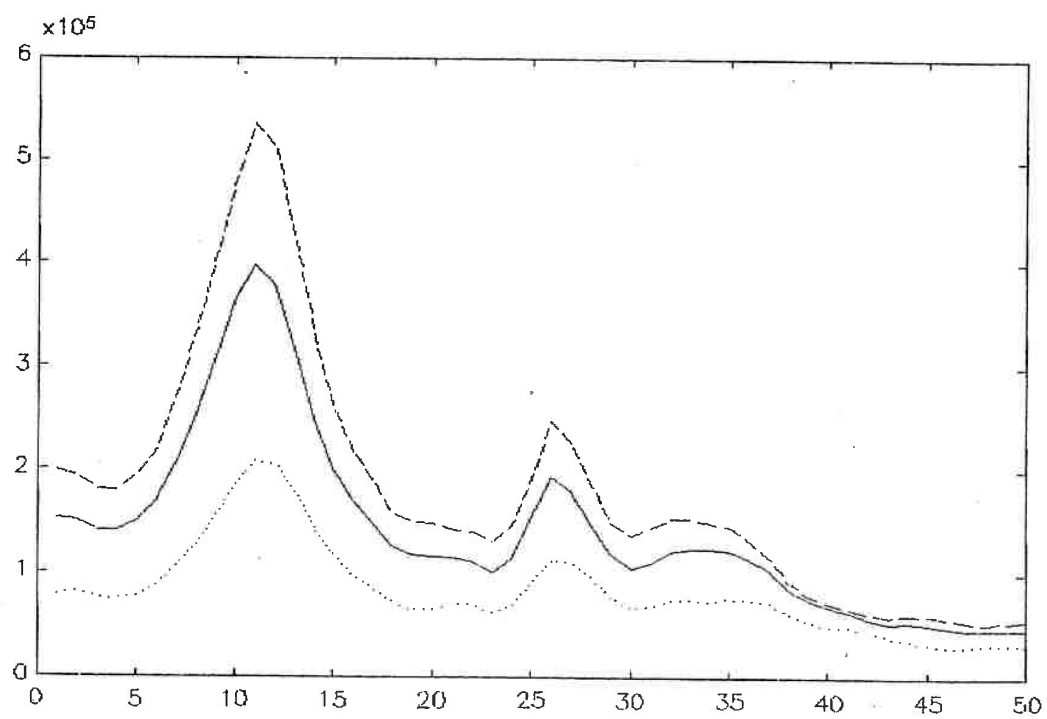


Figure 44 : Fonction de corrélation pour l'avion sur fond de forêt

..... 1 seuil

_____ 2 seuils

----- 3 seuils

Elles permettent de visualiser l'apport des seuillages multiples et pondérés, pour une bonne compréhension de la scène : le maximum représentant l'emplacement de l'avion et d'autant plus marqué que le nombre de seuil est grand.

Il s'agit là de résultats qualitatifs, mais la fonction de cet algorithme est de localiser un mobile en vue de sa poursuite ; les temps d'exécution constituent donc un critère déterminant. Ceux-ci sont aussi liés au nombre de niveaux considérés, nous examinerons cet aspect dans le chapitre suivant.

CHAPITRE III

III - APPLICATION À LA POURSUITE DE CIBLE : ÉTUDE QUANTITATIVE

La procédure de corrélation décrite au chapitre précédent, n'est totalement définie que lorsqu'un choix est fait concernant les paramètres suivants :

- seuil minimal
- nombre de seuils
- poids affecté à chaque seuil

La détermination de ces valeurs a été envisagée d'un point de vue qualitatif (§ II.2.3), à partir de processus aléatoires de caractéristiques connues. Afin de valider les résultats pour des images quelconques, une approche statistique de ces dernières doit être réalisée. Cette opération permet en outre l'évaluation quantitative de la procédure de corrélation : mesure des temps de traitement requis en fonction des paramètres choisis.

III - 1 - Définition statistique des images

Un tableau F représentant une image discrète peut être totalement caractérisé statistiquement par la densité de probabilité jointe des points le constituant [PRA - 74]

$$[3.1] \quad p(F) = p\{F(1,1), F(2,1), \dots, F(m,r)\}$$

Si les valeurs de chaque pixel sont régies par des lois statistiquement indépendantes, la densité jointe est définie par le produit des densités marginales :

$$[3.2] \quad p(F) = p\{F(1,1)\} \times p\{F(2,1)\} \times \dots \times p\{F(m,r)\}$$

Pour les densités jointes les expressions autres que la densité gaussienne sont rarement rencontrées dans la littérature. Mais cette densité est choisie pour représenter les coefficients de transformée unitaire de l'image ; en effet la loi normale est bipolaire et ne convient donc pas pour les luminances, toujours positives, d'une image naturelle.

III - 1 - 1 - Modélisation des données issues du capteur "N-F"

Visualisons l'effet du traitement "N - F" sur un signal binaire positif.



Figure 45: Signal net

signal flou



Figure 46 : signal net-flou

Le résultat de la soustraction est à valeur moyenne nulle. Mais le capteur traitant des images, le signal de sortie est forcé à des valeurs positives.

Soient



Figure 47 : signal de sortie

Les images issues du capteur optoélectronique peuvent être modélisées par un processus de valeur moyenne nulle dont la partie négative est ramenée à zéro. La loi gaussienne peut donc être envisagée pour représenter les luminances de ces images.

Si l'image source est statistiquement stationnaire [PRA - 74], les distributions de probabilité seront les mêmes en chaque point, de plus si le principe d'ergodicité est vérifié, les moyennes sur plusieurs images sont équivalentes à la moyenne spatiale. Sous ces deux hypothèses, la loi de probabilité régissant les points de l'image peut être mesurée par l'histogramme spatial de cette dernière.

En supposant que les niveaux de gris soient répartis suivant une loi normale, l'histogramme de l'image est une demi gaussienne, la partie négative de celle-ci étant ramenée à zéro.

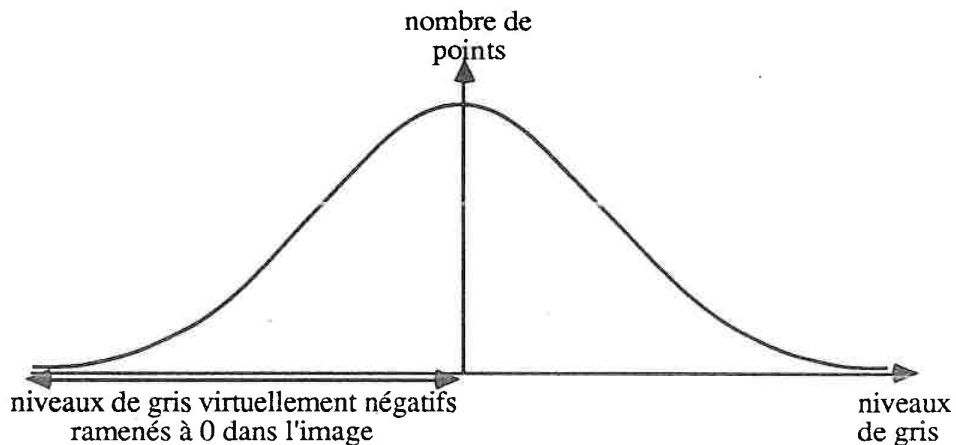


Figure 48 : histogramme pour une loi normale

Exemple :

Afin de vérifier la validité des hypothèses envisagées, nous avons réalisé l'histogramme de l'image d'un bruit traité "net-flou".

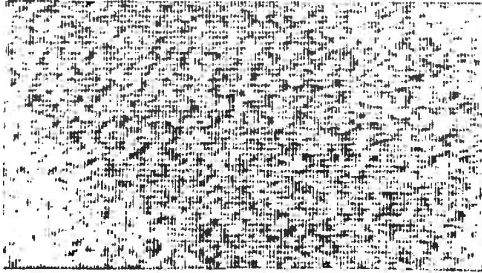


Figure 49: image de bruit

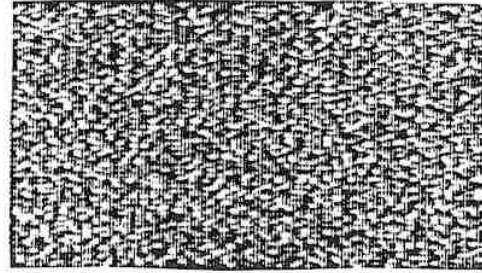


image traitée "net-flou"

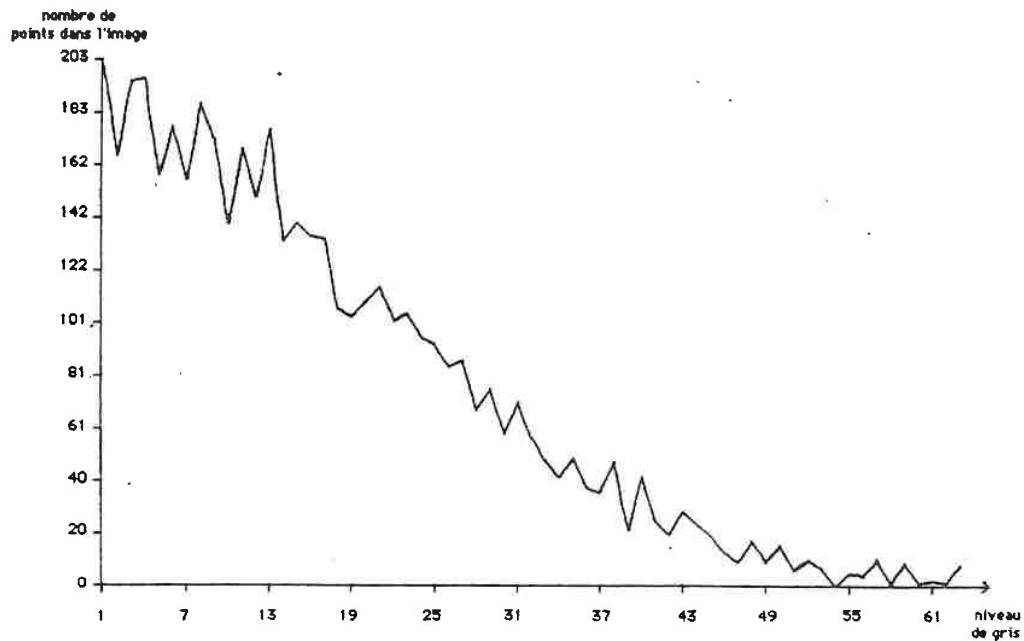


Figure 50 : histogramme de l'image traitée

Le nombre total de points est $N = 9604$. Les $N_0 = 4907$ points ayant un niveau de gris égal à 0 ne sont pas représentés sur le diagramme ci-dessus.

Soient

- x_i : un niveau de gris (de 1 à 63)
- k_i : le nombre de points de l'image au niveau x_i
- N_p : le nombre de points de l'image à niveau de gris positif
- N_n : le nombre de points dont le niveau de gris est 0, mais qui ont virtuellement un niveau de gris négatif symétrique des N_p points
- N'_0 : le nombre de points dont le niveau de gris serait 0 même si des niveaux de gris négatifs étaient envisageables.

Alors

$$[3.3] \quad N_p = \sum_{i=1}^{63} k_i = 4697$$

$$\text{si } N_n = N_p$$

$$[3.4] \quad N'_0 = N_0 - N_n = 210$$

Cette valeur ($N'_0 = 210$) est compatible avec l'histogramme de l'image (fig. 50), l'hypothèse de symétrie "virtuelle" paraît donc réaliste.

D'autre part, l'allure de la courbe est très proche de celle d'une gaussienne.

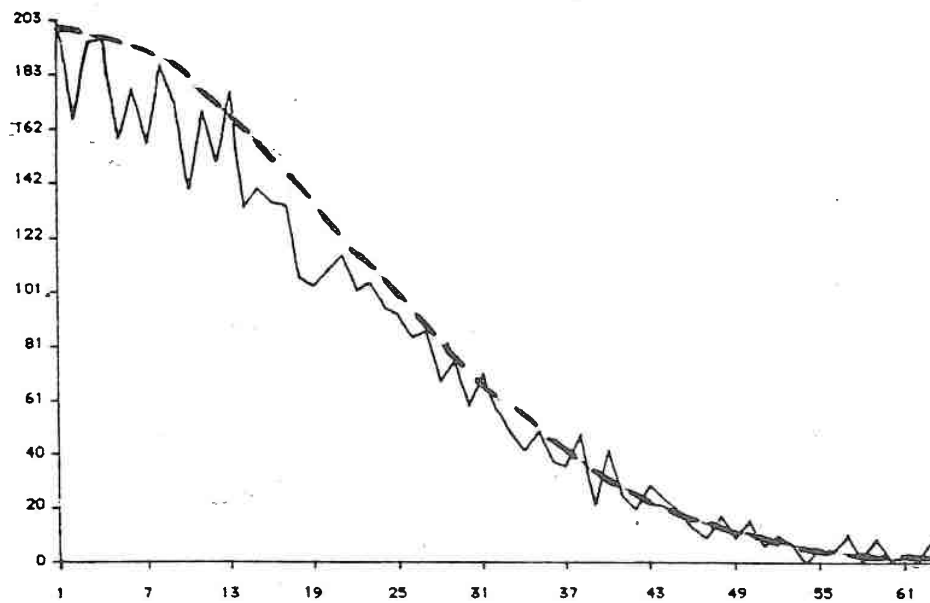


Figure 51 : histogramme ; ----- $e^{-x^2/2\sigma^2}$

Le calcul de l'écart type de la gaussienne devra tenir compte des N_p points virtuellement négatifs, il se fait donc suivant la formule :

$$[3.5] \quad \sigma = \sqrt{\frac{2}{N} \sum_{i=1}^{63} k_i x_i^2}$$

pour l'exemple représenté $\sigma = 20,8$.

Remarque :

La courbe d'histogramme est bruitée, ce phénomène a deux causes :

- l'image est de petite taille, une modélisation statistique serait mieux représentée sur un échantillon plus grand
- les images sont des processus discrets (ici 64 niveaux de gris), ce qui engendre un bruit de quantification que l'on peut caractériser de manière simple [COU - 84].

Si le pas de quantification est Δ et que la discrétisation est réalisée par arrondi, le bruit qui lui est associé est uniformément réparti de moyenne nulle et d'écart type σ_q :

$$[3.6] \quad \sigma_q^2 = \Delta^2 / 12$$

Ce facteur de bruit intervient chaque fois que l'on transpose sur des images, des résultats issus du traitement du signal. Pour le cas considéré, les niveaux de quantification sont constitués par chaque niveau de gris, ainsi on peut calculer :

$$[3.7] \quad \sigma_q = \frac{1}{\sqrt{12}}$$

et le rapport signal sur bruit de quantification :

$$[3.8] \quad \xi_q = \frac{\sigma^2}{\sigma_q^2} = (20,8)^2 \times 12 \approx 5.10^3$$

la faible influence du bruit de quantification permet donc de le négliger.

III - 1 - 2 - Evaluation du nombre de points dépassant un niveau donné dans l'image

La modélisation des signaux image menée est exploitée pour lier le nombre de points à traiter par corrélation avec le niveau de seuil choisi.

III - 1 - 2 - 1 - Aspect théorique

L'approche théorique de référence [LEV - 79] est élaborée sur des signaux temporels continus.

Soient :

$\eta(t)$: processus aléatoire normal

$S(t)$: signal déterministe

$\xi(t)$: bruit normal stationnaire, dérivable en moyenne quadratique, de moyenne nulle, d'écart type σ et de coefficient de corrélation $R(\tau)$.

tels que :

$$[3.9] \quad \xi(t) = \eta(t) - S(t)$$

$$[3.10] \quad \omega_1^2 = -R''_{\xi}(0) = \frac{\int_0^{\infty} \omega^2 F_{\xi}(\omega) d\omega}{\int_0^{\infty} F_{\xi}(\omega) d\omega}$$

avec $F_{\xi}(\omega)$: spectre énergétique du bruit.

L'auteur montre alors que le nombre moyen de dépassements du niveau x_0 par le processus $\eta(t)$ pour un intervalle de temps : $[t, t+T]$, est donné par $\lambda_1(x_0, t, T)$:

$$[3.11] \quad \lambda_1(x_0, t, T) = \frac{1}{2\pi\sigma^2\omega_1 T} \int_t^{t+T} e^{-\frac{[x_0-S(t)]^2}{2\sigma^2}} \int_0^{\infty} y e^{-\frac{[y-S'(t)]^2}{2\sigma\omega_1^2}} dy dt$$

développé sous la forme :

$$[3.12] \quad \lambda_1(x_0, t, T) = \frac{\omega_1}{2\pi T} \int_t^{t+T} e^{-\frac{[x_0-S(t)]^2}{2\sigma^2}} \left[e^{-\frac{[S'(t)]^2}{2\sigma^2\omega_1^2}} + \frac{S'(t) \sqrt{2\pi}}{\sigma\omega_1} F\left(\frac{S'(t)}{\sigma\omega_1}\right) \right] dt$$

où F est la fonction de Laplace : $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$

Soit ${}_1F_1$ la fonction hypergéométrique :

$$[3.13] \quad x \sqrt{2\pi} F(x) - \sqrt{\frac{\pi}{2}} x + e^{-\frac{x^2}{2}} = {}_1F_1\left(-\frac{1}{2}, \frac{1}{2}, -\frac{x^2}{2}\right)$$

d'où

$$[3.14] \quad \lambda_1(x_0, t, T) = \frac{\omega_1}{2\pi T} \int_t^{t+T} e^{-\frac{[x_0 - S(t)]^2}{2\sigma^2}} \left[{}_1F_1\left(-\frac{1}{2}, \frac{1}{2}, -\frac{S'(t)^2}{2\sigma^2 \omega_1}\right) + \sqrt{\frac{\pi}{2}} \frac{S'(t)}{\sigma \omega_1} \right] dt$$

Dans le cas particulier où la composante déterministe ($S(t)$) est absente, soit lorsque le signal est uniquement constitué de bruit.

$$[3.15] \quad S'(t) = S(t) = 0$$

$$[3.16] \quad {}_1F_1\left(-\frac{1}{2}, \frac{1}{2}, 0\right) = 1$$

Il vient :

$$[3.17] \quad \lambda_1(x_0, t, T) = \frac{\omega_1}{2\pi T} \int_t^{t+T} e^{-\frac{x_0^2}{2\sigma^2}} dt$$

Comme le processus $\xi(t)$ est stationnaire, $\lambda_1(x_0)$ ne dépend pas de t .

Le nombre moyen de dépassement par unité de temps est alors

$$[3.18] \quad \lambda_1(x_0) = \frac{\omega_1}{2\pi} e^{-\frac{x_0^2}{2\sigma^2}}$$

De même, il est montré que la durée moyenne de ces dépassements est donnée par $m_1(x_0)$.

$$[3.19] \quad m_1(x_0) = \frac{2\pi}{\omega_1} e^{\frac{x_0^2}{2\sigma^2}} \left[1 - F\left(\frac{x_0}{\sigma}\right) \right]$$

Soit si $x_0 \gg \sigma$

$$[3.20] \quad m_1(x_0) \approx \frac{\sigma \sqrt{2\pi}}{x_0 \omega_1}$$

La donnée intéressante pour la transposition dans le domaine des images est T_0 : le temps moyen pendant lequel le processus $\xi(t)$ dépasse le seuil x_0 sur un intervalle de temps T . A partir des définitions précédentes T_0 est calculé par :

$$[3.21] \quad T_0 = \lambda_1(x_0) \cdot m_1(x_0) \cdot T$$

soit si $x_0 \gg \sigma$

$$[3.22] \quad T_0 = \frac{T\sigma}{x_0 \sqrt{2\pi}} e^{-\frac{x_0^2}{2\sigma^2}}$$

Conformément au modèle décrit précédemment, l'image est considérée comme la réalisation d'un processus aléatoire dont la valeur est le niveau de gris, la suite des lignes de l'image constituant la composante temporelle. L'équivalent d'un intervalle de temps T est alors un nombre de point N (notons que cette grandeur est discrète).

Soit N_d le nombre de points dépassant le seuil s dans une image de N points, l'équivalence donne

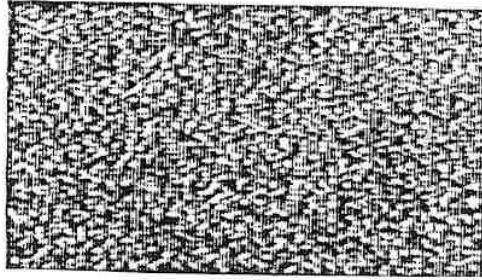
$$[3.23] \quad N_d = \frac{N\sigma}{s \sqrt{2\pi}} e^{-\frac{s^2}{2\sigma^2}}$$

la valeur de σ étant calculée sur toute l'image à partir de la formule [3.5], cette égalité est établie en supposant la présence du signal déterministe comme négligeable ; hypothèse crédible dans la mesure où les contours d'un objet ne représentent qu'une faible proportion de l'image.

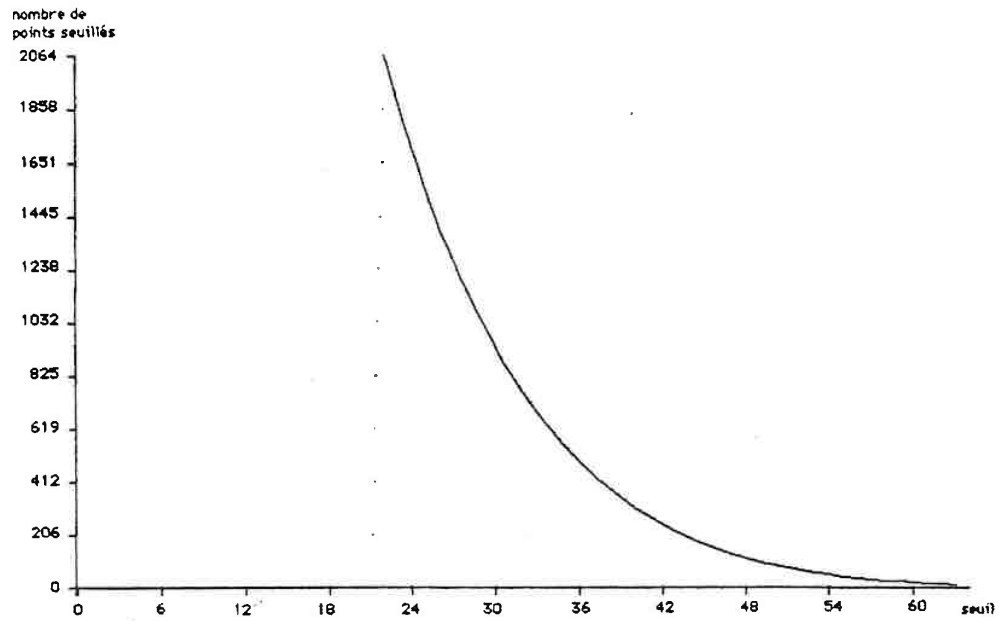
III - 1 - 2 - 2 - Résultats expérimentaux

Afin d'exploiter la formule [3.23] pour des applications pratiques nous avons testé sa validité sur un certain nombre d'exemples [HOR - 89 a]. Dans ce but, nous mesurons, pour une image donnée, par un programme de seuillage, le nombre de point dépassant un niveau de seuil variant de σ à 63. Les valeurs obtenues sont représentées sur une courbe dont l'abscisse est le seuil, et l'ordonnée le nombre de points seuillés (courbe expérimentale). Ensuite, pour cette même image, connaissant le nombre de points la composant (N) et son écart type σ , nous calculons par la formule, les nombres de points N_d dépassant théoriquement les niveaux de σ à 63. Les résultats sont représentés de la même manière (courbe théorique). Nous présentons 6 images et les 12 courbes correspondantes.

IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

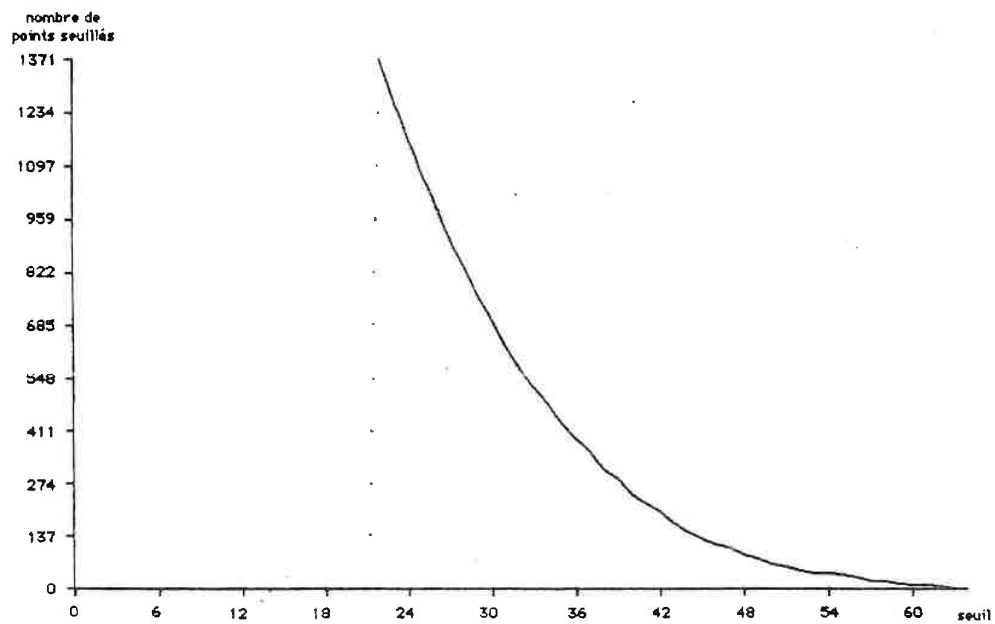
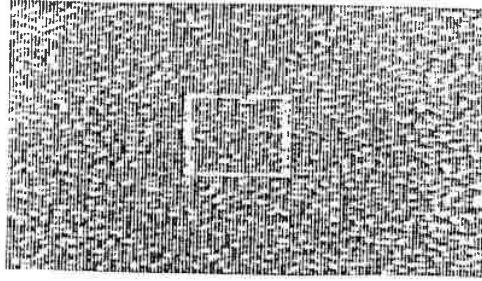
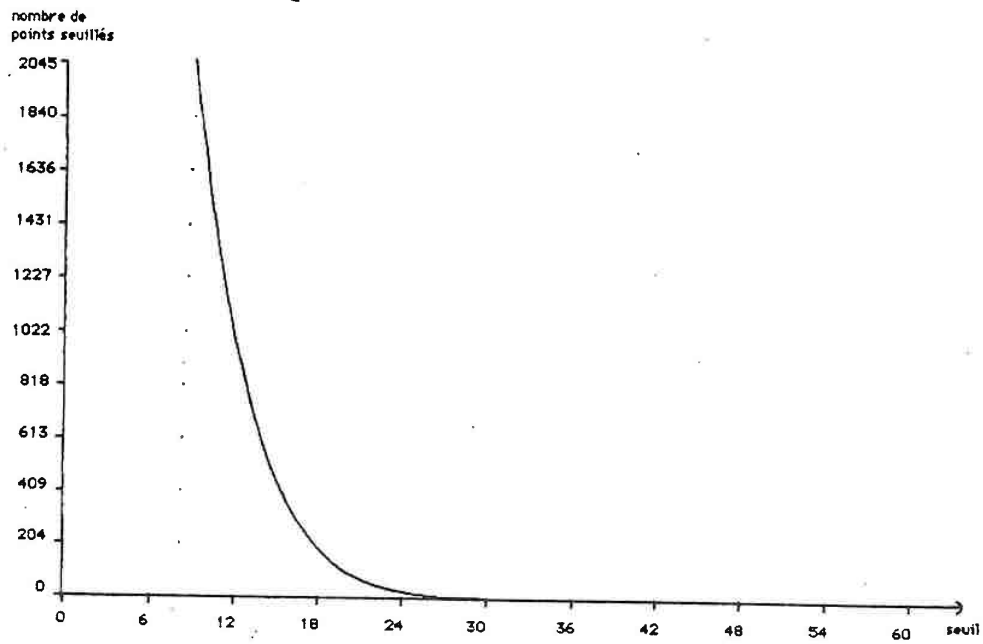


Figure 52 : Bruit gaussien d'écart type $\sigma = 20,8$

IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

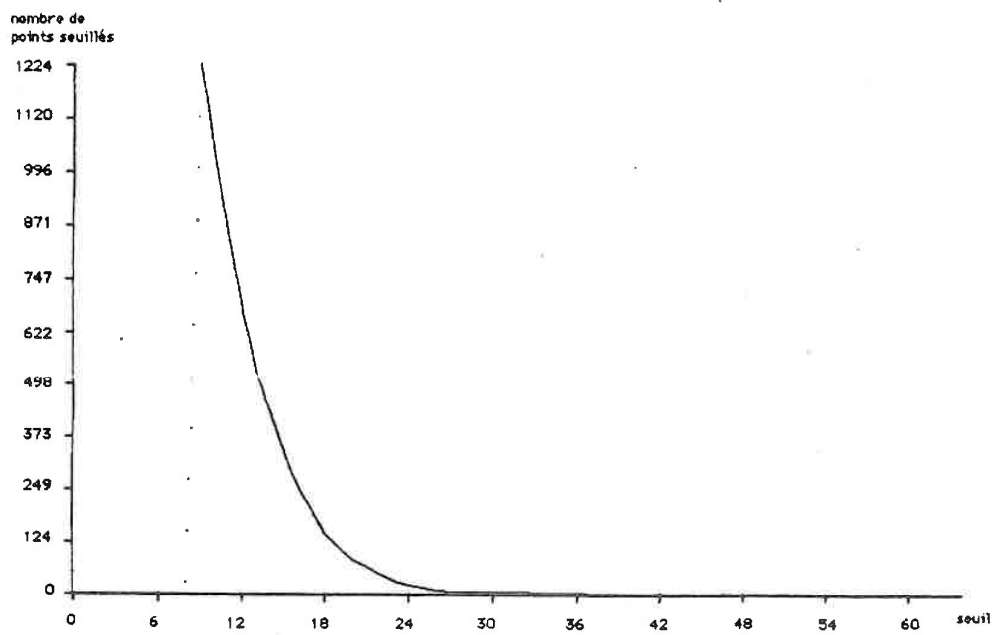
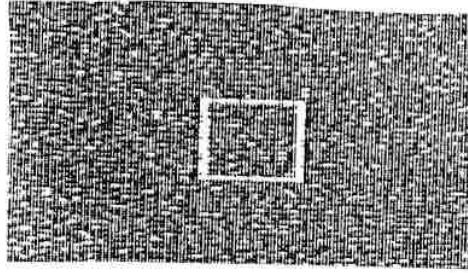
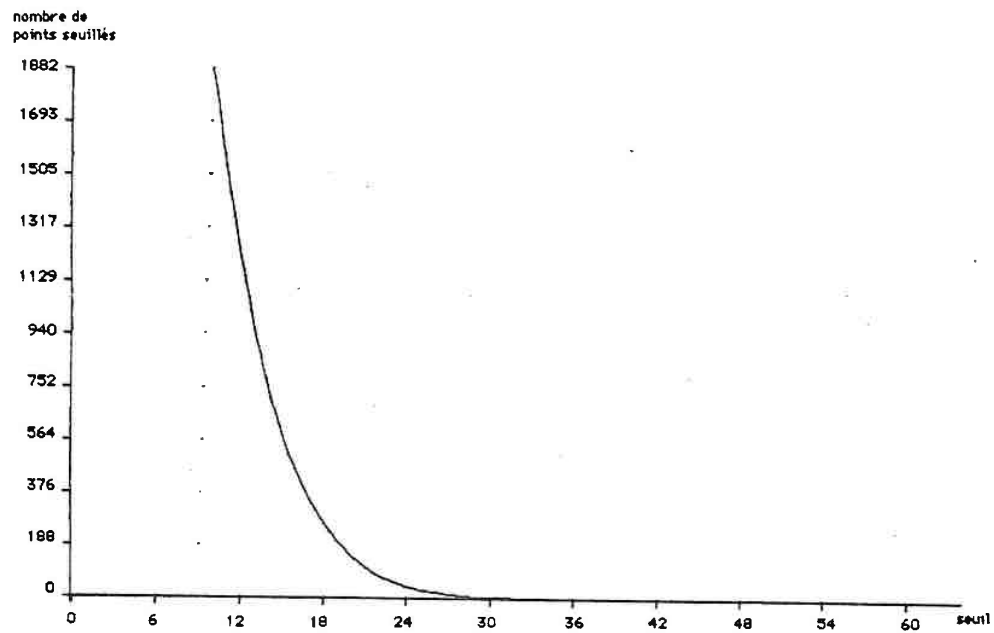


Figure 53 : Bruit gaussien + objet faiblement contrasté $\sigma = 8,5$

IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

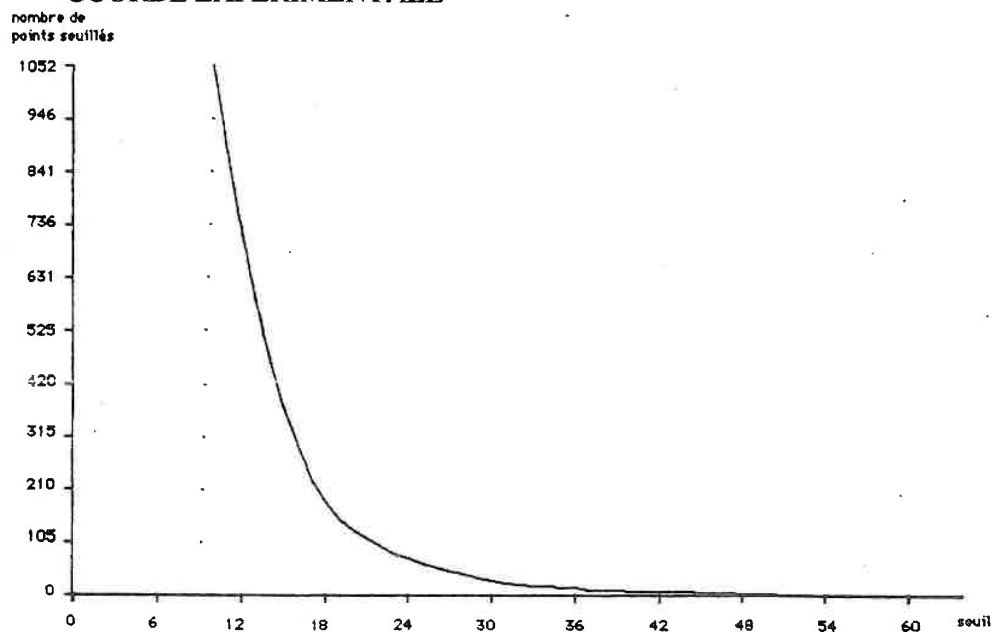
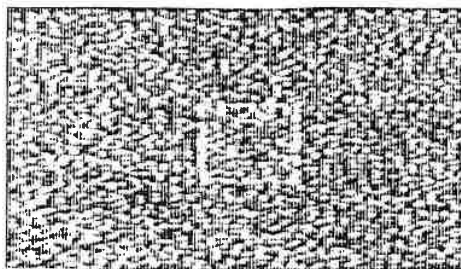
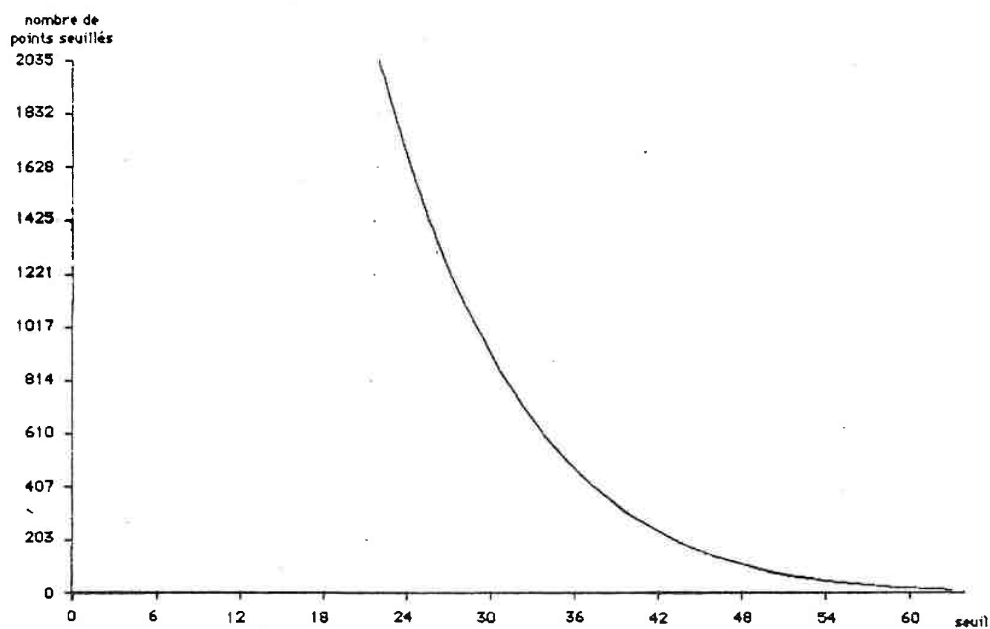


Figure 54 : Bruit gaussien + objet contrasté

IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

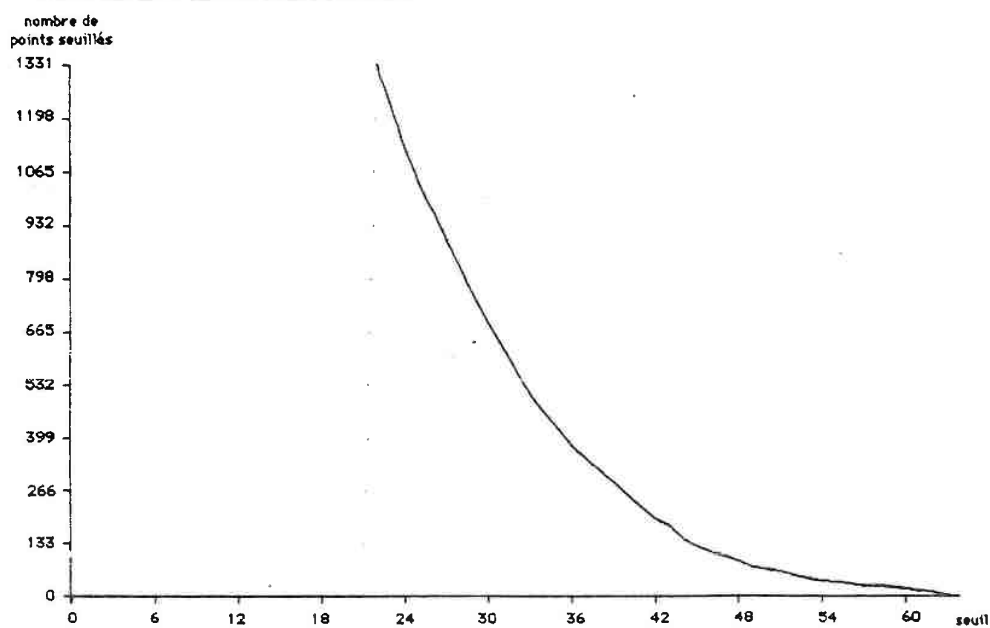
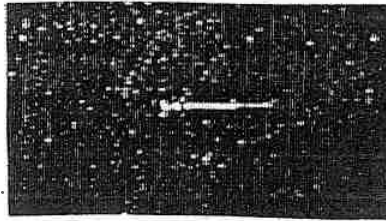
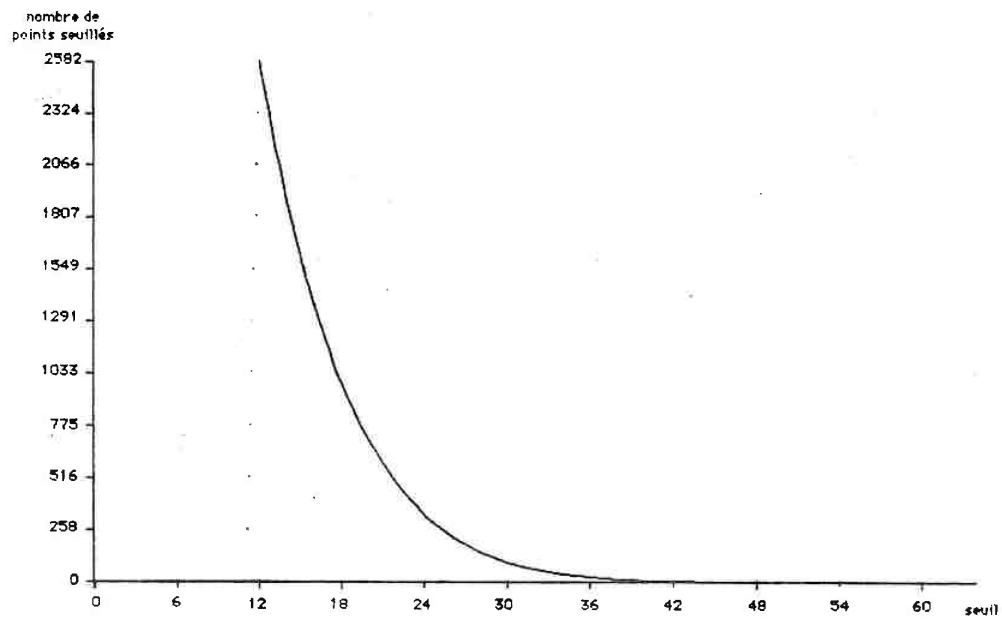


Figure 55 : Fort bruit gaussien + objet contrasté $\sigma = 20,6$

IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

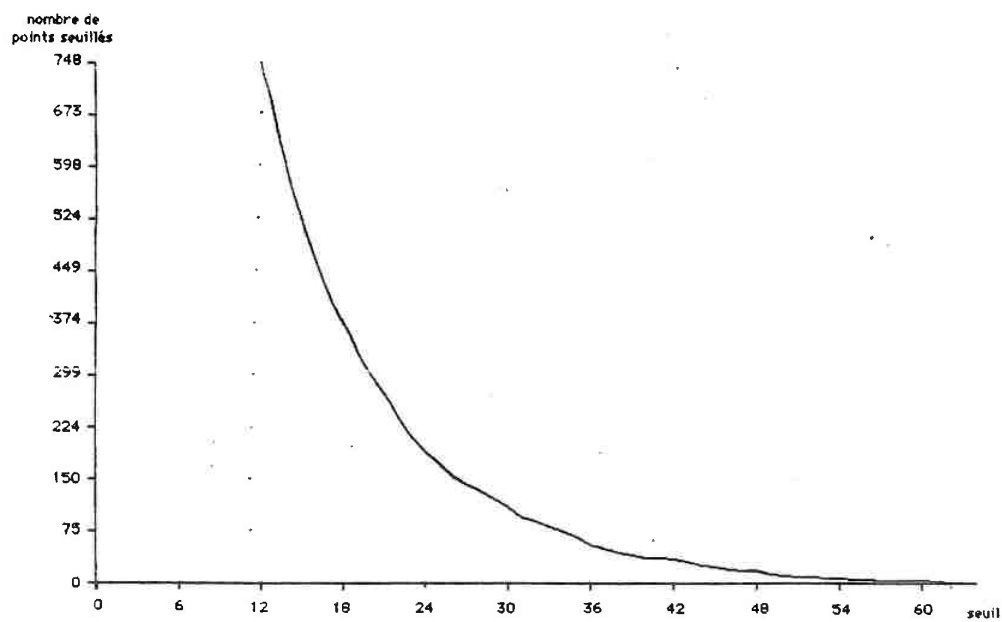
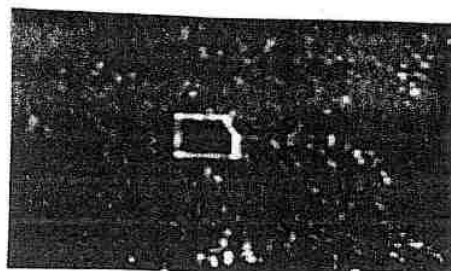
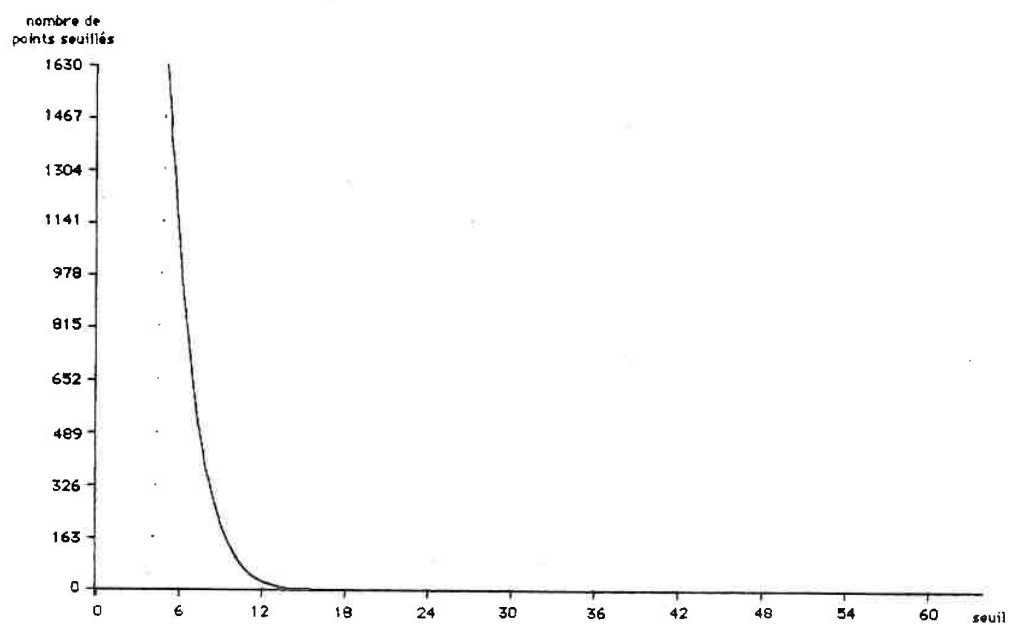


Figure 56 : vis sur plan de travail $\sigma = 12,66$

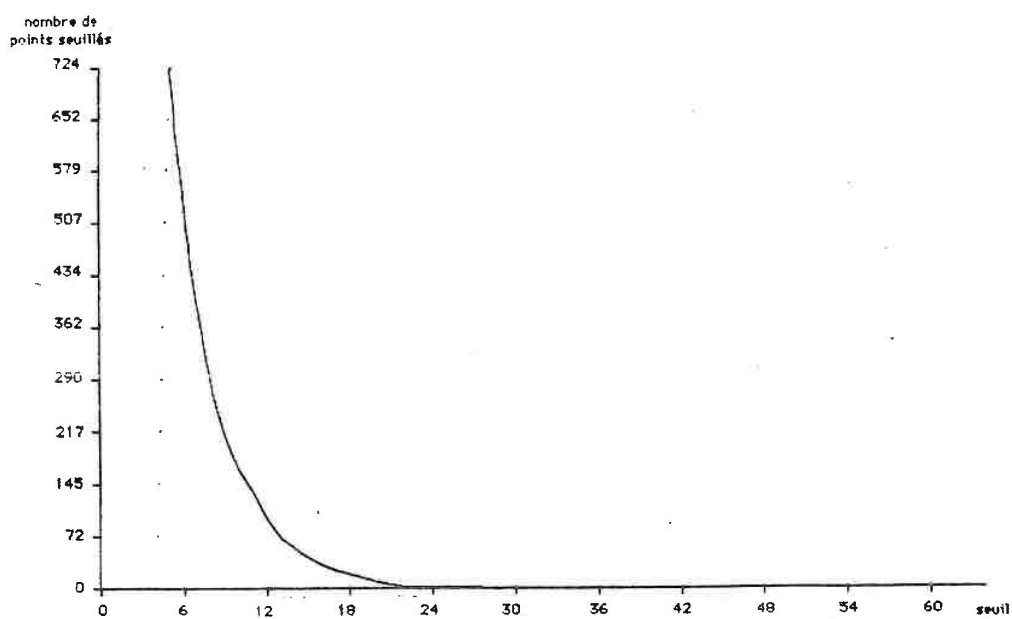
IMAGE



COURBE THEORIQUE



COURBE EXPERIMENTALE

Figure 57 : Véhicule sur paysage $\sigma = 4,5$

Les 4 premières images sont synthétiques (créées numériquement par programme), les deux dernières sont issues de scènes réelles acquises avec la caméra.

Les courbes sont normalisées, les graduations sont donc différentes selon les courbes.

Le premier exemple est l'image dont nous avons tracé l'histogramme précédemment, le modèle gaussien est donc justifié ; les courbes théorique et expérimentale ont exactement la même allure mais un décalage conséquent apparaît en valeur absolue au niveau des graduations, il se reproduit dans tous les exemples. Pour la figure 53, nous avons ajouté à l'image un signal déterministe représentant un objet rectangulaire, son amplitude est faible, les deux courbes sont là encore d'allure très similaire, leur décroissance est plus rapide que pour le cas précédent car l'écart type de l'image est plus faible. L'image suivante (fig. 54) contient un objet plus contrasté, il s'en suit un léger décalage de la courbe expérimentale pour laquelle la décroissance est moins rapide. Le même rectangle contrasté est ensuite couvert par un bruit plus fort (fig. 55) de mêmes caractéristiques que celui du premier exemple, la présence du signal déterministe ne modifie pas les résultats qui sont sensiblement égaux à ceux du premier cas.

Les deux dernières images (fig 56-57) représentent des scènes réelles pour lesquelles l'information déterministe (contour des objets) est plus marquée que le fond. De ce fait, la décroissance des courbes expérimentales est moins rapide que celle des courbes théoriques. Cette réserve étant émise, nous notons une bonne similitude entre les courbes théoriques et expérimentales, révélant le lien réel existant entre le nombre de points seuillés dans l'image et l'écart type (σ) mesuré pour cette dernière.

III - 1 - 2 - 3 - Conclusions

Sur l'ensemble des exemples traités, il apparaît un décalage permanent entre les valeurs théoriques et expérimentales, les premières étant toujours plus grandes. L'équation

$$N_d = \frac{N\sigma}{s \sqrt{2\pi}} e^{-\frac{s^2}{2\sigma^2}}$$

n'est donc pas vérifiée en valeur absolue, le terme $\frac{N\sigma}{s \sqrt{2\pi}}$ s'avère inexact.

Parmi les causes susceptibles d'expliquer ce phénomène, citons la possibilité que les hypothèses requises pour l'établissement de la formule ne soient pas totalement vérifiées (définition des processus aléatoires, calcul de la dérivée seconde du coefficient de corrélation...). Par ailleurs, le passage d'un processus continu à un processus discret (un temps devient un nombre de points) peut engendrer une transformation des caractéristiques λ_1 et m_1 (cf eq. [3.18] et [3.19]) qui se ressentirait sur l'égalité finale.

Quoiqu'il en soit, l'écart observé n'est pas modélisable de manière simple en une fonction des paramètres de l'image.

En revanche, pour une même image, les courbes normalisées théoriques et expérimentales sont très proches, elles sont donc exploitables en ce sens qu'elles justifient la décroissance exponentielle de la formule [3.23]. Ainsi, connaissant le nombre de point dépassant un certain seuil et l'écart type d'une image, il est possible de déterminer le nombre de points dépassant n'importe quel seuil pour cette même image.

Soient

N_1 : nombre de point dépassant un seuil s_1

N_2 : nombre de point dépassant un seuil s_2

Connaissant N_1 , il est possible de calculer N_2 en adaptant la formule [3.23] de manière relative :

$$[3.24] \quad N_2 = N_1 \frac{s_1}{s_2} e^{-(s_2^2 - s_1^2) / 2\sigma^2}$$

Cette formule, justifiée expérimentalement, sera celle que nous utiliserons pour l'évaluation du nombre de points à traiter lors d'une corrélation.

III - 2 - Estimation des temps de traitement

III - 2 - 1 - Influence du nombre de points traités sur le temps d'exécution

Le programme de corrélation que nous avons défini (cf § II.2.3.4) se compose, une fois la phase d'initialisation passée, de quatre étapes :

- 1 - seuillages de l'aire de recherche
- 2 - établissement de la fonction de corrélation
- 3 - recherche du maximum de cette fonction
- 4 - réactualisation des modèles.

Chacune de ces étapes nécessite un temps de traitement différent et variant ou non avec la configuration de la scène traitée. Nous recherchons un modèle simple permettant de lier les paramètres de l'image au temps d'exécution requis pour localiser un objet dans cette dernière. Dans ce but, nous avons réalisé un programme mesurant pour chaque étape son temps d'exécution, et délivrant en outre le nombre de points traités. Lorsque plusieurs seuillages sont effectués, les points traités à chaque niveau sont additionnés, un même point est ainsi comptabilisé plusieurs fois. Nous entendons donc par nombre de points, le nombre d'adresses répertoriées dans les tableaux $recs_k$. Les images de référence ont une dimension de 50 x 50 pixels et sont recherchées dans une aire de 50 x 100 pixels ; les essais sont faits sur des images de différentes natures : éclairage variable, objets différents, et le programme est exécuté avec 1, 2 ou 3 seuils ; il délivre les résultats décrits et interprétés ci-dessous.

- Le temps de seuillage T_s est indépendant du nombre de points sélectionnés. Il varie en fonction du nombre de seuils, mais avec une faible amplitude car le seuillage des niveaux supérieurs est fait sur les seuls points dépassant le niveau inférieur, ils sont relativement rapides. Les limites de T_s sont :

$$[3.25] \quad 0,05 \text{ s} < T_s < 0,07 \text{ s}$$

- Le temps d'établissement de la fonction de corrélation (T_c) dépend du nombre de points à traiter. La stratégie retenue engendre de plus une variation du nombre de décalages et de comparaisons requis en fonction de la configuration de l'image (points seuillés consécutifs ou non). Il apparaît donc que T_c varie linéairement en fonction du nombre de points traités (N) pour une image donnée et un nombre de seuils fixe ; pour une autre image ou un autre nombre de seuils, la fonction liant T_c et N sera encore linéaire, mais d'une autre nature. Afin de visualiser une possible unification des résultats obtenus, nous représentons sur une même courbe l'ensemble des valeurs de T_c et de N issu des différents essais (fig. 58).

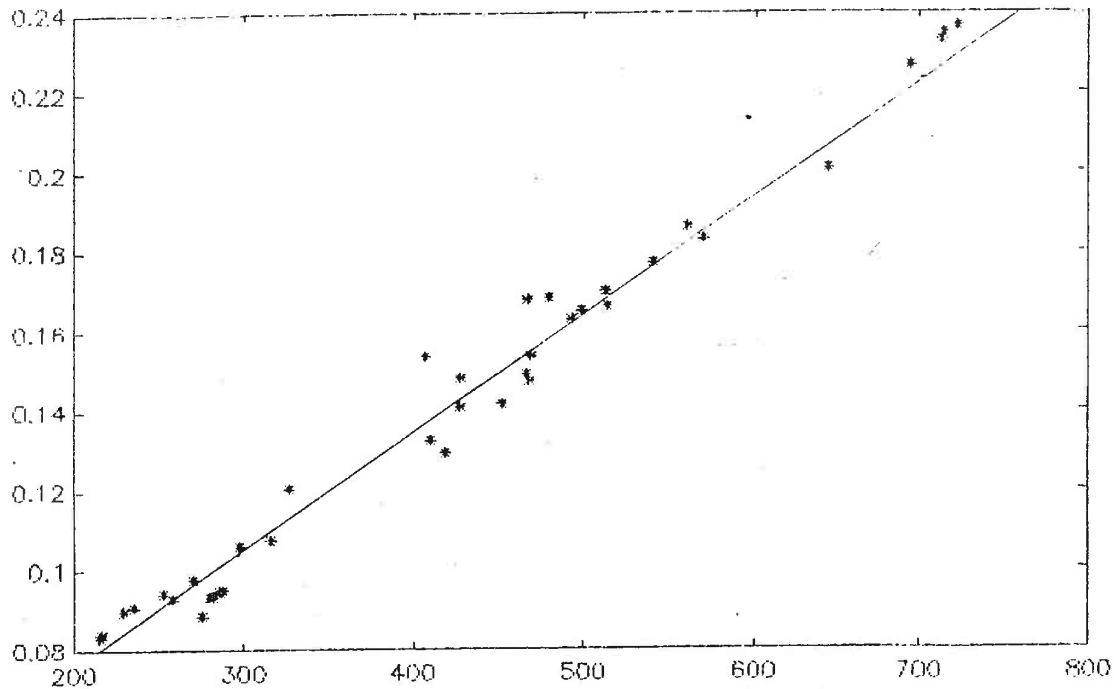


Figure 58 : Temps d'établissement de la fonction de corrélation
en fonction du nombre de points traités

L'ensemble des points constituant presque une droite dont nous allons déterminer l'équation par la méthode des moindres carrés.

En posant :

$$[3.26] \quad \phi = [N(i) \ 1]$$

$$[3.27] \quad \theta = \begin{bmatrix} a \\ b \end{bmatrix}$$

un point quelconque suit l'équation :

$$[3.28] \quad T_c(i) = a N(i) + b + \varepsilon(i)$$

Soit

$$[3.29] \quad T_c = \phi \theta + \varepsilon$$

la solution pour θ rendant ε minimum est :

$$[3.30] \quad \theta = (\phi^T \phi)^{-1} \phi^T T_c$$

Nous obtenons donc

$$[3.31] \quad a = 3.10^{-4} \text{ s/point} \quad b = 1,25 \cdot 10^{-2} \text{ s}$$

et

$$[3.32] \quad \varepsilon = \sqrt{\frac{T}{\varepsilon}} \varepsilon = 1,5 \cdot 10^{-3} \text{ s}$$

l'équation liant T_c et N pour l'ensemble des cas est donc

$$[3.33] \quad T_c = 3 \cdot 10^{-4} N + 1,25 \cdot 10^{-2}$$

avec une erreur moyenne de 1 %.

- Le temps de recherche du maximum (T_R) est constant :

$$[3.34] \quad T_R = 2 \cdot 10^{-4} \text{ s}$$

- Pour la réactualisation du modèle, le temps requis (T_M) est de $1,2 \cdot 10^{-3}$ s par seuil. En effet, il s'agit dans cette étape de chercher parmi les éléments des tableaux de points seuillés de l'aire de recherche (rec_{sk}), ceux représentant les points du modèle et de les placer dans les tableaux de référence (ref_{sk}). La durée de l'opération est la même quelque soit le niveau considéré. Comme T_M est petit devant T_c , nous prendrons, pour l'évaluer, sa valeur maximale :

$$[3.35] \quad T_M = 3,6 \cdot 10^{-3} \text{ s}$$

le temps total de localisation de la cible (T_T) est donc :

$$[3.36] \quad T_T = T_s + T_c + T_R + T_M$$

$$[3.37] \quad T_T = 3 \cdot 10^{-4} N + 9 \cdot 10^{-2}$$

$$[3.38] \quad T_T = aN + T_0 \quad \begin{array}{l} a = 3 \cdot 10^{-4} \text{ s/points} \\ T_0 = 9 \cdot 10^{-2} \text{ s} \end{array}$$

III - 2 - 2 - Application aux choix des seuils

Le choix du seuil minimal a été défini au paragraphe II.2.3.2, en fonction d'un critère qualitatif, de manière à ne pas entraîner une probabilité de fausse alarme trop importante. La mise en place des équations [3.24] liant le nombre de points sélectionnés au niveau du seuil, et [3.37] liant le temps d'exécution au nombre de points traités, va nous permettre d'aborder le problème d'un point de vue quantitatif. C'est à dire que les valeurs de seuil seront choisies en fonction des limites imposées par les temps d'exécution.

Notons qu'il n'est pas envisageable d'établir une relation directe entre le temps de calcul et le seuil, car le nombre de points seuillés est toujours défini de manière relative pour une image donnée.

Dans le cas de la poursuite de cible, le temps est une valeur à minimiser. Nous admettons donc qu'il existe un temps maximum $T_{\max}$ disponible entre deux instants d'échantillonnage pour la corrélation, et au delà duquel les résultats seraient inexploitable. Ce temps dépend du contexte extérieur : de la taille du champs image, de la vitesse maximale de la cible, et des performances de l'ordinateur utilisé. Il définit un nombre maximal de points ($N_{\max}$) pouvant être sélectionnés dans l'image.

D'après [3.38] :

$$[3.39] \quad N_{\max} : (T_{\max} - T_0) / a$$

Si la procédure de corrélation est établie à partir d'un seul seuillage, le niveau de ce dernier ne devra pas dépasser une limite inférieure ($s_{\min}$) correspondant à la sélection de $N_{\max}$ points. La valeur de cette limite peut être calculée à partir d'une mesure expérimentale, pour une image donnée : sachant que $N_{\exp}$ point dépassant le seuil $s_{\exp}$ dans une image d'écart type σ ,

$$[3.24] \rightarrow s_{\min} N_{\max} = s_{\exp} N_{\exp} e^{-\frac{(s_{\min}^2 - s_{\exp}^2)}{2\sigma^2}}$$

Soit

$$[3.40] \quad \frac{e^{-\frac{s_{\min}^2}{2\sigma^2}}}{s_{\min}} = \frac{N_{\max}}{N_{\exp}} \frac{e^{-\frac{s_{\exp}^2}{2\sigma^2}}}{s_{\exp}}$$

Exemple :

la vitesse maximale autorisée pour un objet est

$$v_{\max} = 0,1 \text{ m/s}$$

la marge de vue est de ± 50 pixels, la caméra est à une distance du champ d'objet telle que 1 pixel = 0,28 cm. Le temps maximal disponible pour une corrélation est :

$$T_{\max} = \frac{50 \times 0,28}{10} = 1,4 \text{ s}$$

$$\text{et } N_{\max} = \frac{1,4 - 0,09}{3 \cdot 10^{-4}} = 4366$$

Des mesures préliminaires sur les images à traiter montrent qu'elles sont d'écart type $\sigma = 10$ et que pour un seuillage : $s_{\exp} = 15$ $N_{\exp} = 500$ pixels.

$$\frac{N_{\max}}{N_{\exp}} = 8,7$$

On se reporte alors à la courbe $f(x) = e^{-x^2/2\sigma^2}$

autour du point $x = s_{\exp}$

$$\text{Il vient } \frac{e^{-\frac{s_{\exp}^2}{2\sigma^2}}}{s_{\min}} = 0,02$$

$$\text{l'équation [3.40] donne } \frac{e^{-\frac{s_{\min}^2}{2\sigma^2}}}{s_{\min}} = 0,174$$

et de la courbe on tire $s_{\min} = 5$

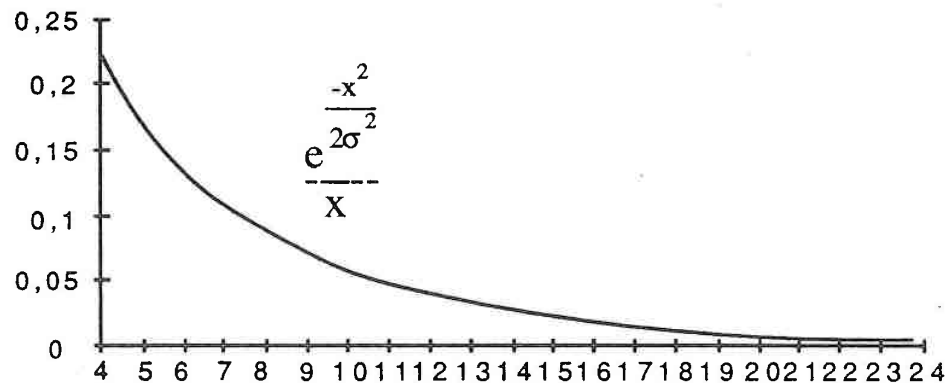


Figure 59 : $f(x) = \frac{e^{-x^2/2\sigma^2}}{x}$

Remarque : pour ce cas précis le résultat obtenu n'est pas en contradiction avec le choix de $s_{\min} = \sigma$ adopté par l'étude qualitative ; en prenant un seuil minimal de σ , il restera du temps disponible pour d'autres seuillages.

Si deux niveaux sont requis pour la corrélation, la somme des points sélectionnés ne devra pas excéder $N_{\max}$ (le lien avec le temps d'exécution étant le même quelque soit le nombre de seuil).

Soient s_1 et s_2 les deux niveaux de seuillage
 N_{m1} et N_{m2} les points sélectionnés correspondant

alors

$$[3.41] \quad N_{m1} + N_{m2} = N_{\max}$$

En définissant chaque valeur en fonction de $N_{\max}$ et $s_{\min}$, il vient:

$$[3.42] \quad N_{\max} \frac{S_{\min}}{S_1} e^{-\frac{s_1^2 - s_{\min}^2}{2\sigma^2}} + N_{\max} \frac{S_{\min}}{S_2} e^{-\frac{s_2^2 - s_{\min}^2}{2\sigma^2}} = N_{\max}$$

$$[3.43] \quad \frac{e^{-\frac{s_1^2}{2\sigma^2}}}{S_1} + \frac{e^{-\frac{s_2^2}{2\sigma^2}}}{S_2} = \frac{e^{-\frac{s_{\min}^2}{2\sigma^2}}}{S_{\min}}$$

Le calcul de s_1 et s_2 est complexe par cette relation, cependant nous pouvons le considérer comme un critère permettant de prévoir les dépassements de temps. Les niveaux de seuils choisis en fonction des critères de décision précédemment cités seront validés s'ils vérifient l'inégalité :

$$[3.44] \quad \frac{e^{-\frac{s_1^2}{2\sigma^2}}}{S_1} + \frac{e^{-\frac{s_2^2}{2\sigma^2}}}{S_2} \leq \frac{e^{-\frac{s_{\min}^2}{2\sigma^2}}}{S_{\min}}$$

Celle-ci pouvant être généralisée pour k seuils :

$$[3.45] \quad \sum_{i=1}^k \frac{e^{-\frac{s_k^2}{2\sigma^2}}}{S_k} \leq \frac{e^{-\frac{s_{\min}^2}{2\sigma^2}}}{S_{\min}}$$

III - 2 - 3 - Recherche de la taille optimale de la fenêtre

Pour la poursuite de cible sans étude de la trajectoire du mobile (cette étude est abordée au chapitre suivant), la fenêtre de recherche est positionnée à chaque étape de manière à placer l'objet en son centre.

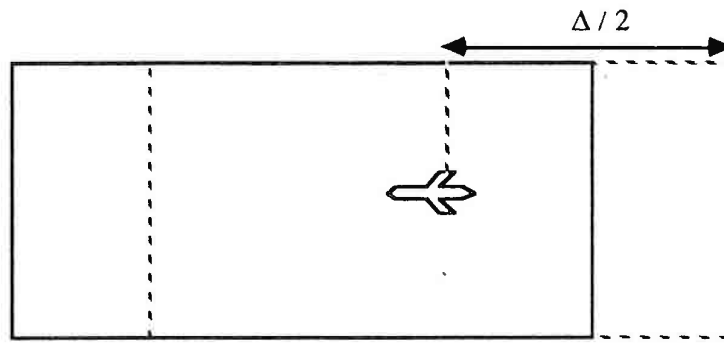


Figure 60 : — fenêtre à t_{n-1} objet à t_{n-1}

----- fenêtre à t_n

Le déplacement maximal autorisé pour l'objet entre deux instants d'échantillonnage est $\Delta/2$ si Δ est la taille de la fenêtre. Il permet de déterminer une vitesse maximale admissible en fonction du temps de traitement T_T :

$$[3.46] \quad v_{\max} = \frac{\Delta}{2T_T}$$

Or T_T est fonction du nombre (N) de points traités, (équation [3.37]), qui est lui même fonction de la taille de la fenêtre. En supposant que l'image est homogène, il existe un rapport de proportionnalité entre ces deux dernières variables :

$$[3.47] \quad N = k\Delta$$

Le coefficient k varie selon les images, il est calculé par une mesure $N_{\exp} = k \Delta_{\exp}$

$$[3.48] \quad T_T = ak\Delta + T_0$$

ainsi $v_{\max}$ peut être déterminé uniquement en fonction de Δ par :

$$[3.49] \quad v_{\max} = \frac{\Delta}{2(ak\Delta + T_0)}$$

Cette fonction de Δ est monotone croissante, ce qui porterait à croire que plus l'aire de recherche est grande, plus la vitesse maximale perceptible est importante. La taille optimale de la fenêtre de recherche serait donc la plus large possible, c'est à dire englobant l'ensemble des points captés par la caméra. Mais deux remarques s'imposent :

- le traitement d'un grand nombre de points implique la création et le maniement de tableaux de grandes dimensions ; ce qui peut poser des problèmes de place en mémoire.

- L'établissement de la relation linéaire entre T et N est effectué sous certaines hypothèses (temps de seuillage constant, temps de réactualisation borné supérieurement), elles peuvent être mises en défaut par des conditions expérimentales trop différentes de celles envisagées lors des essais.

La fonction $v_{\max} = f(\Delta)$ est croissante, elle a pour dérivée $f'(\Delta)$:

$$[3.50] \quad f'(\Delta) = \frac{1}{2} \frac{T_0}{(ka \Delta + T_0)^2}$$

qui est décroissante. Donc si $v_{\max}$ augmente avec Δ , il existe une valeur $\Delta_{\lim}$ à partir de laquelle cette augmentation n'est plus conséquente. Un choix intéressant pour la taille de fenêtre est cette valeur $\Delta_{\lim}$. Le temps de traitement correspondant sera ensuite calculé, afin de mesurer l'écart entre le domaine défini et celui des essais.

Illustrons ce principe sur un exemple :

Soient une image et deux niveaux de seuillage s_1 et s_2 fixés, et délivrant par une mesure :

$$N_{\exp} = 500 \quad \text{pour} \quad \Delta_{\exp} = 100$$

donc $k = 5$

Alors

$$[3.49] \rightarrow v_{\max} = \frac{\Delta}{2 (15.10^{-4} \Delta + 9.10^{-2})}$$

et

$$v'_{\max} = \frac{9.10^{-2}}{2 (15.10^{-4} \Delta + 9.10^{-2})^2}$$

Les représentation de $v_{\max}(\Delta)$ et $v'_{\max}(\Delta)$ sont données ci-dessous pour Δ variant de 1 à 250.

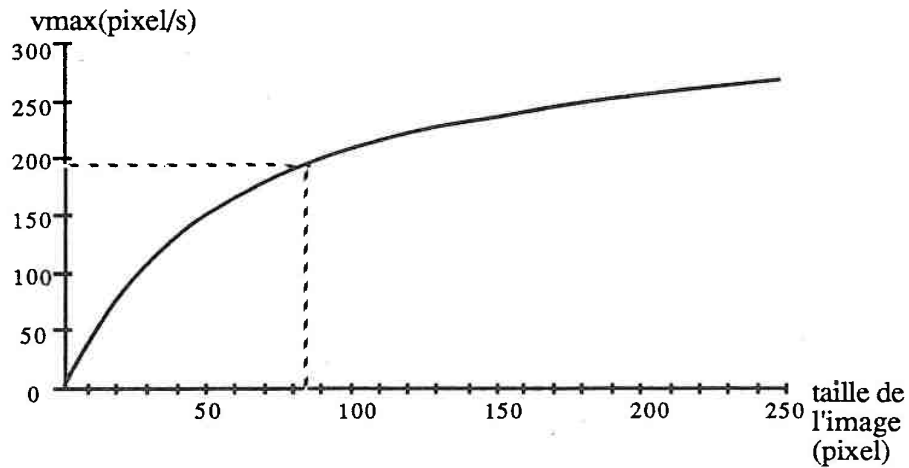


Figure 61 : $v_{\max} = f(\Delta)$ pour $k = 5$

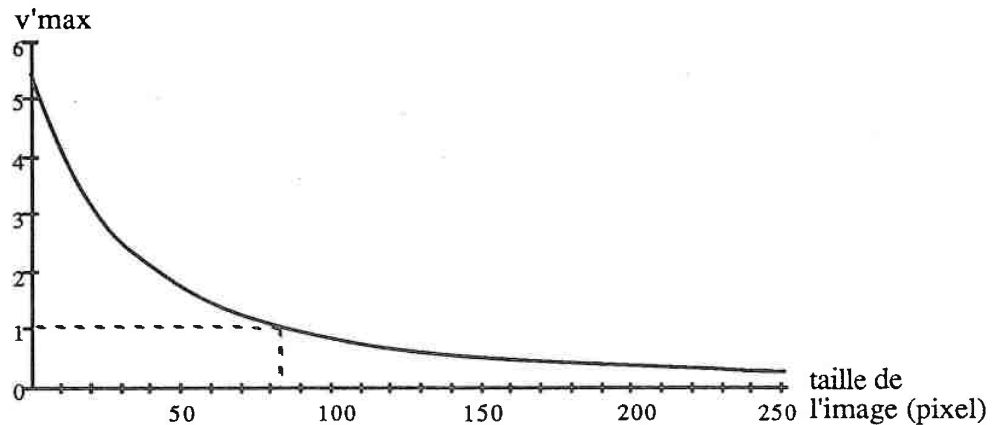


Figure 62 : $v'_{\max} = f(\Delta)$ pour $k = 5$

On détermine à partir de ces courbes une valeur limite de Δ pour laquelle l'accroissement de $v_{\max}$ devient moins sensible. La valeur de Δ correspondant à $v'_{\max} = 1$ paraît une limite commode et justifiée par le changement d'allure de $v_{\max}$ dans cette zone. Ainsi, Δ_{lim} est définie par :

$$[3.51] \quad \Delta_{\text{lim}} / v'_{\max}(\Delta_{\text{lim}}) = 1$$

$$\text{Pour } \Delta_{\text{lim}} = 83 \quad N = 5 \times 83 = 415$$

$$T_c = 3 \cdot 10^{-4} \times 415 + 9 \cdot 10^{-2} = 0,21 \text{ s}$$

Le temps de traitement requis pour la vitesse maximale se situe bien dans l'intervalle des essais (fig. 58), les hypothèses sont donc validées. La fenêtre sera choisie avec une largeur de 83 pixels et la vitesse maximale autorisée sera de 192 pixels / s (la transposition en cm dépend de l'éloignement de la caméra par rapport au plan objet).

III - 2 - 4 - Modèle de corrélation

Le programme final de corrélation est constitué d'une phase préliminaire, incluant les critères définis dans cette partie. En premier lieu, l'écart type de l'image (σ) est calculé à partir de [3.5]. Les seuils sont ensuite établis en proportion de σ , les valeurs expérimentalement performantes sont σ ; $1,3 \sigma$; $1,6 \sigma$

Dans le cas d'une poursuite sans prédiction, il est intéressant de déterminer la taille de fenêtre Δ_{lim} par la formule [3.51] appliquée à partir d'une mesure expérimentale du nombre de points dépassant les seuils choisis pour une taille de fenêtre arbitraire. Le temps maximal disponible pour la corrélation peut alors être calculé à partir de la vitesse maximale envisagée pour la cible. Et la formule [3.45] permet alors de tester si la poursuite est réalisable. Dans le cas où elle ne l'est pas, un réajustement des paramètres du seuillage devra être effectué.

Par la suite, la poursuite sera réalisée avec une phase de prédiction, exploitant les paramètres de la trajectoire du mobile, le temps disponible pour la corrélation ne sera donc plus lié à la vitesse maximale, mais aux variations du mouvement.



CHAPITRE IV

IV - MISE EN OEUVRE DE LA POURSUITE - FILTRAGE DE KALMAN

IV - 1 - Dispositif expérimental

IV - 1 - 1 - Description des différents éléments

La poursuite est réalisée pour un mobile solidaire d'un chariot dont les déplacements sont produits par un moteur pas à pas. Nous avons mis en place une commande optionnelle du moteur par un programme pascal et des routines assembleur implémenté sur un système à base de MC 68000. Il est ainsi possible de donner des trajectoires de différentes natures à l'objet : mouvement uniforme, uniformément accéléré, périodique, à accélération sinusoïdale, bruitée ... De plus, les valeurs envoyées sur les timers du moteur étant enregistrées, la trajectoire du mobile est connue avec précision. Une comparaison valable peut donc être effectuée, entre les résultats délivrés par le système de poursuite et la trajectoire réelle de la cible.

Le logiciel de ce dispositif expérimental est composé de deux parties : génération de trajectoire d'une part et analyse du mouvement d'autre part, qui sont implantés sur deux ensembles matériels distincts.

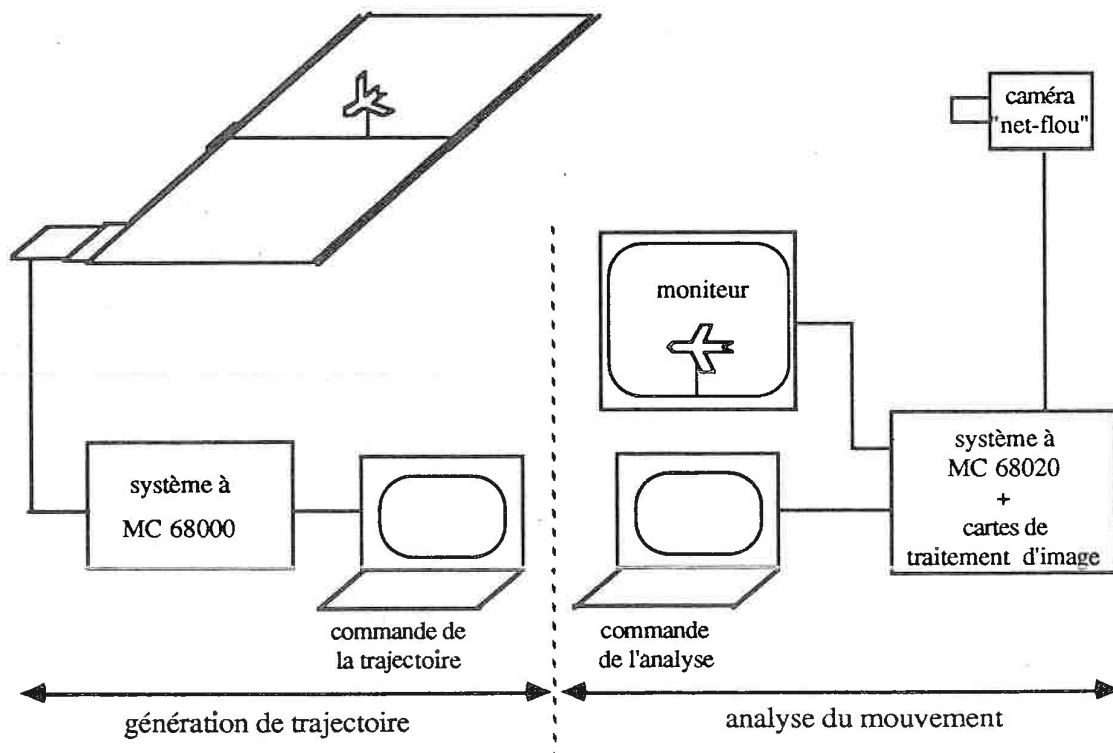


Figure 63 : Représentation du dispositif expérimental

Le système à MC 68020 est lié à deux cartes "Max Vidéo" de traitement d'image, l'une (digimax) servant à numériser les données issues de la caméra, l'autre (framestore) constituant la mémoire d'image. Le contenu de cette dernière est visualisé sur le moniteur.

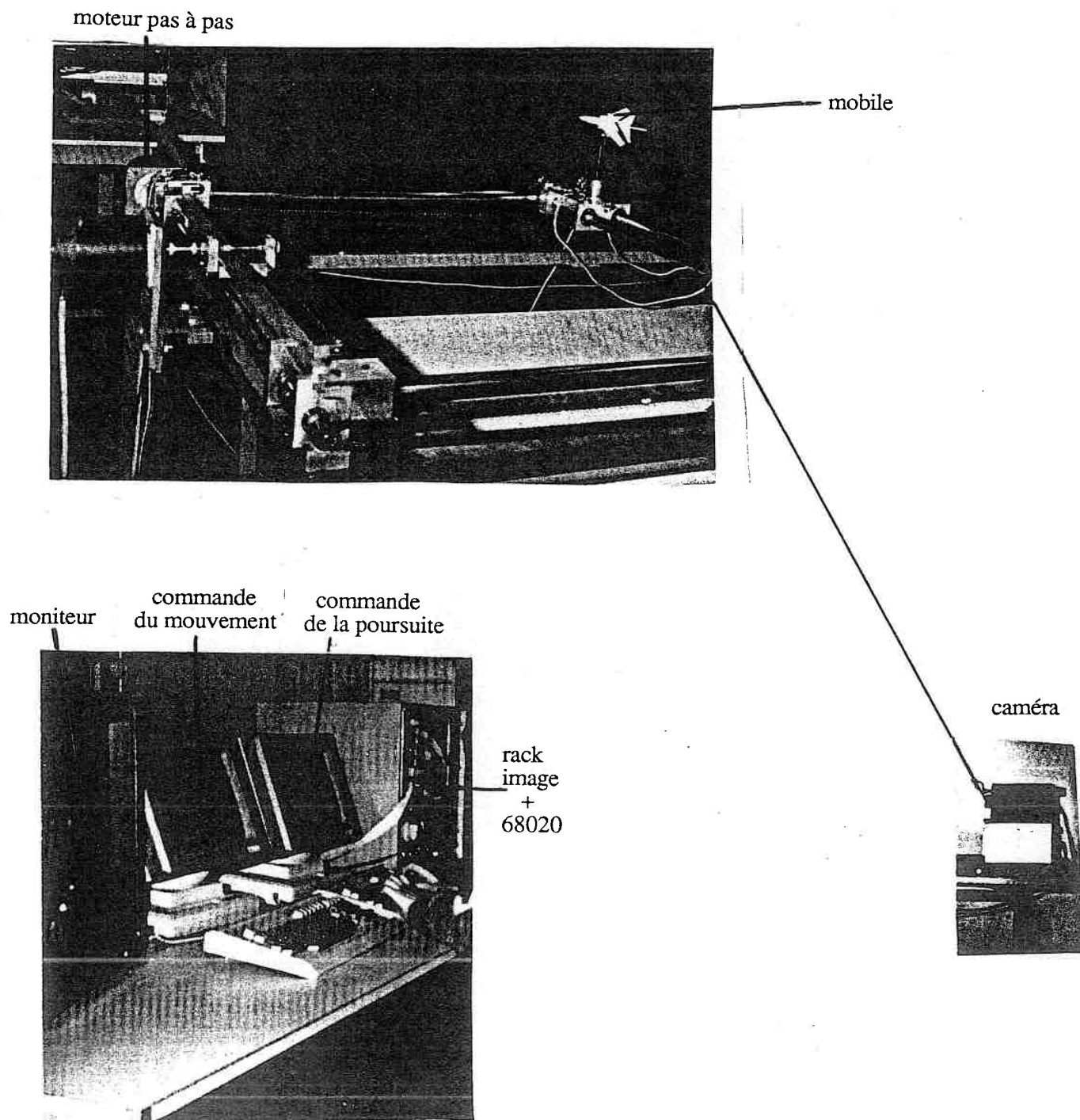


Figure 64 : Photographies du dispositif

IV - 1 - 2 - Premiers essais de poursuite

Le programme de corrélation (cf § II - 2 - 3 - 4) est mis en place sur le système à MC 68020.

La caméra est placée face à la scène suivante :

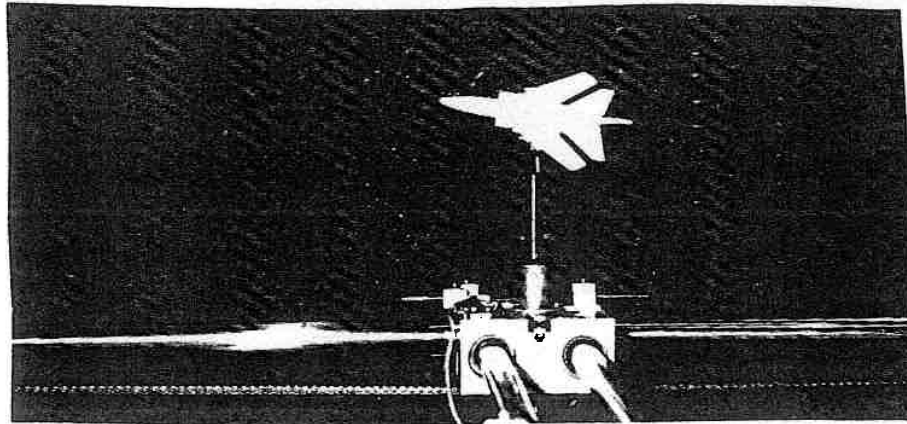


Figure 65 : Scène visualisée

La poursuite est initialisée manuellement en délimitant la fenêtre de référence qui sera prise comme modèle.

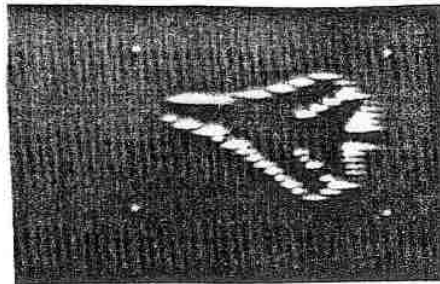


Figure 66 : Fenêtre de référence pour l'avion

Une fois le modèle enregistré, les déplacements de l'objet sont commandés par le générateur de trajectoire ; simultanément le programme de corrélation par restriction est mis en action pour la poursuite du mobile. Après chaque acquisition d'image, la fenêtre de référence est localisée. Le domaine de recherche est ensuite positionné autour de l'emplacement trouvé, et le modèle est réactualisé pour la localisation de la cible à l'instant suivant.

La figure ci-dessous montre l'image dans laquelle le mobile est trouvé, et en surimpression, l'image binaire du nouveau modèle pour le niveau de seuil le plus bas.

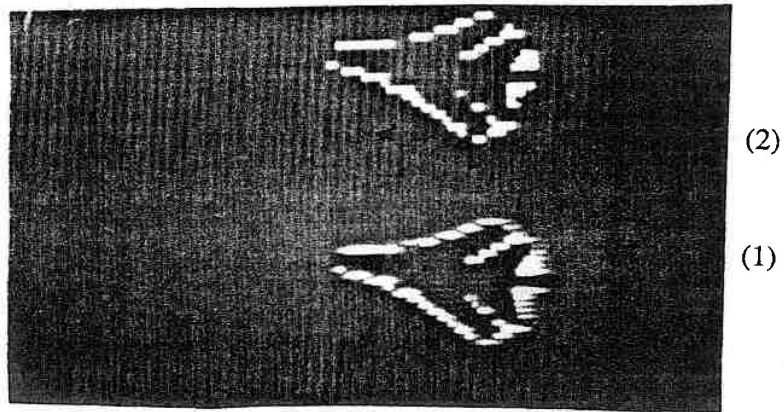


Figure 67 : Image de recherche (1) + nouvelle image du modèle (2)

L'évaluation de la position de la cible dans chaque image, et du temps écoulé entre chaque acquisition permet de connaître la position, la vitesse et éventuellement l'accélération du mobile. De plus, la réactualisation du modèle autorise des variations de luminosité à condition qu'elles ne soient pas trop brutales. Mais le positionnement de l'aire de recherche autour du mobile limite la vitesse perceptible suivant la formule décrite au § III - 2 - 3 :

$$v_{\max} = \frac{\Delta}{2I_T}$$

Une vitesse supérieure entraînera automatiquement le décrochage de la poursuite.

Un autre cas d'échec est rencontré si le mobile est caché (passage derrière un obstacle) à un instant d'acquisition, la réactualisation du modèle est alors réalisée à partir d'une zone quelconque de l'image, et il est impossible de retrouver le mobile.

Enfin si l'objet est partiellement masqué, le nouveau modèle comprendra la cible et le masque ; à l'étape suivante, il est possible que la fenêtre de référence soit localisée sur le masque donnant l'illusion d'une vitesse nulle pour le mobile. Il suffit pour cela que le pic de corrélation du à l'obstacle soit plus important que celui correspondant à l'objet.

Exemple :

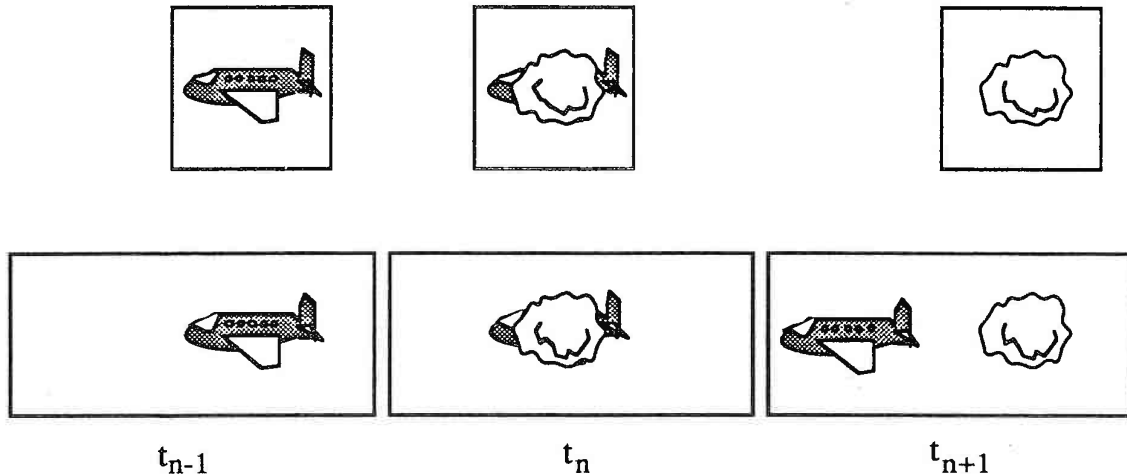


Figure 68 : Décrochage par occlusion

Afin de diminuer les conséquences facheuses de ces cas de figure, d'augmenter la vitesse limite de poursuite et de tenir compte des parasites entachant les mesures, nous mettons en place un module de filtrage des résultats de la corrélation. Parmi les filtres utilisés par ailleurs en poursuite (filtre polynomial, de Wiener, $\alpha - \beta$), nous retiendrons le filtre de Kalman pour son caractère adaptatif, et parcequ'il autorise des variations de la période d'échantillonnage ; cette dernière propriété est intéressante pour la corrélation restreinte où le temps de traitement peut être variable.

IV - 2 - Filtrage de Kalman

Le filtrage est pris ici dans le sens d'une modélisation mathématique du système en vue d'exploiter ses mesures passées et présentes pour obtenir des informations plus précises sur ses variables passées, présentes, ou futures.

IV - 2 - 1 - Notations générales

Il s'agit de l'estimation de l'état d'un processus à partir des mesures bruitées le représentant. Soient, au k ème instant d'échantillonnage, les processus suivants :

X_k = état du système à valeurs dans $\mathbb{R}^n$

Y_k = mesures sur le système, à valeurs dans $\mathbb{R}^m$

w_k = bruit d'état, à valeurs dans $\mathbb{R}^n$

v_k = bruit de mesure, à valeurs dans $\mathbb{R}^m$.

Les bruits d'état et de mesure aux différents instants sont représentés par des suites de variables aléatoires mutuellement indépendantes, gaussiennes, de moyenne nulle:

$$[4.1] \quad E[w_k] = 0 \quad E[w_k w_j^T] = Q_k \delta_{kj}$$

$$[4.2] \quad E[v_k] = 0 \quad E[v_k v_j^T] = R_k \delta_{kj}$$

$$[4.3] \quad E[v_k w_j^T] = 0$$

avec δ_{kj} : symbole de Kronecker

L'évolution du système est décrite par :

$$[4.4] \quad X_{k+1} = \phi_{k+1,k} X_k + w_k$$

et observée par :

$$[4.5] \quad Y_k = H_k X_k + v_k$$

avec

$\phi_{k+1,k}$: matrice de transition régulière du système, de dim (n,n).

H_k : matrice d'observation de dim (m,n).

On se propose de calculer l'estimé optimal $\tilde{X}_k$ et le prédit $\hat{X}_{k+1}$, à partir de k mesures données sur le système, cet estimé est linéaire sans biais et minimise la variance de l'erreur d'estimation.

Notons :

$\tilde{X}_k$: estimé optimal de X_k

$\hat{X}_{k+1}$: prédit de X_{k+1} à partir de k mesures

P_k : matrice de covariance de l'erreur de prédiction

S_k : matrice de covariance de l'erreur d'estimation

K_k : matrice de gain de Kalman.

Les équations de propagation sont :

$$[4.6] \quad \hat{X}_k = \phi_{k,k-1} \tilde{X}_{k-1}$$

$$[4.7] \quad P_k = \phi_{k,k-1} S_{k-1} \phi_{k,k-1}^T + Q_k$$

Les équations d'estimation sont :

$$[4.8] \quad K_k = P_k H_k^T [H_k P_k H_k^T + R_k]^{-1}$$

$$[4.9] \quad \tilde{X}_k = \hat{X}_k + K_k [Y_k - H_k \hat{X}_k]$$

$$[4.10] \quad S_k = [I - K_k H] P_k$$

IV - 2 - 2 - Application à la poursuite de cible

Pour la poursuite de cible, le vecteur d'état est constitué des variables décrivant la trajectoire du mobile, l'exploitation du filtre de Kalman pour ce type de problème a déjà été traitée. A titre d'exemple dans [LEG - 82] les auteurs prennent pour vecteur d'état la position, la vitesse et l'accélération d'un avion dont ils déterminent le mouvement, le filtre assiste la poursuite dans le cas où une partie de l'avion est absente de l'image.

Les choix du vecteur d'état X , du vecteur de mesure Y , de la matrice de transition ϕ , et des covariances des bruits R et Q dépendent des types de mouvements envisagés et des conditions d'observation. Les difficultés de mise en oeuvre résident dans la connaissance des bruits, et du modèle de la trajectoire. En effet, ceux-ci ne se confondent pas toujours avec le cadre prévu par la théorie et un décalage permanent entre la réalité physique et les équations tentant de la décrire peuvent conduire à des phénomènes de divergence du filtre.

Par exemple, si le bruit d'état est sous évalué, un écart de la cible par rapport à la trajectoire prévue sera filtré comme un changement de direction et la recherche de l'objet à l'étape suivante se fera dans une zone erronée. En revanche, si ce même bruit est surévalué, le phénomène inverse apparaît et un réel changement de direction sera interprété comme un bruit permanent, engendrant une perception biaisée du mouvement. Dans les équation [4.6] [4.7] [4.8] [4.9] [4.10], il apparaît qu'une forte valeur du bruit d'état (Q_k) donne une grande importance à chaque nouvelle mesure alors que si le bruit de mesure (R_k) est prépondérant, chaque nouvelle donnée sera moins déterminante que le modèle décrivant l'évolution du système ($\phi_{k,k-1}$).

Les conditions d'application ne permettent pas forcément de faire un choix justifié pour ces valeurs. Décrivons brièvement quelques études menées dans ce domaine. Le modèle de base de [XUE - 86] suppose que le mobile se déplace à vitesse constante ; R_k et Q_k sont évalués simultanément au filtrage par un algorithme récurrent. Lorsqu'une manœuvre est décelée, un terme d'accélération est introduit dans le vecteur d'état. Il s'agit dans cet exemple de palier aux limites de la modélisation par des critères complémentaires d'estimation.

Dans une autre optique, les auteurs de [SAS - 87] proposent un calcul du gain optimal de Kalman K en regroupant plusieurs mesures, ce lorsque le modèle du système est bien connu et le rapport signal sur bruit indéfinissable mais peu variable dans le temps ; le gain K n'est pas déduit des équations [4.7] [4.8] et [4.10] mais est évalué à partir des séquences d'innovation. Par une approche analogue, un dispositif de poursuite de cibles manoeuvrantes est décrit dans [PER - 79 et 80] ; les variables d'état sont : les écarts de position, les vitesses angulaires, et les accélérations en site et latérales de la cible. Les constantes de temps du système sont supposées nettement supérieures aux périodes d'échantillonnage ; ainsi un gain optimal K est calculé pour plusieurs étapes par l'algorithme d'approximation stochastique de Hampton, en assurant la décorrélation des innovations. Le filtre établi permet le calcul des accélérations cinématiques et des ordres de guidage à partir des mesures des écarts en position.

Les procédures mentionnées recherchent itérativement les valeurs de R_k Q_k ou K_k de manière à éviter le phénomène de divergence du filtre lorsque les paramètres du système varient. Leur application suppose donc que les transformations survenant sur le modèle sont à constante de temps très grande devant la période d'échantillonnage.

Une autre approche, applicable pour des variations plus rapides du système, recherche un comportement sous optimal du filtre en évitant le phénomène de divergence par limitation de l'erreur d'estimation. Nous choisissons dans cette optique un modèle simple, en considérant le mobile animé par un mouvement uniforme [HOR - 89 b].

Si x est sa position :

$$[4.11] \quad X_k = \begin{bmatrix} x_k \\ x'_k \end{bmatrix} \quad \phi_{k,k-1} = \begin{bmatrix} 1 & T_{k-1} \\ 0 & 1 \end{bmatrix}$$

$$H = [1 \quad 0]$$

Afin d'éviter la perte de la cible lorsqu'une accélération apparaît, nous choisissons une valeur variable pour l'écart type du bruit d'état.

La détection d'une manœuvre est assurée par l'observation de la suite des innovations :

$$[4.12] \quad E(k) = y_k - H \hat{X}_k$$

Le processus d'innovation révèle l'écart existant entre représentation modélisée (grandeurs prédites) et représentation réelle (grandeurs mesurées) du phénomène.

On utilisera la distance $D(k) = |E(k)|$; par comparaison de D à un seuil de décision s , le système adapte le niveau de filtrage dès qu'une forte probabilité de manœuvre est observée. Pour maintenir la poursuite dans ce cas, on se propose d'augmenter le coefficient Q à une valeur q_m qui sera déterminée en fonction des variations limites attendues. Par ce moyen, les coefficients du gain de Kalman K_k sont augmentés donnant ainsi une importance plus grande aux innovations.

Le schéma de principe du filtrage adapté est alors le suivant :

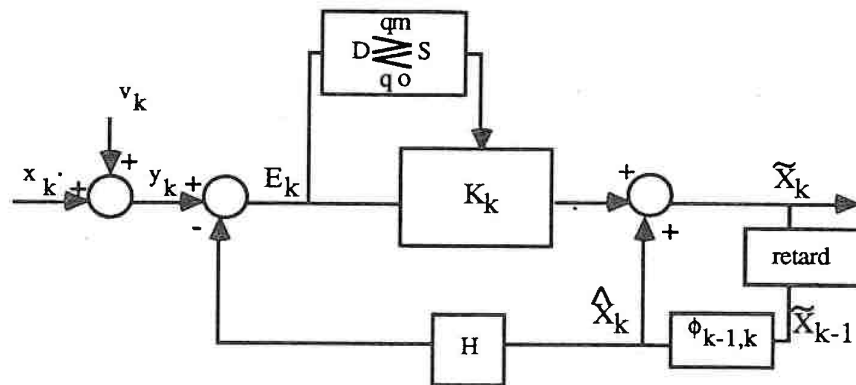


Figure 69 : Représentation du filtre adapté

L'initialisation est assurée par la mesure des positions aux premiers instants d'échantillonnage et détermination des valeurs S_1 et Q_1 suivant les procédures décrites dans [BOZ - 83], en supposant que le mobile ne manoeuvre pas lors des premières acquisitions.

$$[4.13] \quad \tilde{X}_1 = \begin{pmatrix} y_1 \\ \frac{y_1 - y_0}{T_1} \end{pmatrix}$$

$$[4.14] \quad S_1 = \begin{pmatrix} \sigma^2 & \sigma^2 / T_1^2 \\ \sigma^2 / T_1^2 & 2\sigma^2 / T_1^2 \end{pmatrix}$$

$$[4.15] \quad Q_1 = \begin{pmatrix} q_0 T_1^3 / 3 & q_0 T_1^2 / 2 \\ q_0 T_1^2 / 2 & q_0 T_1 \end{pmatrix}$$

avec σ : écart type du bruit de mesure.

IV - 2 - 3 - Simulation de la poursuite

Le choix du modèle sans terme d'accélération pour décrire l'évolution du mobile ($\phi_{k-1,k}$) est guidé par les limites, en temps d'exécution, imposées par l'application. Afin de justifier notre option et de visualiser l'effet du filtre, nous réalisons des essais par simulation. Tout d'abord exposons les courbes données dans [CIC - 87]. Il s'agit de filtrer les valeurs des positions (fictives) d'une cible en mouvement rectiligne bruité avec changement de direction. La première figure [fig. 70] illustre la divergence du filtre non adapté : il interprète les mesures comme si elles étaient issues d'un mobile se déplaçant à vitesse constante et considère les changements de direction comme des bruits. Les trois courbes suivantes illustrent la procédure d'adaptation due aux changements de valeur de q (q_0, q_m), lorsque D dépasse le seuil s .

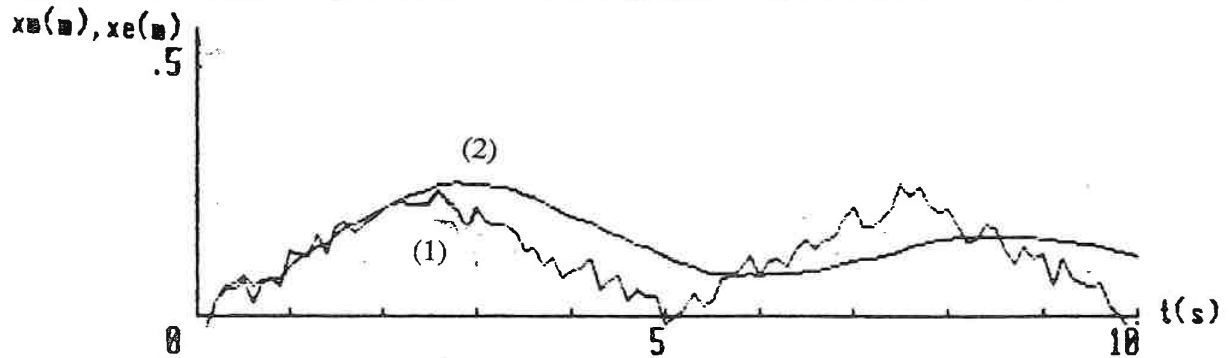


Figure 70: Position mesurée (1) et estimée (2) sans adaptation

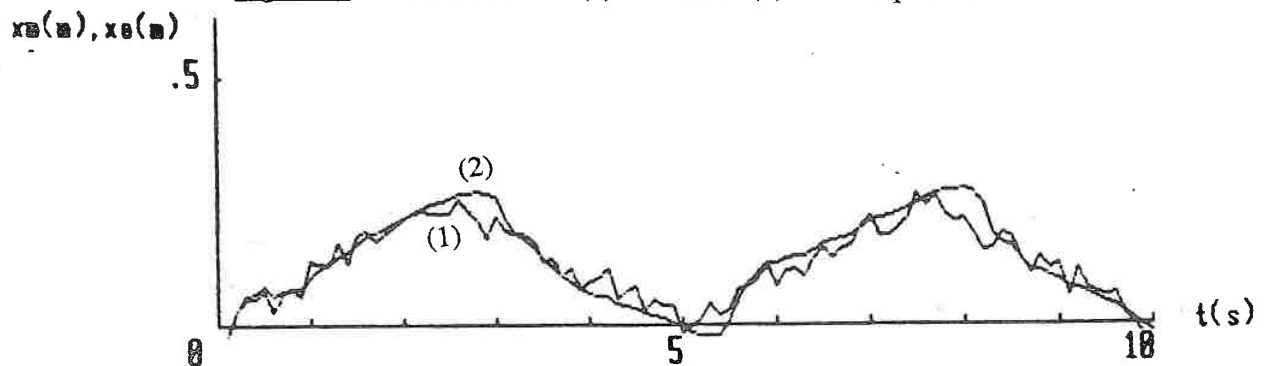


Figure 71 : Position mesurée (1) et estimée (2) avec adaptation

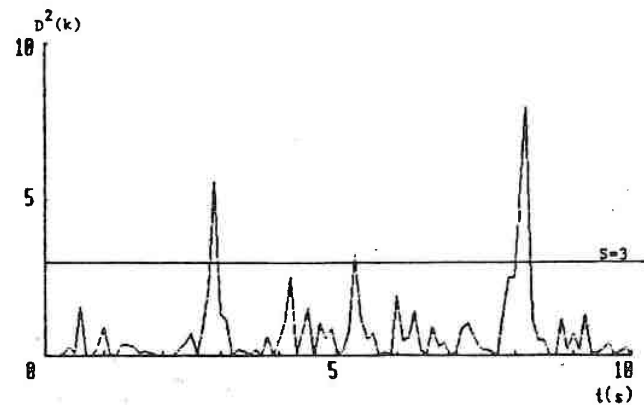


Figure 72 : Distance de décision

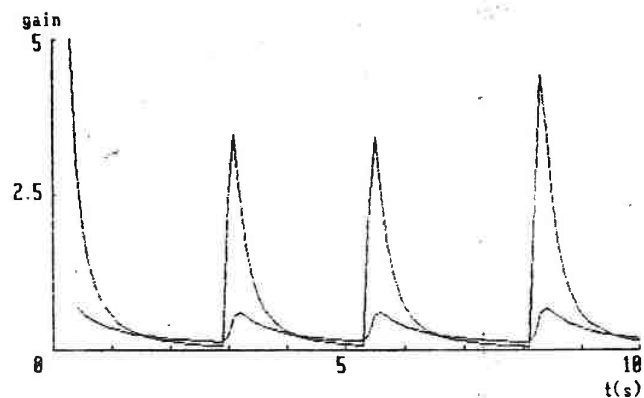


Figure 73 : Gains de Kalman

Le gain de Kalman est amplifié dans les zones de manoeuvre, assurant une analyse adaptée des données, après quoi les calculs se poursuivent suivant les modalités classiques. Pour le réglage des coefficients q_0 , q_m et s , nous verrons pratiquement qu'ils dépendent des conditions expérimentales dans lesquelles la poursuite est réalisée, et des performances souhaitées. Cependant afin de définir leurs ordres de grandeur, nous testons sur des trajectoires réelles mesurées par analyse d'image et enregistrées, les effets du filtrage. Notons que cette simulation est un peu biaisée car l'effet du filtre n'influe pas sur la mesure suivante, ce qui est le cas lorsque la position prédite détermine l'emplacement du domaine de recherche.

La trajectoire choisie est sinusoïdale et volontairement bruitée, la figure 74 illustre l'effet du filtre non adapté avec $q_0 = 0,5$. Le facteur de bruit est totalement éliminé, mais il existe un temps de retard d'environ 1 seconde entre les valeurs mesurées et filtrées, un tel phénomène mènerait à la perte de la cible.

Par le filtrage adapté en gardant $q_0 = 0,5$, nous recherchons s et q_m de manière à filtrer partiellement le bruit sans engendrer un retard trop marquant. Les valeurs les plus satisfaisantes sont : $q_m = 10$ $s = 0,1$ (fig. 75), elles serviront de base pour le système final .

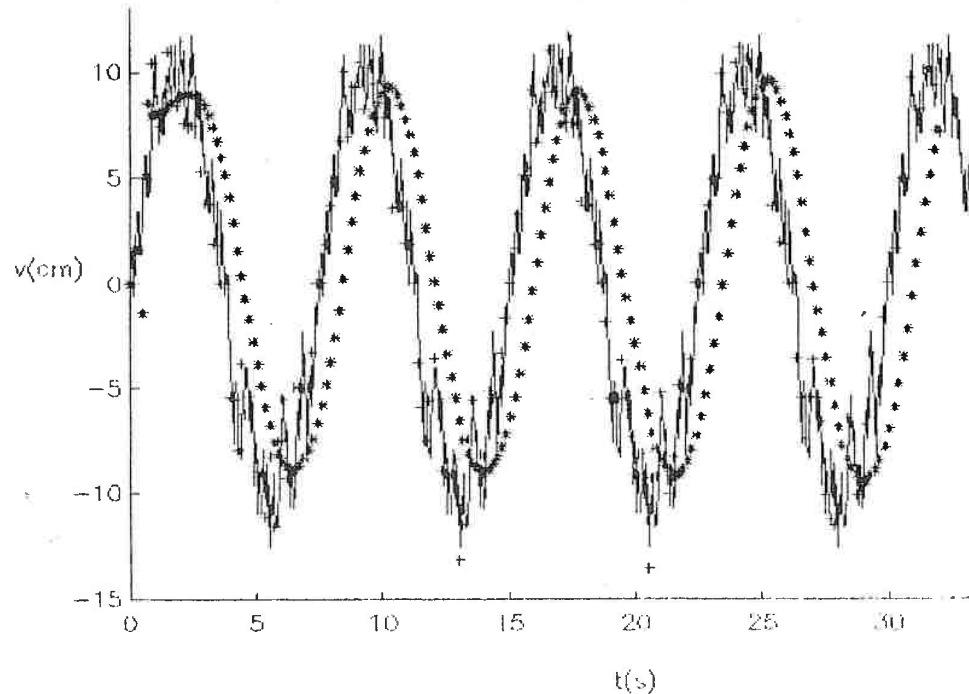


Figure 74

- Trajectoire générée par les moteurs (volontairement bruitée)
- + : mesures obtenues par analyse d'image
- * : résultats donnés par le filtrage

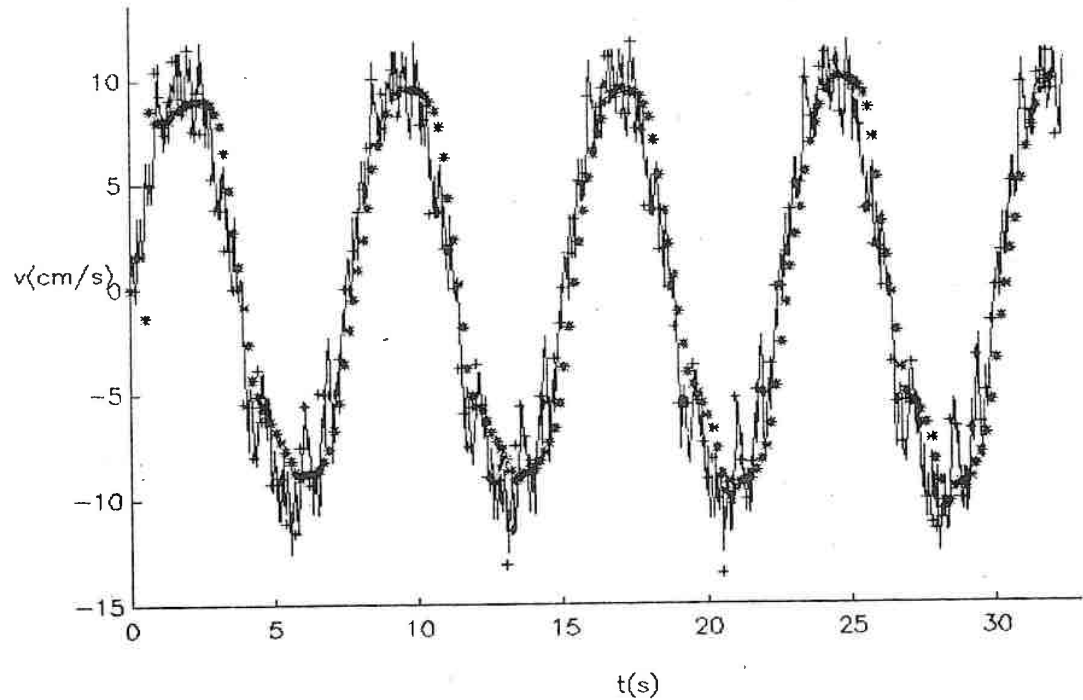


Figure 75 :

- Trajectoire générée par les moteurs (volontairement bruitée)
- + : mesures obtenues par analyse d'image.
- * : résultats donnés par le filtrage

IV - 2 - 4 - Mise en oeuvre sur le dispositif - Résultats

L'estimation et la prédiction sont maintenant intégrées au système de poursuite, le programme est implémenté sur le coprocesseur rapide (68881), le module d'analyse d'image donne la position de l'image de référence (de 25 sur 25 pixels) dans l'aire de recherche (de 25 sur 40 pixels) qui est ensuite déplacée en relation avec la position prédite pour l'étape suivante. Des essais préliminaires montrent que le seuil de basculement du bruit d'état ($s = 0,1$) obtenu en simulation est mal approprié pour une fenêtre de recherche de si petite taille ; la décision $q = q_m$ ne survenant que lorsque l'objet est sorti de la fenêtre. C'est pourquoi nous choisissons $s = 0,01$ et gardons les autres paramètres à leurs valeurs d'essais qui semblent convenir. L'amélioration apportée par le filtrage concerne tout d'abord la possibilité de suivre un mobile animé d'une vitesse supérieure à la vitesse limite définie précédemment (Fig. 76-77).

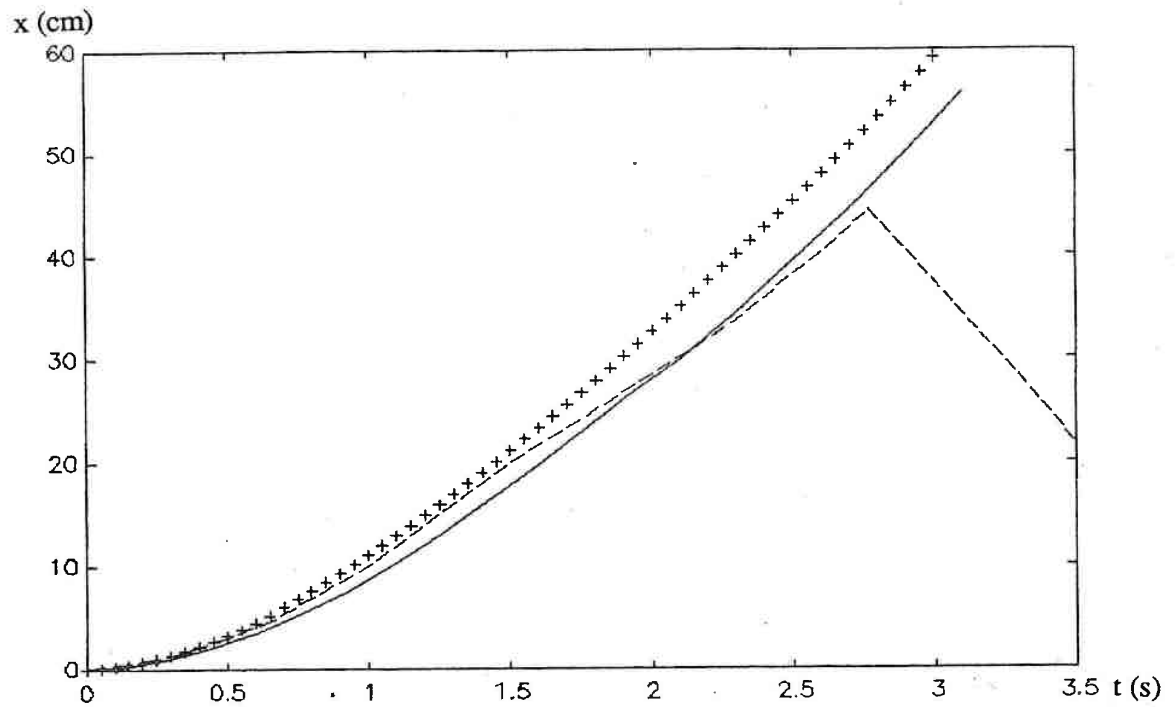


Figure 76 : + position programmée

position mesurée ----- sans filtrage _____ avec filtrage

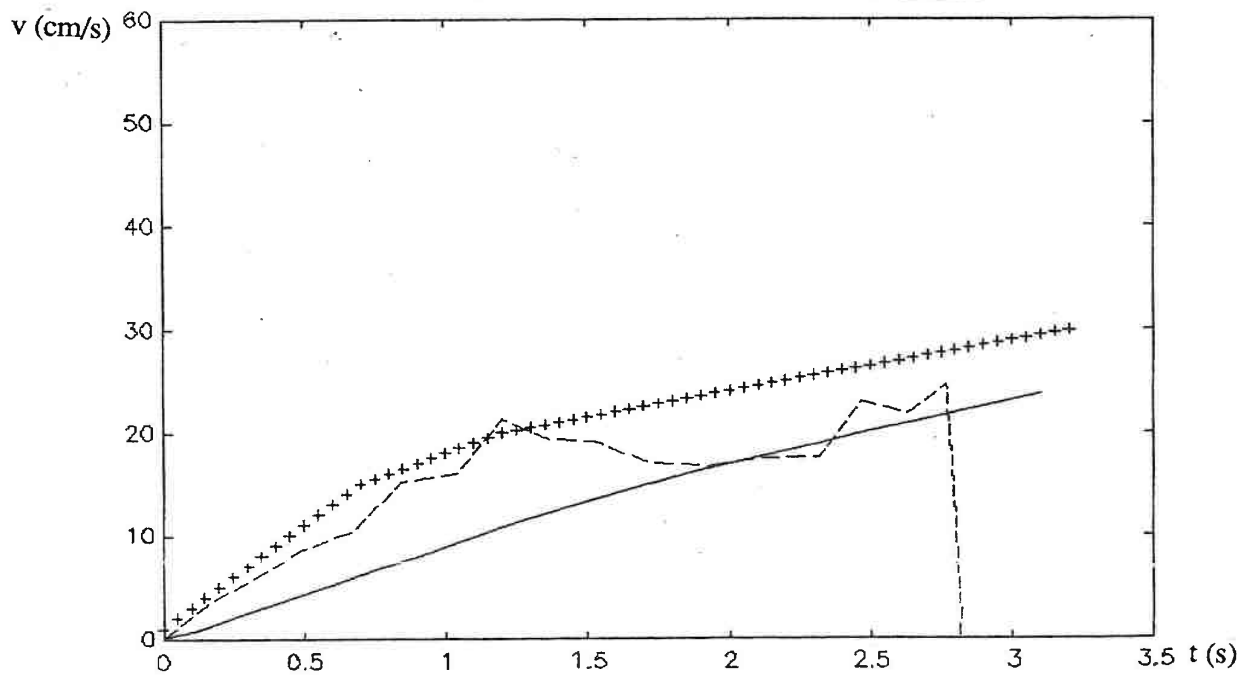


Figure 77 : + vitesse programmée

vitesse mesurée ----- sans filtrage _____ avec filtrage

Les résultats représentés sont obtenus pour une période d'échantillonnage de 0,08 s.

Remarque : Le temps requis pour le filtrage est de 0,3 ms valeur négligeable devant le temps de corrélation évalué au chapitre III. Mais les périodes d'échantillonnage envisagées pour les essais sont plus longues, nous les avons volontairement dilatées afin d'observer avec les moyens dont nous disposons des phénomènes ne se produisant normalement qu'à des vitesses plus grandes. C'est aussi pour cette raison que nous avons réduit la fenêtre de recherche. Les résultats obtenus dans cette partie sont donc essentiellement qualitatifs.

L'apport du filtre est aussi remarquable lorsque la cible est masquée pendant plusieurs instants d'échantillonnage, les équations de propagation ($\hat{X}_k = \phi_{k,k-1} \tilde{X}_{k-1}$) permettent de continuer la poursuite un certain temps sans aucune mesure nouvelle, alors que dans le même cas une poursuite sans filtrage s'égare. Pour les figures 78 - 79 - 81, la cible est animée d'un mouvement rectiligne uniforme lorsqu'elle est cachée à 20 cm de son point de départ, les périodes d'échantillonnage sont en moyenne de 0,6 s.

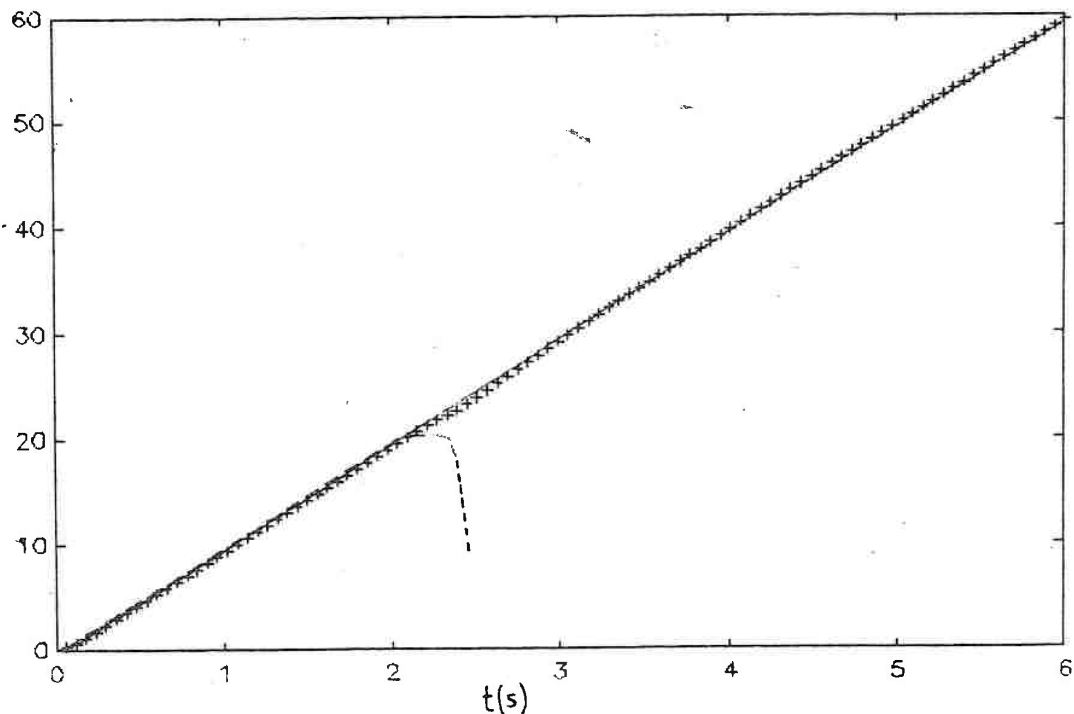


Figure 78 : — Position commandée pour l'objet en cm
 ---- Position perçue sans filtrage (décrochage)
 + Position perçue avec filtrage

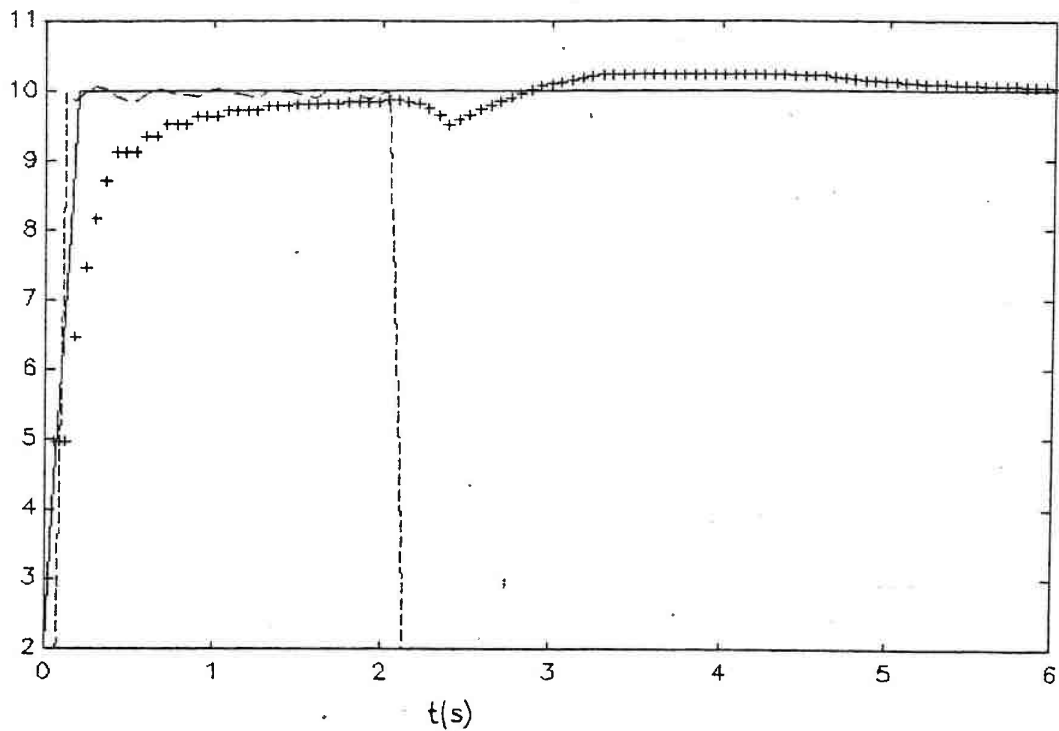


Figure 79 : — Vitesse commandée (en cm/s)
 ---- Vitesse perçue sans filtrage
 + Vitesse perçue avec filtrage

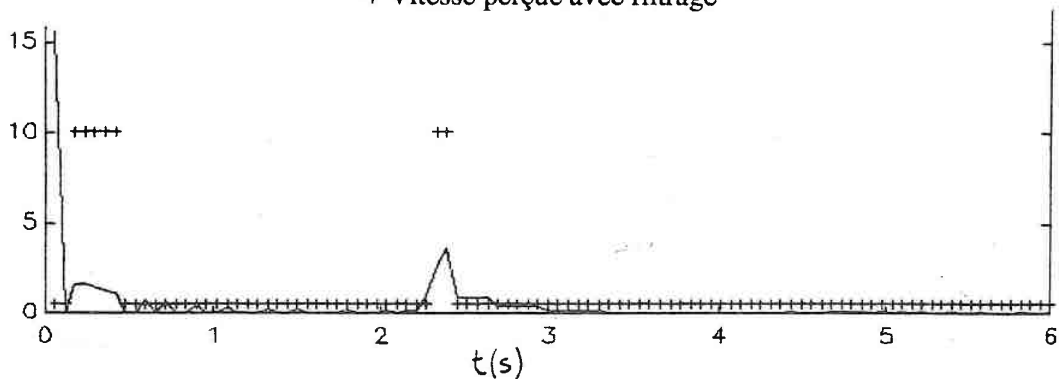


Figure 80 : — D distance déterminant les basculements (x 100);
 + q ($q_0 = 0,5$ $q_m = 10$).

Nous recherchons ensuite l'effet des paramètres q_m et s sur l'analyse des déplacements. Pour cela, un mouvement à vitesse variable est programmé pour le chariot, et un obstacle masque la cible à 30 cm de son point de départ. les enregistrements suivants sont établis avec deux valeurs différentes de q_m les courbes en trait plein sont obtenues pour $q_m = 20$ et celles en pointillés avec $q_m = 10$.

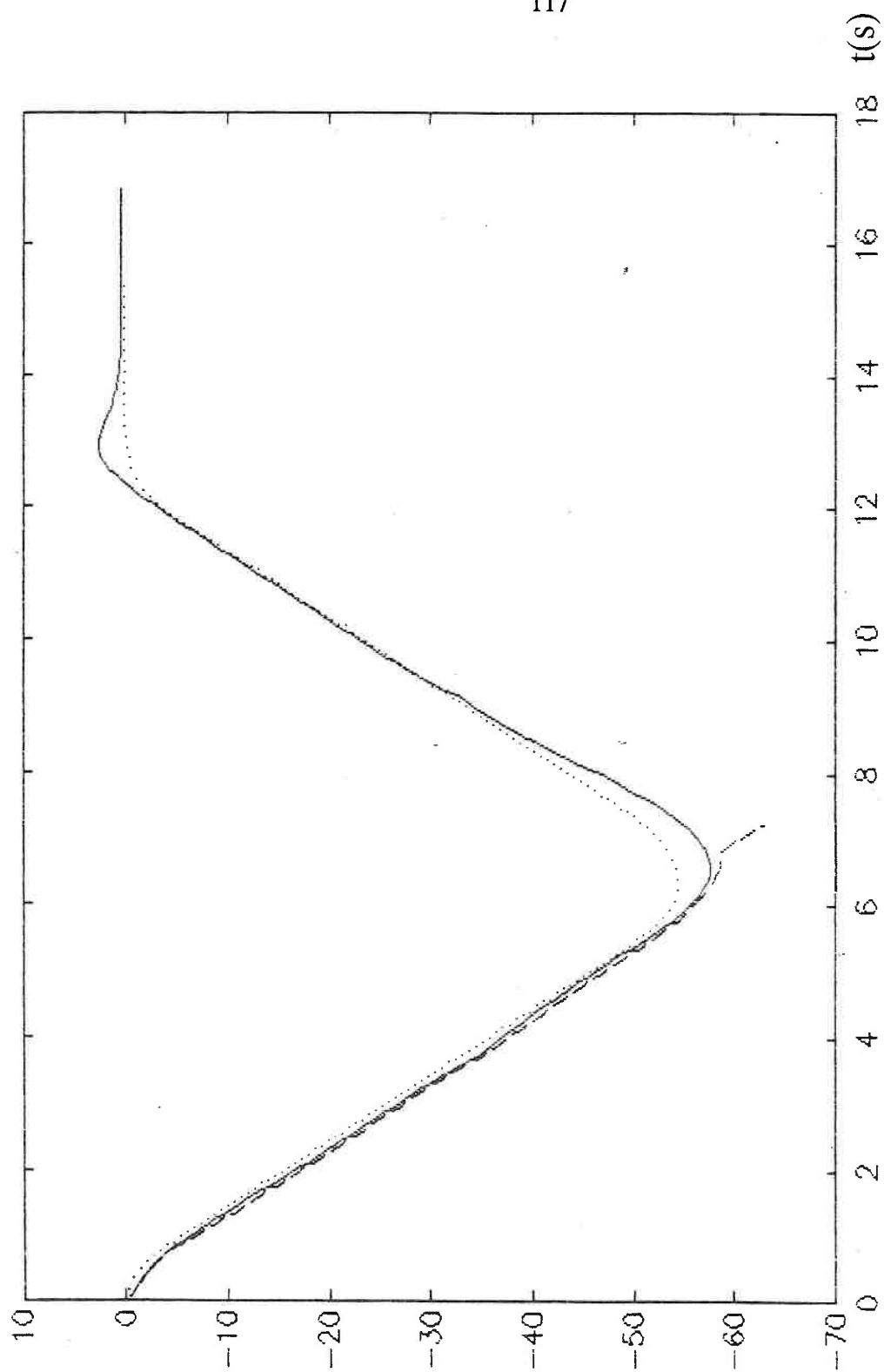


Figure 81
..... x commandé en cm
----- x calculé

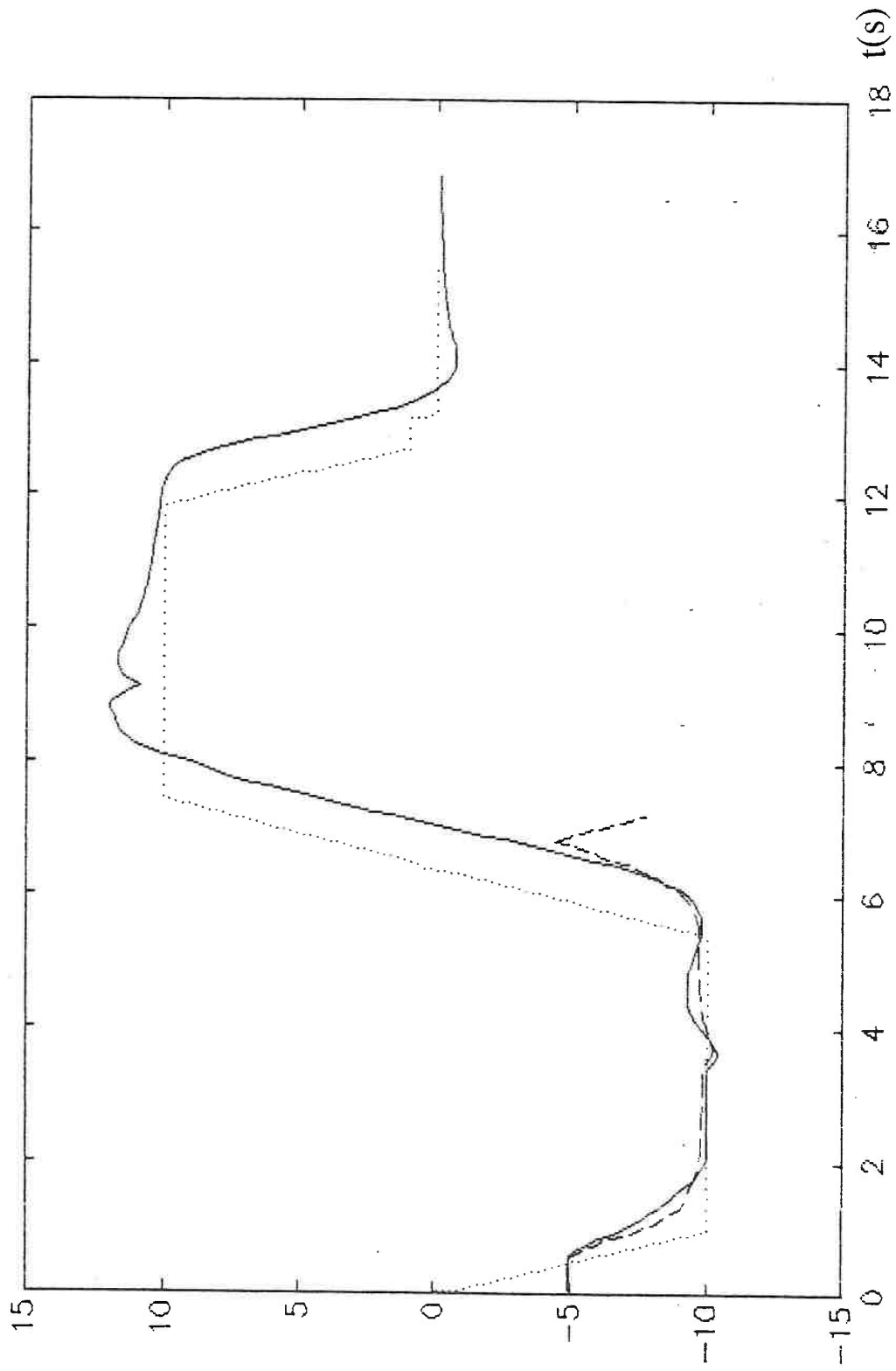


Figure 82
..... $v_{\text{commandé}}$ en cm/s
----- $v_{\text{calculé}}$

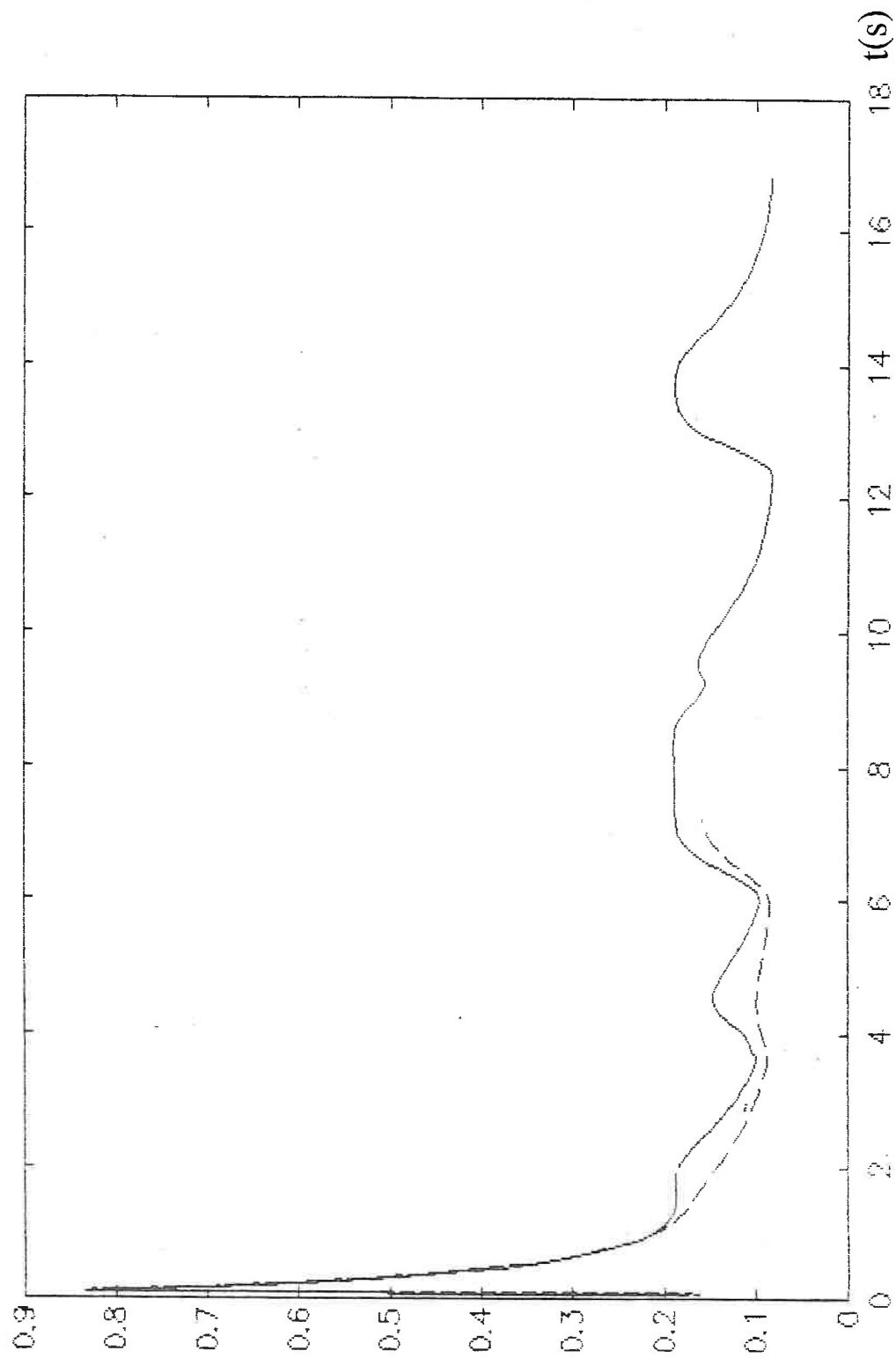


Figure 83 : K_1 gain de Kalman en position

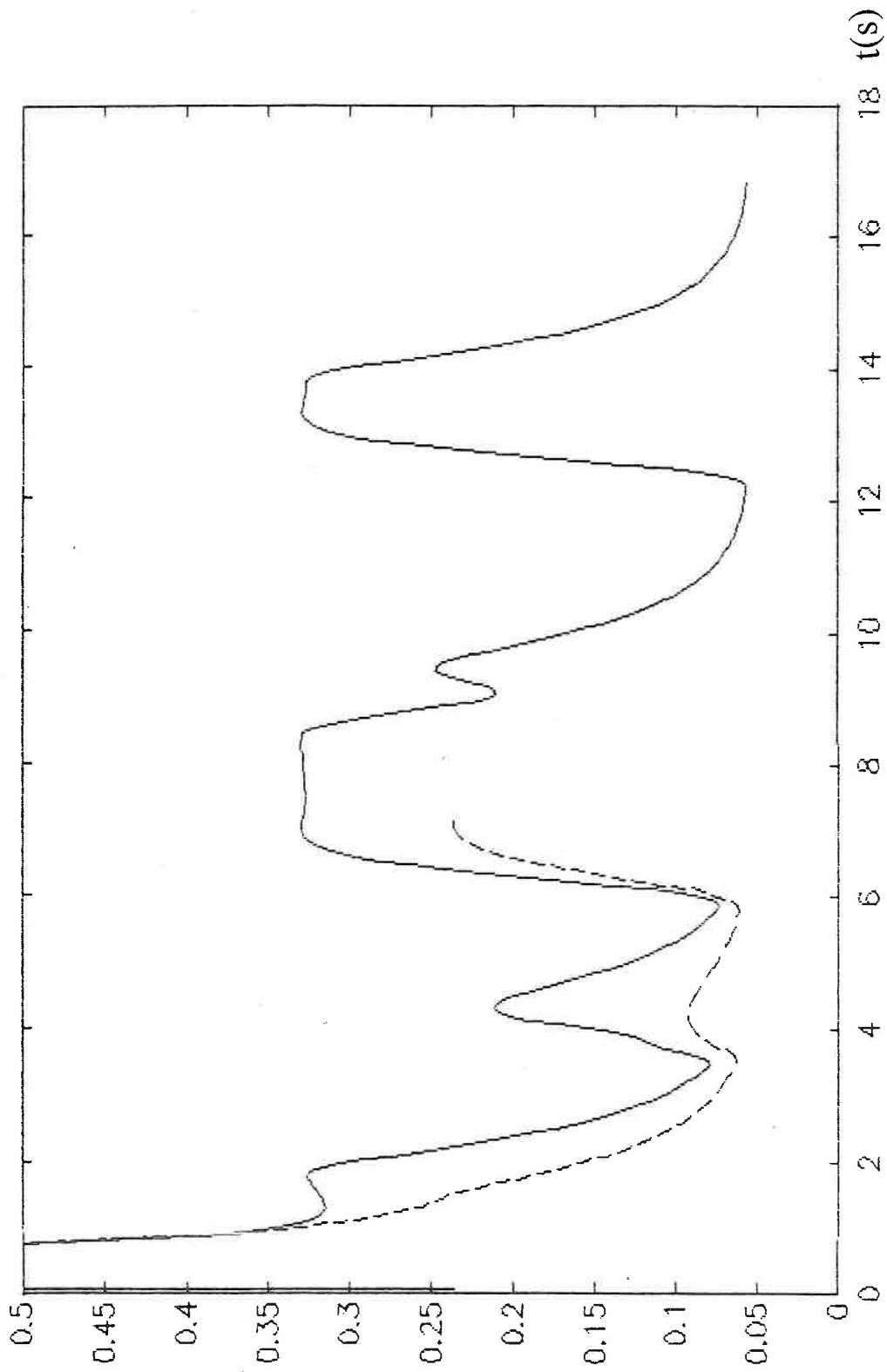


Figure 84 : K_2 gain de Kalman en vitesse

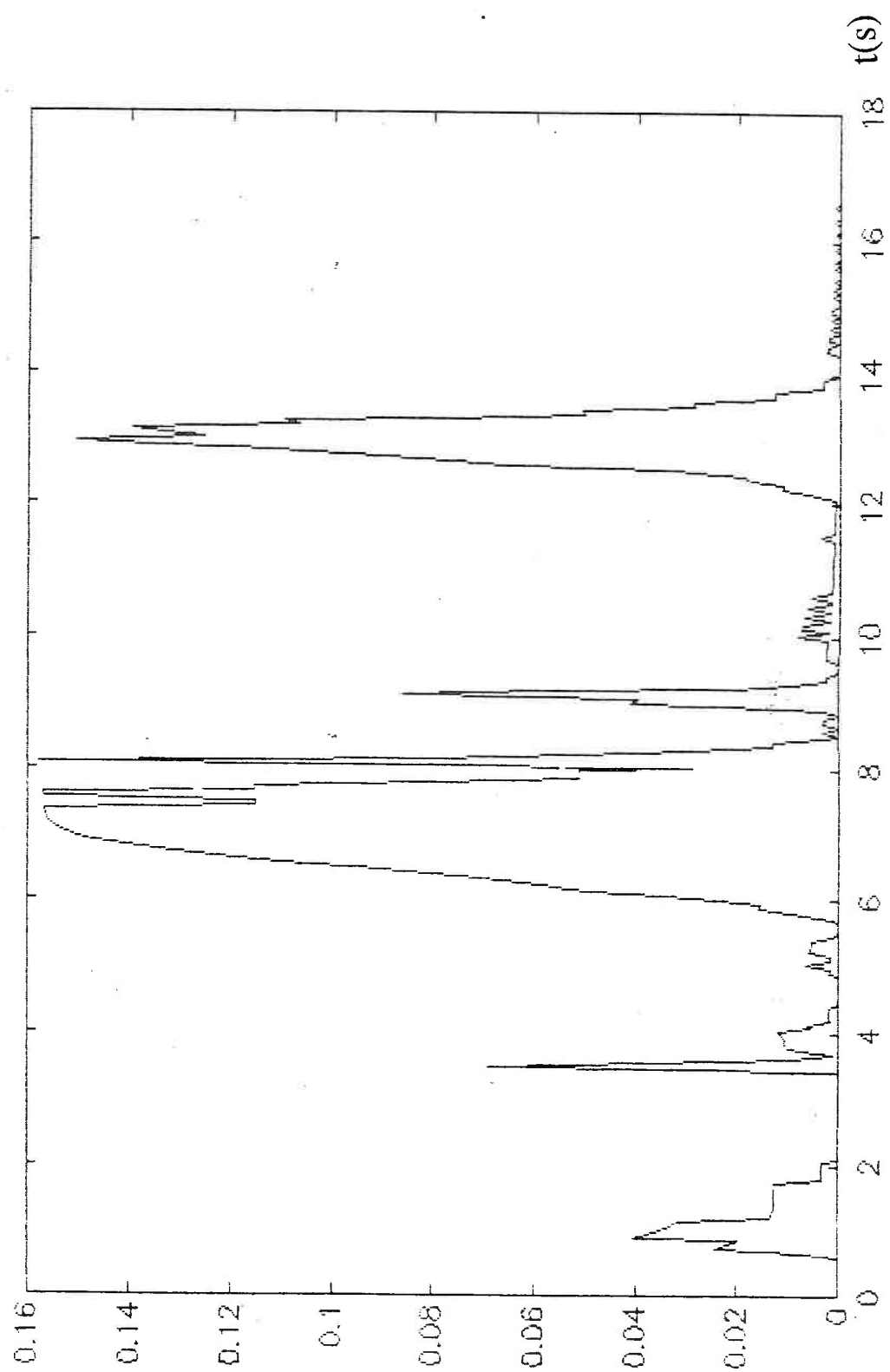


Figure 85 : D : distance de décision

Quelque soit la valeur de q_m l'occlusion ne pose pas de problème, notons que l'écart par rapport à la trajectoire réelle est dans ce cas moins marqué pour $q_m = 10$ que pour $q_m = 20$. En revanche, lors du changement de direction, le décrochage survient pour $q_m = 10$. Les valeurs des gains de Kalman (K_1 et K_2) montrent que l'impulsion donnée lors du basculement est plus importante lorsque q_m est grand, cela s'avère déterminant pour l'exemple choisi. La courbe des distance (fig. 85) montre qu'un seuil de 0,1 présenterait l'avantage de n'augmenter le valeur de q qu'en présence d'un changement de direction. Nous avons vu précédemment que le choix n'est pas applicable avec une fenêtre de recherche de petite taille, c'est pourquoi nous l'avons testé à partir d'une fenêtre de recherche plus grande (25 x 50 pixels) et les résultats obtenus montrent un meilleur comportement en présence d'occlusion (la valeur de q restant à q_0).

Cet accroissement de la taille de fenêtre permet aussi d'envisager des accélérations plus fortes, il est alors nécessaire d'accroître q_m , jusqu'à $q_m = 40$ le comportement du filtre est amélioré. Mais si la valeur de q_m est trop grande, lorsqu'une fausse alarme de manoeuvre est perçue (causée par une information parasite), la valeur mesurée à l'étape suivante est déterminante et peut diriger la recherche de la cible dans une direction erronée.

Enfin donnons un résumé de l'effet de chacun des paramètres sur la qualité de la poursuite, et les écueils envisageables;

- q_0 règle le niveau de filtrage du bruit apparaissant sur la trajectoire mesurée
 - + q_0 grand $\rightarrow$ filtre sans effet
 - + q_0 petit $\rightarrow$ système à constante de temps élevée mais ayant un bon comportement en l'absence de manoeuvre.
- q_m impose le taux d'adaptation en présence de manoeuvre
 - + q_m faible $\rightarrow$ le filtre est équivalent au modèle sans adaptation
 - + q_m grand $\rightarrow$ les valeurs récentes sont prépondérantes en présence de manoeuvre entraînant des risques d'instabilité.
- s détermine jusqu'où les écarts peuvent être considérés comme du bruit.
 - + s faible $\rightarrow$ un bruit sera pris pour une manoeuvre (basculement trop fréquent)
 - + s grand $\rightarrow$ Q sera toujours égal à q_0 .

Le filtrage de Kalman est adaptatif dans sa conception en ce sens qu'il est défini de manière récurrente en relation avec chaque nouvelle mesure ; le facteur de bruit variable ($q_0 - q_m$) introduit un deuxième niveau d'adaptation (sur le gain de Kalman K), c'est pourquoi il est impossible de le caractériser suivant les critères habituels de représentation des filtres.

Nous avons donc illustré les performances de notre système en observant son comportement face à différents types de mouvement et de perturbations et donnons des limites d'application en relation avec les conditions expérimentales dans lesquelles nous avons opéré. Les résultats obtenus débouchent sur une procédure de réglage des paramètres du filtre en fonction des types de trajectoires susceptibles d'être rencontrées dans le cadre d'un système de poursuite donné.

Par ailleurs, nous avons testé l'effet des variations de périodes d'échantillonnage, bien qu'il soit toujours souhaitable de les rendre minimum ; en l'absence de manoeuvre leur variation est presque sans effet. Mais lorsque le filtre bascule sur la valeur $q = q_m$, le problème est ramené au cas sans filtrage pour lequel une vitesse de décrochage est définie.

Une amélioration peut être envisagée pour les temps d'exécution, en exploitant l'indépendance des corrélations de chaque niveau (cf Chap. II) par l'utilisation d'une architecture informatique parallèle.

Enfin les techniques évoquées et les réalisations effectuées n'ont pas été envisagées pour l'analyse de mouvements en deux dimensions. Le dispositif expérimental n'a pas été conçu dans cette optique, et il ne permet pas d'effectuer des déplacements suivant deux directions différentes dans le plan perpendiculaire à la caméra. Le suivi d'objets animés d'un tel mouvement constitue cependant une extension proche de notre travail, une technique possible consisterait à exploiter la réactualisation de la fenêtre pour maintenir la poursuite. Une comparaison périodique du modèle actuel, au modèle existant quelques itérations plus tôt permettrait de détecter un déplacement dans la 2ème direction.

Exemple :

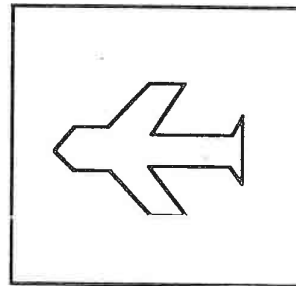
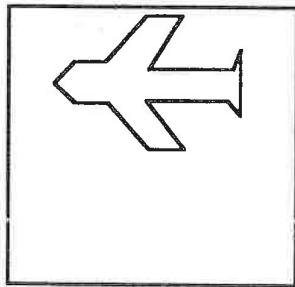


Figure 86 : modèle actuel

modèle précédent

Lorsqu'une différence notable entre les deux modèles est perçue, l'aire de recherche est décalée vers la direction de meilleure correspondance entre ces derniers.

Cette procédure n'est envisageable que dans le cas où l'axe de recherche est proche de l'axe du mouvement. Dans ce but, une possibilité de pivotement de la caméra serait nécessaire, lorsque le décalage se reproduirait plusieurs fois dans la même direction.

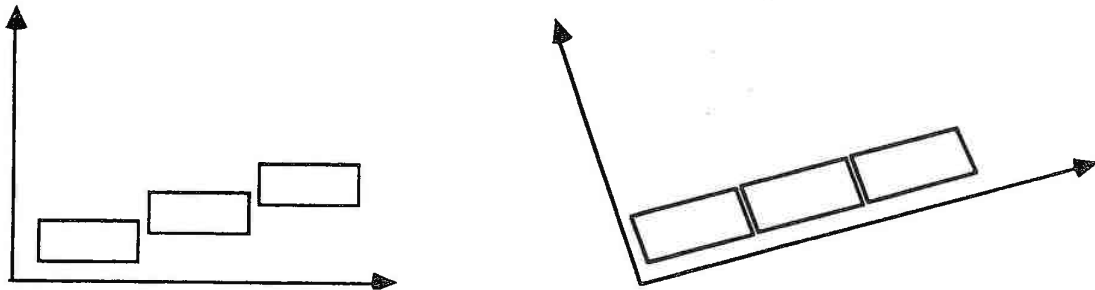


Figure 87 : Rotation du capteur

Ce principe serait envisageable dans le cadre d'un système comportant un asservissement de la caméra.

CONCLUSION

Le travail réalisé met en évidence la faisabilité d'un système de poursuite de cible par l'association extracteur de contour, corrélateur, et filtrage.

La réalisation et l'implémentation de l'ensemble des éléments constituant ce modèle a permis de déterminer les caractéristiques de chacun d'entre eux : modélisation des images issues du capteur, évaluation des temps de traitement requis par le corrélateur, estimation de la qualité des fonctions délivrées par ce dernier, et mesure des performances du filtrage de Kalman adapté.

Les différentes parties du système comportent des paramètres réglables pour chaque exécution (niveau de seuillage, modèle de bruit ...). Les essais effectués fournissent des critères de choix pour les valeurs qui leur seront affectés.

La représentation des images des contours a été testée pour l'application considérée, une prolongation de cette recherche pourrait être des investigations plus approfondies sur ce point.

Le filtre de Kalman s'avère approprié ; en effet il est faible consommateur de temps de traitement, et le mode d'adaptation choisi autorise son application dans différents contextes. Une amélioration intéressante serait de l'exploiter sur un dispositif comportant un asservissement de la position de la caméra afin d'élargir les domaines d'investigation pour le mobile.

Les extensions à apporter au banc d'essais en vue d'une exploitation industrielle concernent l'initialisation de la poursuite, la parallélisation de l'algorithme de corrélation, et éventuellement le câblage de certaines fonctions afin d'en diminuer le temps d'exécution.

BIBLIOGRAPHIE

- [AGG - 81] JK. AGGARWAL - LS. DAVIS - WN. MARTIN
Correspondance processes in dynamic scene analysis
Proceeding of the IEEE, vol 69 n° 5 - May 1981
- [BAR - 72] DI. BARNEA - HF. SILVERMAN
A class of algorithms for fast digital image registration
IEEE transactions on computers - Vol. C 21 n° 2 - Février 1972
- [BOL - 88] P. BOLON - BT. CHIA
Contribution à la comparaison d'opérateurs locaux de détection de contours
TIPI, Aussois, 1988
- [BON - 88] P. BONNIN - E. PAUCHON
Analyse de mouvement par imagerie infrarouge utilisant une segmentation coopérative points d'intérêt vs régions
PIXIM - 88 - 1ère Conférence annuelle sur l'image numérique
Paris - 24-28 Octobre 1988
- [BOU - 88] P. BOUTHEMY
Modèle et méthodes pour l'analyse du mouvement dans une séquence d'image
2ème Atelier Scientifique TIPI - Aussois - Savoie - 19-22 Avril 1988
- [BOZ - 83] C. BOZZO
Le filtrage optimal et ses applications aux problèmes de poursuite
Ed. Technique et documentation tome III
- [BRE - 86] P. BREUER - Z. FAZEKAS - I. RENYI - I. SZABO
Speeding up locating templates using parrallel matching
Congrès ICPR - Paris - 27-31 Octobre 1986
- [BUR - 82] PJ. BURT - C. YEN - X. XU
Local correlation measures for motion analysis a comparative study
IEEE Conf. on pattern recognition and image processing
Las Vegas - June 1982

- [CAR - 87] S. CARLSON - C. REILLO
Contour based motion analysis with application to prediction and segmentation
Proc. of the 5th Scandinavian Conference on Image Analysis IAPR
Stockholm 1987
- [CIC - 84] J. CICCOTELLI
Automatisation du contrôle magnétoscopique par système de vision hybride "Net-Flou"
Thèse 3ème Cycle, Génie Electrique, INPL 27 Juin 1984
- [CIC - 87] J. CICCOTELLI
Contribution du capteur d'image hybride "N-F" à la poursuite automatique de cible - Application à la robotique
Thèse d'état, INPL, 1987
- [COO - 79] DB. COOPER
Maximum likelihood estimation of Markov-Process blob boundaries in noisy images
I.E.E.E. Transaction on PAMI, Vol PAMI 1, n° 4, Octobre 1979
- [COU - 84] F. DE COULON
Théorie et traitement des signaux
Traité d'Electricité, Vol. 6, Ed. Georgi, Lausanne, 1984
- [COU - 85] RY. COULON - JP. CHARRAS
Trajectométrie automatique par caméra vidéo
5ème Congrès AFCET RdF et IA - Grenoble - Novembre 1985
- [DIE - 86] N. DIEHL - HB. BURKHARDT
Planar motion estimation with a fast converging algorithm
Congrès ICPR - Paris - 27-31 Octobre 1986
- [DUB - 84] X. DUBOIS - M. REBUFFET - M. GIRARD
PRIAM System de poursuite par reconnaissance et interprétation automatique du mouvement
Note ETCA - 1984
- [FAU - 88] OD. FAUGERAS
Vision artificielle en dimension 3
2ème atelier scientifique TIPI - Aussois - 19-22 Avril 1988

- [GAR - 76] G. GARRET - E. REAGH - E. HIBBS
Detection threshold estimation for digital area correlation
IEEE Trans. on Systems Man and Cybernetics - Janvier 1976
- [GOO - 72] JW. GOODMAN
Introduction à l'optique de Fourier et à l'holographie
Ed. Masson & Cie, 1972
- [GUI - 84] Y. LE GUILLOUX
Détermination automatique du mouvement dans une séquence d'image.
Intérêt pour l'interprétation.
Thèse de Docteur Ingénieur - 29 Juin 1984 - ENST
- [HAL - 79] EL. HALL
Computer Image Processing and Recognition
Academic Press, 1979
- [HAL - 80] EL. HALL - DI. DAVIES - ME. CASEY
The selection of critical subset for signal image and scene matching
IEEE Transaction on PAMI, Vol. PAMI 2, n° 4 - July 1980
- [HOR - 81] B. HORN - B. SCHUNK
Determining optical flow
Artificial Intelligence 17 - p 185-203 - 1981
- [HOR - 89 a] O. HORN - J. CICCOTELLI - R. HUSSON
Restrictive correlation evaluation of processing times in target tracking
SPIE/SPSE Symposium on Electronic Imaging
Advanced Devices and Systems, L.A., U.S.A, Janvier 1989
- [HOR - 89 b] O. HORN - D. WOLF - R. HUSSON
Filtrage de Kalman adaptatif appliqué à la poursuite de cible
Congrès AIPAC'89 : Traitement avancé de l'information pour l'automatisation en contrôle de procédé
Nancy, France, 3-5 Juillet 1989
- [JAI - 81] JR. JAIN - AK. JAIN
Displacement Measurement and its application in interframe image coding
IEEE Transaction on communications
Vol. com 29 n° 12 - Décembre 1981

- [JAI - 84 a] R. JAIN
Difference and accumulative difference pictures in dynamic scene analysis
Image and vision computing - Vol 2 n° 2 Mai 1984
- [JAI - 84 b] RC. JAIN
Segmentation of frame sequences obtained by a moving observer
IEEE Transaction on PAMI - Vol PAMI 6 n° 5 Septembre 1984
- [KAP - 85] S. KAPPA GANTULA - KR. RAO
Motion compensated interframe image prediction
IEEE Trans. on communication - Vol. COM 33 n° 9 - Sept. 1985
- [KOG - 83] T. KOGA et Al
A 1,5 Mb/s interframe codec with motion compensation
Proc. on Int. Conf. Communications - Boston - June 1983 - pp 71-75
- [LEG - 82] GR. LEGTERS - JR. and TY. YOUNG
A Mathematical Model for computer image tracking.
IEEE Transactions on pattern analysis and machine intelligence
Vol. PAMI - 4, n° 6, Nov. 1982
- [LEG - 82] GR. LEGTERS - TY. YOUNG
A mathematical model for computer image tracking
IEEE Transaction on PAMI
Vol. PAMI 4 n° 6 - Novembre 1982.
- [LET - 83] Ph. LETELLIER
Le Mouvement
INRIA "Vision par ordinateur" 13-17 Juin 1983
- [LEV - 79] -B. LEVINE
Fondements théoriques de la radiotechnique statistique
Edition Mir, Tome 3, Moscou 1979
- [LI - 85] X. LI - RC. DUBES
The first stage in two stage template matching
IEEE Transactions on PAMI - Vol. PAMI 7 - n° 6 - November 1985
- [LIF - 81] J. LIFERMANN
Les principes du traitement statistique du signal
Ed. Masson, Paris, 1981

- [MIL - 88] C. MILLOUR - A. LANUSSE
Intégration de mécanismes préattentifs en analyse par vision de scène dynamique
2ème atelier scientifique TIPI Aussois - Savoie - 19-22 Avril 1988
- [MOH - 88] R. MOHR
Sur l'appariement modèle perception
2ème Atelier TIPI Aussois - Savoie - 19-22 Avril 1988
- [MUN - 81] C. MUNTEANU
Evaluation of the sequential similarity detection algorithm applied to binary images
Pattern Recognition - Vol 13 - n° 2 1981
- [NAG - 86] H.H. NAGEL
Image Sequences - Ten (Octal) Years
From phenomenology towards a theoretical Foundation
Congrès ICPR - Paris - 27-31 Octobre 1986
- [NAR - 87] C. NARATHONG - RM. INIGO - JF. DONER - ES. Mc VEY
An algorithm for motion prediction using a biological visual sensor
Proc. IEEE International Conference on robotics automation
March April 1987 - North Carolina
- [OLI - 84] CJ. OLIVER
An analysis of template matching in image registration
Optica Acta - 1984 - Vol 31 n° 2
- [OLS - 87] TJ. OLSON - L. BUKYS - CM. BROWN
Low level image analysis on MIMD architecture
Proc. IEEE 1st international conf. on computer vision
June 8.11.1987 - London England
- [PAV - 86] T. PAVLIDIS
A critical survey of image analysis method
Congrès ICPR - Paris - 27-31 Octobre 1986
- [PER - 79] D. PERRIOT - MATHONA
Contribution à l'étude et à la mise en oeuvre d'algorithmes de filtrage de Kalman adaptatif
Thèse Docteur Ingénieur - Université de Paris Sud

- [PER - 80] D. PERRIOT - MATHONA
Le filtrage de Kalman adaptatif. Application à la poursuite de cible manoeuvrantes
Revue technique Thomson CSF - Vol. 12 n° 1
- [PIE - 86] F. PIERRE - Y. HERVE - R. EUGENE - C. DRUMAN - S. WENDLING
Description d'une machine de traitement d'images temps réel : l'approche multi-parallélisme
PIXIM - 88 1^{ère} conf. annuelle sur l'image numérique
Paris 24-28 Octobre 1988
- [PRA - 78] WK. PRATT
Digital Image Processing
Wiley & Sons, New York, 1978
- [ROD - 85] V. RODRIGUES - L. MASSIP-Pailhes - S. CASTAN
La détermination du vecteur déplacement et le modèle de la covariance temporelle
5^{ème} Congrès AFCET RdF - IA - Grenoble Novembre 1985
- [ROS - 82] A. ROSENFELD - A. KAK
Digital Picture Processing
Academic Press, 2nd Edition 1982
- [SAM - 84] R. SAMY
Traitement de séquence d'image télévision
Thèse de 3^{ème} cycle - 9 Mai 1984
- [SAM - 86] RA. SAMY - CA. BOZZO
Matching images using motion features
Congrès ICPR - Paris - 27-31 Octobre 1986
- [SAS - 87] JZ. SASIADEK - PJ. WOJUK
Tactile sensor signal processing using an adaptative Kalman Filter
Proc. IEEE International Conference on Robotics and Automation
North Carolina
- [SEC - 87] JP. SECILLA - N. GARCIA - J. CARRASCOSA
Locating templates in noisy environments
Proc. of the 5th Scandinavian conference on image analysis IAPR
Stockholm - June 1987

- [SIL - 87] J. SILANDER - O. SILVEN - I. VIRTUNEN
Experiments with an image pipelined processor
Proc. of the 5th Scandinavian Conference on Image Analysis
IAPR - Stockholm - 1987
- [SKL - 78] J. SKLANSKY
On the Hough Technique for curve detection
IEEE Transactions on computers - Vol C27 n° 10 - Octobre 1978
- [STE - 85] P. STELMASZYK - P. BONNET - JG. POSTAIRE
Analyse de scènes dynamiques par recherche des contours en mouvement
5ème Congrès AFCET RdF et IA Grenoble Novembre 1985
- [SUG - 86] SA. SUGIMOTO - H. MATSUKI - Y. ICHISKA
Implementation of tracking and extraction of moving objects in
successive frames
Applied Optics - Vol 25 n° 6 15 March 1986
- [TSA - 81] RY. TSAI - TS. HUANG
Estimating three dimensional motion parameters of a rigid planar patch
IEEE Transaction on acoustics speech and signal processing
Vol. ASSP 29, n° 6 - December 1981
- [TSA - 83] RY. TSAI - TS. HUANG
Uniqueness and Estimation of 3-D motion parameters and surface
structures of rigid objects
Research Report IBM
- [TZI - 88] G. TZIRITAS
Estimation récursive - itérative du mouvement apparent dans une
séquence d'image
2ème atelier scientifique TIPI Aussois - Savoie - 19-22 Avril 1988
- [VAN - 77] GJ. VANDERBRUG - A. ROSENFELD
Two stage template matching
IEEE Transactions on computers - Vol C 26 n° 4 - April 1977
- [WAL - 80] IP. WALLACE - OR. MITCHELL
Analysis of three dimensional movement using Fourier Descriptors
IEEE PAMI Vol 2 n° 6 Novembre 1980

- [WIK - 87] J. WIKLUND - GH. GIRANLUND
Image sequence analysis for object tracking
Proc. of the 5th Scandinavian Conference on image analysis IAPR
Stockholm 1987
- [WOL - 89] D. WOLF - J. CICCOTELLI - O. HORN - R. HUSSON
Assessment of motion parameters of a moving object through image
processing
SPIE/SPSE Symposium on Electronic Imaging
Advanced Devices and Systems, L.A., U.S.A, Janvier 1989
- [WON - 78] RY. WONG - EL. HALL
Sequential hierarchical scene matching
IEEE Transactions on computers - Vol. C-27 n° 4 - April 1978
- [XUE - 86] Y. XUESHAN - RR. MOHLER
Adaptative Kalman Filter for Maneuvering Target Tracking
Congrès Robotique et IA - Toulouse - 18/20 Juin 1986
- [ZAV - 88] B. ZAVIDOVIQUE - P. BONNIN - C. MILLOUR
Méthodes de détection de mouvement, application à la poursuite
Ecole d'automne Thomson - 6-10 Septembre 1988 - Article n° 13



AUTORISATION DE SOUTENANCE DE THESE
DU DOCTORAT DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

VU LES RAPPORTS ETABLIS PAR :

Monsieur WENDEL, Professeur, ENSPS Strasbourg,
Monsieur HATON, Professeur, INRIA-CRIN/Université de Nancy I.

Le Président de l'Institut National Polytechnique de Lorraine,
autorise :

M ademoiselle HORN Odile

à soutenir devant l'INSTITUT NATIONAL POLYTECHNIQUE DE
LORRAINE, une thèse intitulée :

"Etude et mise en oeuvre d'un algorithme de poursuite de cible par
analyse d'image"

en vue de l'obtention du titre de :

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Spécialité : "Génie Electrique"

Fait à Vandoeuvre le, 15 Juin 1989

Le Président de l'I.N.P.L.,

M. GANTOIS

2, avenue de la Forêt de Haye - B.P. 3 - 54501 VANDOEUVRE CEDEX

Téléphone : 83. 57. 48. 48 - Télex : 961 715 F - Télécopie : 83. 57. 49. 55

**Etude et mise en oeuvre
d'un algorithme de poursuite de cible
par analyse d'image**

Mots clés : analyse de scène dynamique, corrélation rapide, extraction de contours, filtre de Kalman, poursuite de cible, estimation de trajectoire.

Résumé : Les travaux décrits dans ce mémoire concernent l'analyse de scène dynamique appliquée à la poursuite de cible. Une présentation des techniques employées pour l'étude des mouvements par traitement d'image est faite dans un premier temps.

L'approche développée se compose de trois parties :

- un capteur optoélectronique délivre les contours de la scène en temps réel,
- une technique originale de corrélation rapide "par restriction" détermine la position du mobile pour chaque acquisition d'image,
- un filtre de Kalman adaptatif calcule la trajectoire de la cible et prédit sa position pour l'étape suivante.

Ces différents éléments sont détaillés et les critères de choix pour leur réalisation sont exposés. D'un point de vue quantitatif, une modélisation des images des contours est établie et testée, elle permet l'évaluation du temps d'exécution requis pour la localisation de la cible. L'ensemble des modules est mis en oeuvre, et des essais sont réalisés à partir d'un mobile de trajectoire connue et commandée. Les résultats issus du système de poursuite peuvent ainsi être comparés aux paramètres du mouvement réel.

**A target tracking algorithm by picture processing :
design and realization**

Keywords : Dynamic scene analysis, fast correlation, edge extraction, Kalman filter, target tracking, estimation of trajectory.

Abstract : The researches explained in this thesis deal with the dynamic scene analysis for target tracking. At first we make a survey of the techniques used for motion estimation by picture processing.

The developed approach is composed of three parts :

- an optoelectronic sensor generates edge picture in real time,
- an original technique of fast correlation "by restriction" determines the position of the moving object for each picture acquisition,
- an adaptive Kalman filter computes the trajectory of the target and predicts its next position.

We detail these different elements, and expose the choice criterium for their realization. In a quantitative point of view, we establish and test a modelization of edge pictures. It allows the assessment of the processing time required for target localization. All the parts have been achieved, and tests have been realized on a moving object with a known and constrained trajectory. The results from the tracking system are compared with real parameters of the movement for different kinds of displacement.