

**PERFORMANCES COMPAREES DE METHODES DE
POURSUITE DE CIBLES MOBILES EN ECHOGRAPHIE
ULTRASONORE**

DEVEVEY C. * - DENIS F. **
CACHARD C. * - GIMENEZ G. *

* LTSU, URA 1216, INSA bat. 502, 69621 Villeurbanne CEDEX
** LEMI, UCB Lyon I bat. 710, 69622 Villeurbanne CEDEX

Résumé: Quatre méthodes classiques de poursuite sont comparées en fonction de leur temps de calcul et de leur précision. La validation est effectuée sur des images échographiques de bulles en ascension libre et les résultats sont confrontés à des mesures par vélocimétrie Doppler.

Abstract: The accuracy and the processing time of four classical tracking methods are compared. The validation is made on echographic images of free rising bubbles and the results are collated to Doppler velocimetry measures.

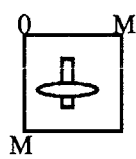
1 - INTRODUCTION

La poursuite d'objets mobiles dans une séquence d'images (tracking) a donné lieu à de nombreux travaux au cours des dernières années. La plupart des techniques ont été développées en vue d'applications précises que ce soit la robotique [SAG-89], l'imagerie satellite [AGG-81] ou le codage vidéo [BIE-86].

Dans cette étude, les comportements de quatre méthodes "classiques" sont testés sur des images échographiques ultrasonores en présence d'un bruit multiplicatif important (speckle). Les principaux critères de comparaison sont la précision des résultats et, en particulier l'erreur cumulée après plusieurs itérations, ainsi que le temps de calcul en vue d'une application temps réel.

2 - METHODES DE POURSUITE

Chacun des algorithmes présentés nécessite le pointage initial de l'objet à poursuivre pour extraire une fenêtre de référence contenant la cible. Le travail consiste alors à trouver dans les images suivantes des fenêtres de même taille et de mêmes caractéristiques correspondant à la cible après son déplacement.



fenêtre de référence $M \times M$

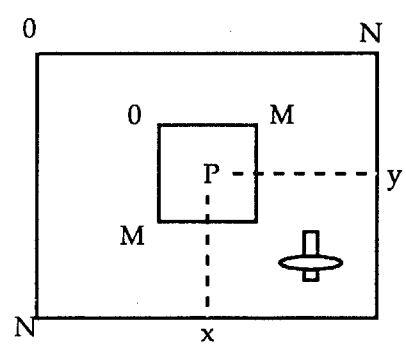


image $N \times N$ et fenêtre image $P(x, y)$ de taille $M \times M$

Fig. 1: définition des notations

Les notations utilisées sont les suivantes (Fig. 1):
 - R est la fenêtre de référence de $M \times M$ points contenant l'objet à poursuivre et obtenue par pointage initial;
 - J représente l'image de $N \times N$ points ($M < N$) contenant une fenêtre image $P(x, y)$ centrée en (x, y) et de taille $M \times M$;



a - Corrélation

Le coefficient de corrélation est une mesure de ressemblance entre deux signaux. Sa définition en deux dimensions permet donc de comparer les deux fenêtres référence et image. L'étude de Burt [BUR-82] a montré qu'en présence d'un bruit important, il faut normer le coefficient de corrélation pour avoir des résultats corrects. Avec les notations précédentes, l'intercorrélation $C(x, y)$ entre la fenêtre de référence R et une fenêtre P de l'image J de taille $M \times M$ centrée en (x, y) est définie par la formule suivante (1):

$$C(x, y) = \frac{\sum_{i=1}^M \sum_{j=1}^M R(i, j) \cdot P(i+x, j+y)}{\left\{ \sum_{i=1}^M \sum_{j=1}^M R^2(i, j) \right\}^{1/2} \left\{ \sum_{i=1}^M \sum_{j=1}^M P^2(i+x, j+y) \right\}^{1/2}}$$

avec: $-1 \leq C(x, y) \leq 1$ (valeur normée)

Ce coefficient est calculé pour toutes les fenêtres $P(x, y)$ centrées sur un des points de la zone de recherche. La cible est identifiée au point de meilleure correspondance, c'est à dire au point (a, b) où l'intercorrélation est maximale:

$$\forall (x, y) \neq (a, b) \quad C(x, y) < C(a, b)$$

b - Statistiques de la fenêtre

Si la fenêtre de référence et la fenêtre image sont identiques, les histogrammes de leurs niveaux de gris sont eux aussi identiques. Cette méthode cherche à trouver dans la zone de recherche, l'histogramme de la fenêtre image P qui ressemble le plus à celui de la fenêtre de référence R . Cette mise en correspondance s'effectue par un calcul d'erreur entre les statistiques des deux fenêtres (moyenne, variance, etc ..). La correspondance est réalisée au point (a, b) où cette erreur est minimale.

c - Displaced Frame Difference (DFD)

La DFD mesure la distance entre la fenêtre image et la fenêtre de référence en effectuant la somme des valeurs absolues de leurs différences pixel à pixel. Elle est utilisée pour estimer des mouvements de translation en codage vidéo [THO-89].

Avec les notations de la figure 1, la DFD entre les fenêtres R et P pour un déplacement $D = (x, y)$ entre les dates t et $t-T$ est défini par:

$$DFD(X, T) = \sum_{i=1}^M \sum_{j=1}^M |R_{k-1}(i, j) \cdot P_k(i+x, j+y)| \quad (2)$$

Le déplacement estimé est celui qui minimise la DFD.

d - Estimateur de déplacement

Une approche utilisée par Bierling [BIE-86] puis par Grotz [GRO-89] pour le codage de séquences d'images vidéo consiste à minimiser itérativement la différence entre le déplacement estimé à la $i^{\text{ème}}$ itération D_i et le déplacement réel D . Une approximation au premier ordre du développement en série de cette différence $D - D_i$ permet d'évaluer le déplacement. Le calcul est effectué à partir des espérances mathématiques (E) de diverses valeurs calculées sur des fenêtres de même taille extraites des images correspondant aux dates t (indice k) et $t-T$ (indice $k-1$). Soient:

- FD (Frame Difference), la différence intertrame des luminances I au même point (x, y) des trames k et $k-1$:

$$FD(x, y) = I_k(x, y) - I_{k-1}(x, y);$$

- G_x (respectivement G_y), la moyenne des gradients directionnels suivant x (respectivement suivant y) au point (x, y) sur les images k et $k-1$:

$$G_x(x, y) = \frac{1}{2} \cdot \left[\frac{\partial}{\partial x} I_k(x, y) + \frac{\partial}{\partial x} I_{k-1}(x, y) \right]$$

$$G_y(x, y) = \frac{1}{2} \cdot \left[\frac{\partial}{\partial y} I_k(x, y) + \frac{\partial}{\partial y} I_{k-1}(x, y) \right]$$

- et DEN, la quantité:

$$DEN = E(G_x^2) \cdot E(G_y^2) - E^2(G_x \cdot G_y).$$

La différence $d = D - D_i = (dx, dy)$ est alors calculée de manière itérative jusqu'à l'obtention de $dx = dy = 0$ avec rafraîchissement des fenêtres de calcul à chaque itération par les formules suivantes :

$$dx = \frac{E(G_x \cdot G_y) \cdot E(FD \cdot G_y) - E(FD \cdot G_x) \cdot E(G_y^2)}{DEN} \quad (3)$$

$$dy = \frac{E(G_x \cdot G_y) \cdot E(FD \cdot G_x) - E(FD \cdot G_y) \cdot E(G_x^2)}{DEN} \quad (4)$$

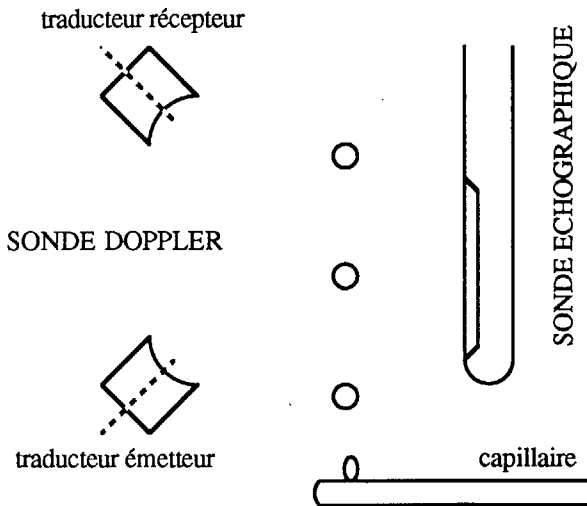


Fig 2: Dispositif expérimental d'acquisition des signaux.

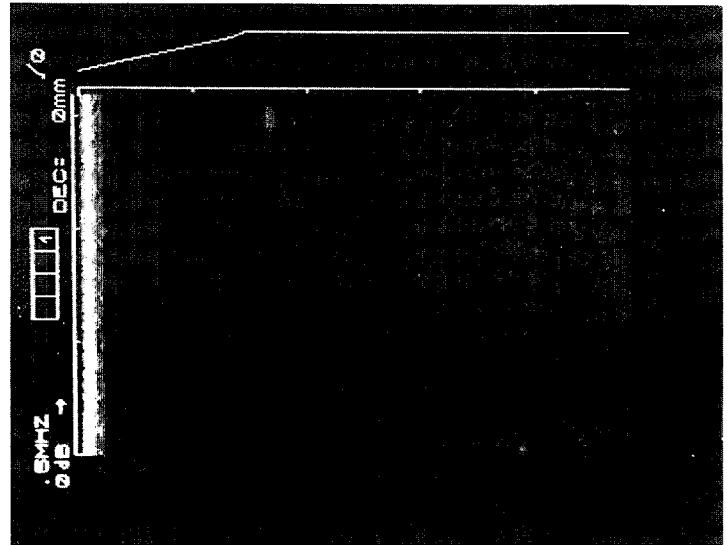


Fig 3: image de bulles

3 - RESULTATS

La validation expérimentale a été effectuée sur des images échographiques de bulles d'azote en ascension libre dans de l'eau (Fig. 2). Cette configuration qui présente l'avantage de travailler sur des cibles parfaitement maîtrisées, est très proche de la situation rencontrée en échographie médicale. Les images utilisées présentent les défauts habituels des images d'échographie médicale, à savoir un faible contraste et un niveau de bruit important (Fig. 3). De plus les célérités ainsi calculées peuvent être dans ce cas confrontées à des mesures de célérité réalisées par vélocimétrie Doppler ultrasonore [CAC-86, GIM-90] afin de contrôler la précision des résultats (Fig. 4).

Les mesures obtenues à l'aide des statistiques des niveaux de gris de la fenêtre sont entachées d'erreurs. Ce n'est pas surprenant compte tenu de la sensibilité des statistiques au bruit et de l'importance du speckle sur les images échographiques. De plus, en prenant l'histogramme d'une fenêtre, l'information contenue dans la localisation des niveaux de gris est perdue, ce qui tend à augmenter la probabilité de fausse correspondance.

La corrélation et la DFD donnent des résultats satisfaisants et presque identiques, ce qui s'explique facilement car les deux méthodes minimisent une distance:

- la maximisation du coefficient de corrélation revient à minimiser l'écart quadratique moyen:

$$\epsilon_1 = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M (X_{ij} - b \cdot R_{ij})^2$$

- pour la minimisation de la DFD, c'est l'erreur absolue que l'on prend en compte:

$$\epsilon_2 = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M |X_{ij} - R_{ij}|$$

Il faut cependant noter que le temps de calcul de la DFD (une somme de différences) est plus faible que celui de l'intercorrélation (une somme de produits et une normalisation).

Pour ces deux méthodes, si le déplacement n'est pas de trop forte amplitude, la taille de la fenêtre optimale s'étend sur une gamme importante (plus ou moins cinq pixels dans les deux directions). La corrélation possède un léger avantage de robustesse sur la DFD.

L'estimateur de déplacement manque de constance dans ses résultats, il est en particulier très sensible à la manière dont sont calculés les gradients qui interviennent dans sa formulation. Cet inconvénient est pallié par l'utilisation d'un masque de Sobel moins sensible au bruit. Plus grave, pour que l'estimateur fonctionne, la fenêtre d'extraction de paramètres sur chacune des images doit contenir entièrement la cible avant et après déplacement. Si le mouvement est de forte amplitude, la fenêtre doit donc être de très grande taille, mais la cible devient petite vis à vis du reste de la fenêtre et cesse d'être significative dans les calculs: le déplacement estimé est nul. Une fréquence d'échantillonnage élevée minimise le déplacement entre les deux trames, mais l'estimateur donne de toute manière de très mauvais résultats, si la vitesse des cibles est élevée.



4 - CONCLUSIONS

La DFD donne des résultats satisfaisants pour un temps de calcul deux fois inférieur à celui de la corrélation (Fig. 5). Plus rapide encore, l'estimateur de déplacement ne peut être utilisé que si la taille de la cible est du même ordre de grandeur que l'amplitude du déplacement.

Le principal avantage de ces méthodes de traitement d'image est la poursuite d'objets se déplaçant à vitesse variable (la mesure de vitesse est dans ce cas délicate à réaliser par vélocimétrie Doppler) et leur principal inconvénient est la nécessité d'un pointage initial de la cible.

	seq. 1	seq. 2	seq. 3
Doppler cm/s	2,5	3,2	3,1
DFD cm/s	2,5	3,2	2,8
corrélation cm/s	2,5	3,1	2,9
statistiques cm/s	1,9	4,3	3,7

Fig. 4: Vitesse estimée par les quatre méthodes pour trois séquences d'images.

	estim.	DFD	corrél.	stat.
T	0,1 s	2,4 s	5,5 s	3,8 s

Fig. 5: temps de calcul pour des fenêtres d'extraction de paramètres 30x30 (M) et un déplacement maximum de dix pixels sur un PC 386 muni d'un coprocesseur.

Références:

[AGG-81] AGGARWAL J.K., DAVIS L.S., MARTIN W.N, Correspondence processes in dynamic scene analysis, *Proc. of the IEEE*, 1981, Vol. 69, No 5, p. 562-572.

[BIE-86] BIERLING M., THOMA R., Motion compensating field interpolation using a hierarchically structured displacement estimator, *Signal processing*, 1986, Vol. 11, p. 387-404.

[BUR-82] BURT P.J., YEN C., XU X., Local correlation measures for motion analysis, a comparative study, *Proc. Conf. on Pattern Recognition and Image processing*, Las Vegas, 1982, p. 269-274.

[CAC-86] CACHARD C., GIMENEZ G., Mesure ultrasonore de bulles en mouvement dans un liquide, *Revue d'Acoustique*, 1986, Vol. 78, p.19-23.

[GIM-90] GIMENEZ G., CACHARD C., VRAY D., Use of an analytic signal to model interaction between an acoustic wave and a moving target with a time- dependent velocity, *IEEE trans. on Ultrasonics, Ferroelectrics and Frequency Control*, 1990, Vol. 37, No 3, p. 196-204.

[GRO-89] GROTZ K., MAYER J.U., SUESMEIER G.K., A 64 kbits/s videophone codec with forward analysis and control, *Signal Processing: Image Communication*, 1989, Vol. 1, No 2, p. 103-115.

[SAG-89] SAGE D., *Poursuite d'objets par analyse d'une séquence d'images couleurs*, Thèse de doctorat, Automatique et Traitement du Signal, INPG, Avril 1989

[THO-89] THOMA R., BIERLING M., Motion compensating interpolation considering covered and uncovered background, *Signal Processing: Image Communication*, 1989, Vol. 1, No 2, p. 191-212.