

ESTIMATION DU MOUVEMENT PAR FILTRAGES SPATIO-TEMPORELS RECURSIFS

D. PELLERIN, A. SPINEI, A. GUERIN-DUGUE

Laboratoire de Traitement d'Images et Reconnaissance de formes - INPG
46 Avenue Félix Viallet, 38031 Grenoble Cedex, France
E-mail : name@tirf.inpg.fr

RÉSUMÉ

Nous présentons une méthode fréquentielle de calcul du flot optique à base de filtres spatio-temporels de Gabor, inspirée de l'architecture introduite par Heeger en 1987. Cette méthode, dite de filtrages combinés, associe des filtres de sélectivités différentes. Elle permet ainsi d'estimer un flot optique de qualité en s'appuyant sur la détermination préalable des masques des objets en mouvement. Nous proposons une implantation récursive très efficace des filtres de Gabor 3D permettant de réduire le volume de calculs.

1. INTRODUCTION

Nous présentons une méthode de calcul du flot optique dans les séquences d'images à base de filtres spatio-temporels orientés, inspirée de l'architecture introduite par Heeger [1]. Cette méthode repose sur le couplage d'une étape de détection des zones en mouvement (masques) suivie d'une étape d'estimation de vitesse. Pour réduire le volume des calculs, nous proposons une implantation récursive très efficace des filtres.

2. APPROCHE PAR BANC DE FILTRES DE GABOR 3D

Le spectre associé à une image $i(x,y)$ en translation à vitesse uniforme $v(v_x, v_y)$ occupe, dans le domaine des fréquences spatio-temporelles, un plan passant par l'origine [1] d'équation :

$$f_t = v_x f_x + v_y f_y$$

où f_x, f_y représentent les fréquences spatiales et f_t la fréquence temporelle. Le principe des méthodes fréquentielles consiste à localiser ce plan d'énergie non nulle ou plan de vitesse au moyen de filtres spatio-temporels orientés, pour obtenir ensuite directement les composantes du vecteur vitesse (v_x, v_y) .

Les filtres spatio-temporels ne sont pas sélectifs à une vitesse particulière mais plutôt réglés sur des fréquences spatio-temporelles : ils ne permettent donc pas à eux seul l'estimation de vitesse. Heeger a proposé une méthode de calcul du flot optique combinant les réponses fournies par un banc de filtres spatio-temporels de Gabor implantés par somme de convolutions séparables 1D [1].

Un filtre de Gabor 3D à phase cosinus est représenté

ABSTRACT

We present a frequential method to calculate optical flow based on spatio-temporal Gabor filters, inspired from the architecture of Heeger introduced in 1987. This method, named combined filtering, associates filters of different selectivities. It produces a very good optical flow based on the previous determination of the masks of the moving objects. We propose a very efficient recursive implementation of the 3D Gabor filters, which reduces the number of calculus.

comme le produit d'une fonction spatio-temporelle Gaussienne et d'une fonction cosinus :

$$g(x, y, t) = \frac{1}{(\sqrt{2\pi})^3 \sigma_x \sigma_y \sigma_t} \exp\left(-\left\{\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} + \frac{t^2}{2\sigma_t^2}\right\}\right) \\ \times \cos(2\pi(f_{x0}x + f_{y0}y + f_{t0}t))$$

où (f_{x0}, f_{y0}, f_{t0}) est la fréquence centrale du filtre et $(\sigma_x, \sigma_y, \sigma_t)$ l'écart-type associé à la fonction spatio-temporelle Gaussienne 3D. La somme des carrés des réponses de deux filtres de Gabor en quadrature (l'un à phase cosinus, l'autre à phase sinus mais d'orientation et de largeur de bande identique) donne une mesure d'énergie indépendante de la phase du signal. Le filtre obtenu est appelé filtre d'énergie de Gabor.

L'architecture de filtrage proposée par Heeger consiste à placer un ensemble de 36 filtres orientés d'énergie de Gabor répartis en trois gammes de vitesse. Pour chaque gamme de vitesse, 12 filtres sont positionnés sur un cylindre à différentes fréquences spatiales et temporelles (Figure 1). L'estimation de vitesse est formulée comme une minimisation de l'écart quadratique entre les réponses mesurées des filtres pour une image quelconque et les réponses prédites pour une image de bruit blanc en translation.

La méthode de Heeger fournit une très bonne estimation des vitesses mais elle présente l'inconvénient majeur d'utiliser des filtres dont le noyau de convolution spatial est de taille importante (23 coefficients). Cela conduit à un effet de lissage spatial excessif et donc une mauvaise détection des zones en mouvement, ainsi qu'à un volume de calculs très important.

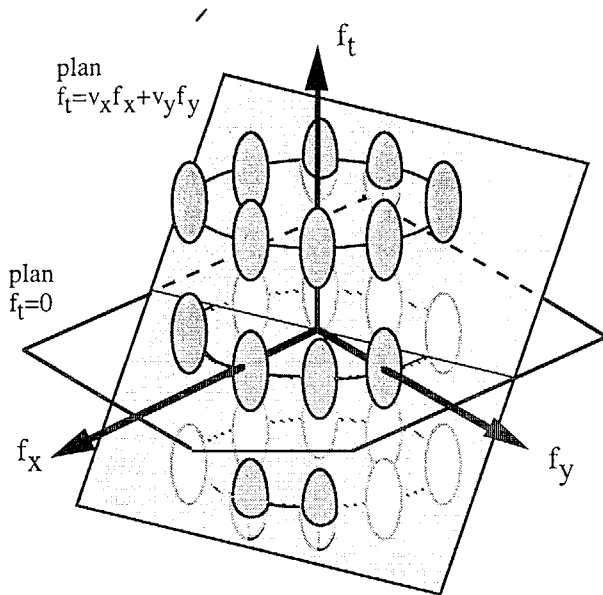


Figure 1 : Spectres de 12 filtres spatio-temporels orientés de Gabor pour une seule gamme de vitesse. Chaque filtre a la forme d'une paire d'ellipsoïdes (Gaussienne 3D) centrées à (f_x0, f_y0, f_t0) et $(-f_x0, -f_y0, -f_t0)$.

3. METHODE DE FILTRAGES SPATIO-TEMPORELS COMBINES

Nous proposons une méthode, dite de filtrages combinés, permettant à la fois la détection des zones de mouvement et une estimation de vitesse de qualité. Cette méthode consiste à faire coopérer, à l'instar du système visuel des vertébrés, deux mécanismes distincts pour la détection des zones de mouvement et l'estimation de vitesse. Nous avons choisi d'agir seulement sur la sélectivité des filtres spatiaux, c'est à dire les écart-types (σ_x, σ_y) , pour exécuter chacune de ces deux tâches, et de conserver globalement l'architecture décrite par Heeger.

Heeger a choisi pour les filtres de Gabor 3D les valeurs d'écart-types $\sigma_t=1$ (en temporel) et $\sigma_s=\sigma_x=\sigma_y=4$ (en spatial) sans réelle justification. Nous avons étudié l'effet du choix de l'écart-type σ_s grâce à des séquences d'images test. La séquence d'images test de la figure 2 se compose d'un carré en translation à droite et vers le bas sur un fond immobile. Le carré et le fond sont constitués de textures de type bruit blanc (spectre quasi uniforme).

Nous définissons deux critères d'erreur :

- l'erreur sur le flot optique EF s'exprime comme l'écart quadratique moyen (normalisé) entre les vecteurs vitesses mesurées et vitesses prédites en chaque pixel de l'image I.

$$EF(\%) = \frac{\sum_{i \in I} |v_{mi} - v_{pi}|^2}{\sum_{i \in I} |v_{mi}|^2} \times 100$$

- l'erreur sur le masque EM s'exprime comme l'écart (normalisé) entre le nombre total de pixels détectés en mouvement dans l'image (D) et le nombre exact de pixels du carré en mouvement (C).

$$EM(\%) = \frac{\sum_{i \in D} p_i - \sum_{j \in C} p_j}{\sum_{j \in C} p_j} \times 100$$

Ces critères d'erreur nous ont permis de sélectionner deux valeurs d'écart-type intéressantes ($\sigma_s=4$ pour EF minimum et $\sigma_s=2.4$ pour EM minimum). Les résultats obtenus sur la séquence test sont présentés dans la figure 2 et le tableau 1.

Nous observons que :

- pour des filtres spatiaux d'écart-type importants ($\sigma_s=4$), le flot optique estimé est de bonne qualité. Le phénomène de lissage spatial conduit malheureusement à un flot optique qui déborde largement les objets en mouvement.

- à l'inverse, pour les filtres spatiaux d'écart-type plus faible ($\sigma_s=2.4$), le flot obtenu comporte de nombreux vecteurs vitesses pour lesquels le module est sous-estimé. Ces filtres permettent toutefois une bonne détection des zones en mouvement.

Notre méthode de filtrage combiné consiste à effectuer successivement deux estimations de vitesses :

- la première, réalisée avec des filtres peu sélectifs, permet par seuillage et dans le cas d'une caméra fixe, la détermination des masques des objets en mouvement

- la seconde, réalisée avec des filtres très sélectifs, permet d'estimer correctement le flot optique mais seulement sur les objets en mouvement. Le flot optique final est alors constitué du second flot calculé affiché seulement sur les masques des objets en mouvement.

σ_s	EF (%)	EM (%)
4	11.35	47.08
2.4	13.65	27.03

Tableau 1 : Erreurs sur le flot optique (EF) et sur les masques (EM) en fonction de l'écart-type.

Nous présentons les résultats obtenus pour la séquence d'images réelle intitulée "Taxi de Hambourg" (Fig. 3). Dans cette scène de rue, filmée avec une caméra fixe, un piéton et trois véhicules sont en mouvement avec des vitesses comprises entre 0.3 et 3 pixels par image.

4. IMPLANTATION RECURSIVE DES FILTRES D'ENERGIE

L'implantation directe (par convolution) des filtres de Gabor 3D (ou plus précisément des filtres d'énergie) constitue une étape extrêmement coûteuse en calculs. Nous proposons une implantation récursive beaucoup plus efficace des filtres d'énergie, inspirée des travaux de Unser [2]. La sortie $y(k)$ d'un filtre de Gabor 1D (exprimé sous forme complexe) excité par un signal d'entrée $x(k)$ s'écrit :

$$y(k) = x(k) * \left[K \exp\left(-\frac{k^2}{2\sigma^2}\right) \exp(2\pi j f_0 k) \right]$$

d'où en réorganisant l'expression,

$$y(k) = \left[x(m) \exp(-2\pi j f_0 m) \right] * K \exp\left(-\frac{k^2}{2\sigma^2}\right) \exp(2\pi j f_0 k)$$

Ainsi, un filtrage de Gabor peut être décomposé en trois étapes successives : une modulation, un filtrage passe-bas (enveloppe gaussienne) et une démodulation (Figure 3).

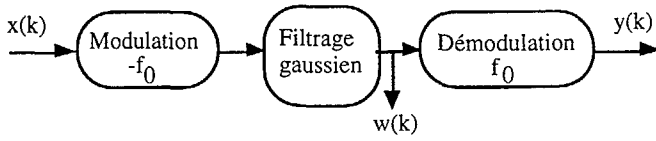


Figure 3 : Décomposition du filtrage de Gabor

Pour les filtres d'énergie, l'étape de démodulation pour la recomposition du signal est inutile. En effet, en appelant $w(k)$ la sortie intermédiaire dans la décomposition du filtrage, on montre sans difficulté que :

$$|w(k)|^2 = |y(k)|^2$$

Le filtrage passe-bas gaussien est réalisé au moyen d'un filtre récursif, qui offre l'avantage d'avoir un nombre de coefficients indépendant de la sélectivité. Nous avons choisi des filtres d'ordre 3 ayant une erreur quadratique moyenne avec un filtrage gaussien très faible (EQM=0.15% pour $\sigma_s=4$). Les coefficients de ces filtres ont été calculés à partir des expressions fournies par Young [3]. Ces filtres, dont les coefficients sont des fonctions cubiques de la sélectivité, se décomposent en une partie causale et une partie anticausale :

$$H(z) = \frac{A_2}{b_0 - b_1z^{-1} - b_2z^{-2} - b_3z^{-3}} \times \frac{A_2}{b_0 - b_1z^1 - b_2z^2 - b_3z^3}$$

L'implantation récursive des filtres (spatiaux) permet une accélération des calculs pour le filtrage d'un facteur 5 environ par rapport à une implantation directe (tableau 2).

Implantations	filtrage (ms/pixel)	flot optique (s/flot)
Directe	0.82	27.5
Récursive	0.16	16.8
Récursive (filtres combinés)	0.32	23.4

Tableau 2 : Temps d'exécution pour la phase de filtrage d'une part et l'ensemble du calcul d'estimation d'un flot optique sur des images 128*128 pixels avec une station de travail Sparc 10 d'autre part.

5. CONCLUSION

La méthode de filtrages spatio-temporels combinés permet, à partir de la méthode de Heeger, d'améliorer la qualité du flot optique pour des séquences d'images avec caméra fixe. Nous avons décrit une implantation récursive des filtres d'énergie qui permet une réduction importante de la durée du filtrage. Cette implantation ouvre deux voies : celle d'une architecture nouvelle comportant plus de filtres pour l'élargissement des gammes de vitesses, et celle d'une implantation analogique des filtres (circuits VLSI).

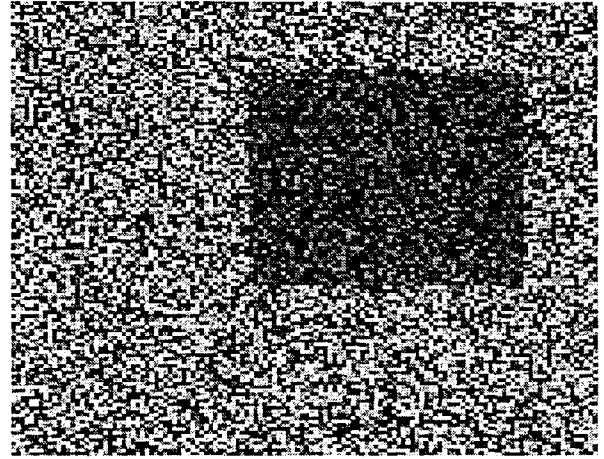
Bibliographie :

[1] D.J. Heeger, "Model for the extraction of image flow",

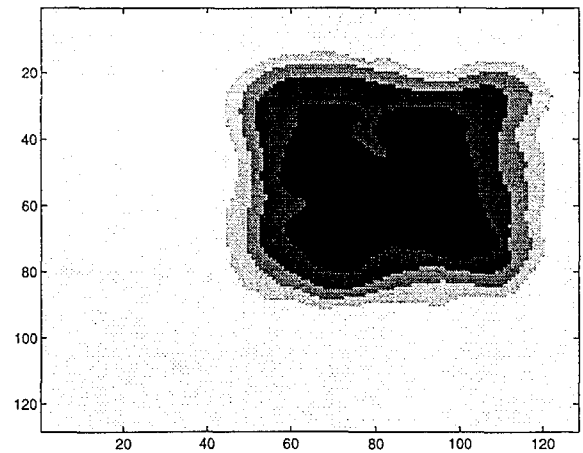
Journal of the Optical Society of America A., vol. 4, n° 8, 1455-1471, 1987.

[2] M. Unser, "Fast Gabor-like windowed Fourier and continuous wavelet transforms", NCCR Report 143/93, National Institutes of Health, 1993.

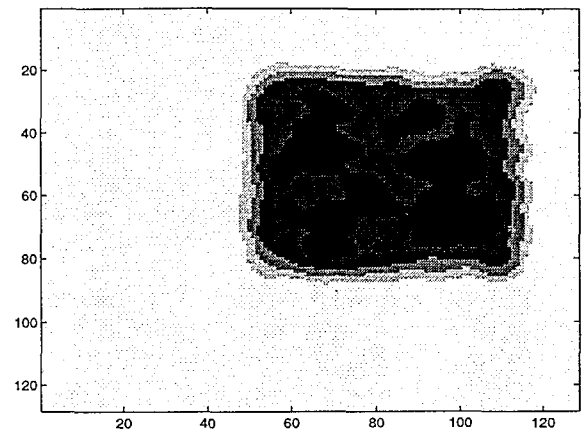
[3] I.A. Young and L.J. van Vliet, "Recursive implementation of the Gaussian Filter", Soumis à Signal Processing, 1994.



a) Séquence test



b) $\sigma_s=4$

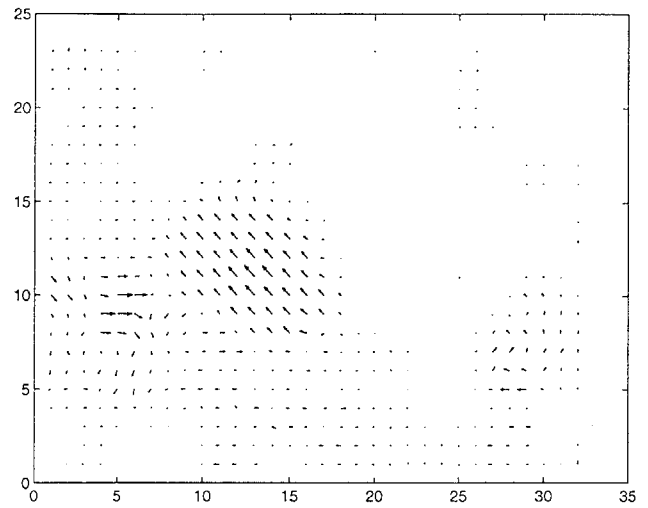
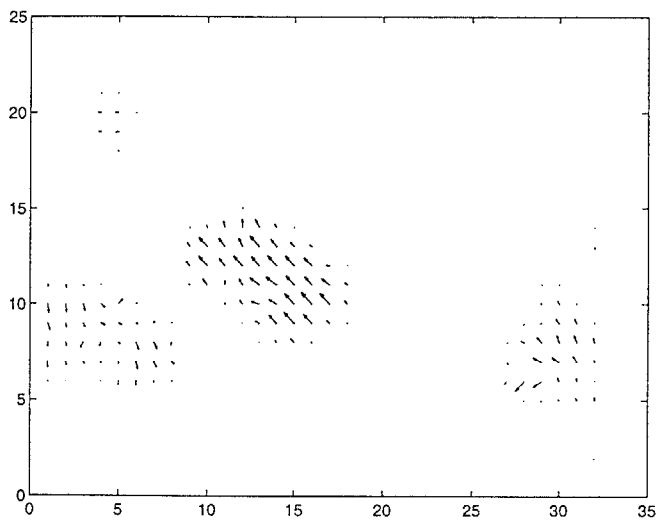
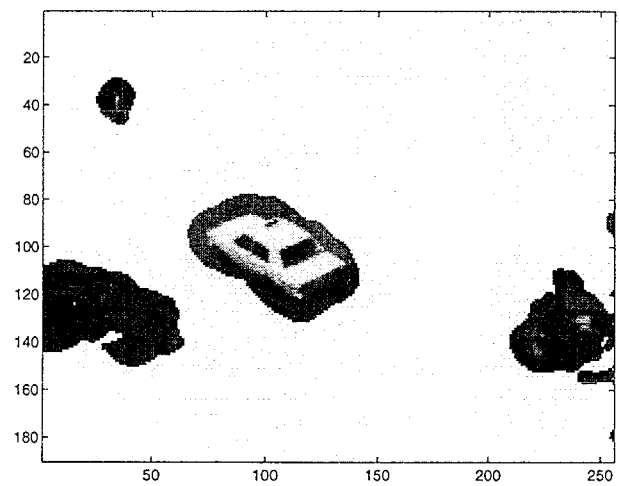


c) $\sigma_s=2.4$

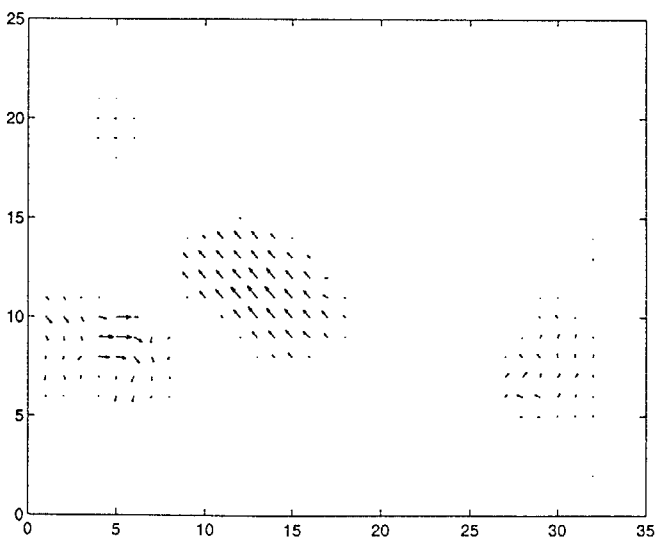
Figure 2 : a) Séquence test d'un carré texturé en translation, Répartition des modules des vecteurs vitesses représentés sous forme de niveaux de gris pour b) $\sigma_s=4$, c) $\sigma_s=2.4$.



a) Séquence initiale

b) $\sigma_s=4$ c) $\sigma_s=2.4$ 

d) masques



e) Filtrage récursifs combinés

Figure 3 : a) Séquence d'images "Taxi de Hambourg", b) et c) Flots optiques résultants pour des filtres de sélectivité différente (implantation récursive), d) masques des zones en mouvement obtenus à partir de c), e) Flot optique par la méthode des filtrages combinés (pour l'affichage, les flots optiques sont moyennés sur des carrés 8*8).