

Schémas d'estimation de mouvement multirésolution-multigrille en codage de séquences d'images

N. BAAZIZ, C. LABIT

IRISA/INRIA, Campus de Beaulieu 35042 Rennes Cedex, France
Tel: 99 36 20 00 Telex: UNIRISA 950 473 F

Résumé

Dans le cadre des schémas de codage en sous-bandes de séquences d'images, plusieurs approches multigrille et multicontrainte pour l'estimation de champs de vecteurs mouvement sont présentées ici, sur différentes décompositions pyramidales. L'estimation multicontrainte par approche Wiener permet d'introduire simultanément et efficacement la contribution de plusieurs bandes fréquentielles. Ce champ estimé de mouvement, de taille moindre, effectue une reconstruction par compensation de mouvement de qualité comparable aux usuelles méthodes monorésolution plus coûteuses. Des résultats portant sur la rapidité de convergence et l'erreur de reconstruction sont illustrés sur séquences réelles.

1 Introduction

1.1 Contexte de l'étude

Récemment un intérêt grandissant est apparu pour les représentations multirésolution en analyse et codage d'images numériques. Ces techniques consistent à décomposer l'image en un ensemble de canaux fréquentiels à largeur et à recouvrement spectral adaptés suivant les caractéristiques des filtres utilisés et permettent de définir des modèles de représentation et de structuration efficaces des données image en reliant le contenu spectral à la position spatiale du signal pour une meilleure interprétation ou compression de celui-ci. L'apparition de ces techniques (voir [1]) s'est produite de manière dispersée au sein de divers champs d'application : méthodes de décomposition en sous-bandes par des filtres miroirs en quadrature, transformations pyramidales [8], transformée en ondelettes [3],[5] qui fournit un cadre mathématique cohérent pour la définition de telles transformées.

Notre objectif étant le codage de séquences d'images télévisuelles, il importe de disposer de formes de représentation adéquates pour la compression et le rendu de l'information spatio-temporelle. Nous nous sommes intéressés dans ce contexte à deux points essentiels :

- l'obtention d'une décomposition dyadique pyramidale qui, en elle-même, possède de bonnes propriétés de localisation espace-fréquence, de représentation compacte et hiérarchique des données ainsi que d'estimation de mouvement au sein des bandes fréquentielles.
- la définition d'approches Multirésolution-Multigrille-Multicontrainte [9],[1] pour l'estimation de mouvement et la comparaison de celles-ci avec des approches plus usuelles (estimation hiérarchique de mouvement sur une pyramide passe-bas par exemple [8]).

Abstract

For image sequence coding using subband decomposition, several multigrid and multiconstraint methods for optical flow estimation are presented here. We use different pyramidal decompositions and show the efficiency of Wiener estimation scheme to simultaneously take advantage of several subbands and consecutively, having a smallest motion vector field, to obtain a quality of motion compensated reconstruction, quite similar to usual monoresolution approaches but with a least computational cost. Results about the convergence speed and reconstruction errors are illustrated using real image sequences.

1.2 Les approches Multirésolution-Multigrille-Multicontrainte

Le cadre méthodologique dans lequel s'inscrit l'ensemble de ces approches peut être décrit ainsi. Un banc de filtres d'analyse (phase de décomposition) permet de décomposer le spectre du signal d'image en M sous-bandes (avec repliement de spectre entre bandes lié au processus de décimation par M , les filtres passe-bandes n'étant pas idéaux). Un traitement T est opéré alors; il peut s'agir par exemple d'une opération non-linéaire de quantification ou de prédiction avec compensation de mouvement. Enfin, un banc de filtres de synthèse (phase de reconstruction) recompose le signal dans le domaine spatio-temporel. Ce schéma d'Analyse/Synthèse crée naturellement une structuration **multirésolution** du signal d'entrée. Pour définir clairement les notions d'algorithmique **multigrille** et celle **multicontrainte**, nous pouvons qualifier le processus de traitement T (qui dans les paragraphes sera un estimateur-compensateur de mouvement)

- de **traitement multigrille**[1],[4],[9] s'il utilise de manière hiérarchique plusieurs niveaux de résolutions; ceci fera l'objet du paragraphe 3 dans le cadre restrictif de l'estimation-compensation de mouvement.
- de **traitement multicontraint** s'il utilise de manière simultanée plusieurs bandes fréquentielles; ces techniques seront décrites au paragraphe 4 et résolues du point de vue estimation de mouvement par une approche Wiener [2].

De plus, ces notions loin d'être antinomiques peuvent s'assembler au sein d'un même schéma d'estimation.



2 Contexte méthodologique

2.1 Typologie des décompositions pyramidales

Une typologie stabilisée des diverses décompositions pyramidales a été proposée dans [1]. Les critères de classification portent sur

- la nature de la décomposition fréquentielle réalisée: passe-bas ou passe-bande
- la propriété d'orthogonalité de cette décomposition
- la propriété de séparabilité des filtres utilisés

Une comparaison a été effectuée [1] entre plusieurs méthodes de représentation multirésolution, leur efficacité est évaluée en termes d'entropie, de complexité de mise en œuvre et de qualité de reconstruction. Ainsi, ont été testées sur des séquences synthétiques et réelles les transformations pyramidales suivantes :

- la pyramide passe-bas avec noyau Gaussien, triangulaire, ou trimodal
- la pyramide passe-bande non-orthogonale (Laplacienne ou pseudo-Laplacienne).
- les pyramides passe-bande orthogonales (ondelettes de Daubechies [3]).
- les pyramides passe-bande quasi-orthogonales (ondelettes bi-orthogonales).
- la pyramide passe-bande basée sur un noyau 2D non-séparable (Quinconce).

Les conclusions qui sont apparues à l'issue de ces expérimentations montrent que certaines décompositions créent des structures spatiales orientées; le critère entropique appliqué sur le signal décomposé montre un gain par rapport au signal originel et offre une discrimination possible entre les différentes décompositions. Les meilleures performances sont atteintes par les pyramides ondelette et quinconce. Nous retiendrons le cas ondelette pour traiter de l'estimation de mouvement sur pyramides passe-bande.

2.2 L'estimation de mouvement en monorésolution

Le cadre algorithmique d'estimation de mouvement choisi comme base de référence en monorésolution est issu des méthodes différentielles. Celles-ci sont basées sur l'hypothèse essentielle d'**invariance de la luminance** c.à.d que, pour tout pixel p , connaissant son vecteur mouvement projeté apparent exact d ,

$$DFD(p, d) = I(p, t) - I(p - d, t - dt) = 0 \quad (1)$$

Une méthode d'estimation intitulée *méthode de Wiener modifiée* est introduite ici (figure 2) et validée en monorésolution (Tableau 1.a) [1]. Cette méthode combine deux algorithmes pel-récurrents connus. Le premier est l'algorithme de Walker-Rao [10] , basé sur une technique de minimisation par approche gradient qui peut être résumée par l'algorithme d'estimation itératif suivant:

$$\hat{d}^i = \hat{d}^{i-1} - \frac{\epsilon}{2} \cdot DFD(p, \hat{d}^{i-1}) \cdot \nabla I(p - \hat{d}^{i-1}, t - 1) \quad (2)$$

avec :

$$\epsilon = \frac{1}{\|\nabla f(p - \hat{d}(i-1), t-1)\|^2}$$

où i désigne la $i^{\text{ème}}$ itération et $\nabla I = (I_x, I_y)$ le gradient spatial. La seconde méthode introduite par [2] utilise l'approche Wiener et s'exprime sous la forme du processus itératif d'estimation suivant:

$$\hat{d}^i = \hat{d}^{i-1} -$$

$$\left(\begin{array}{cc} \sum_{j=1}^N (I_x^{(j)})^2 + \mu & \sum_{j=1}^N I_x^{(j)} I_y^{(j)} \\ \sum_{j=1}^N I_x^{(j)} I_y^{(j)} & \sum_{j=1}^N (I_y^{(j)})^2 + \mu \end{array} \right)^{-1} \left(\begin{array}{c} \sum_{j=1}^N I_x^{(j)} DFD^{(j)} \\ \sum_{j=1}^N I_y^{(j)} DFD^{(j)} \end{array} \right)$$

où la sommation sur N pixels porte sur un voisinage spatial causal dans ce cas monorésolution, voisinage qui sera étendu à plusieurs bandes dans le cas multirésolution-multicontraite du paragraphe 4.

Le Tableau 1.a fournit les performances de ces algorithmes dans le cadre monorésolution en termes de pourcentages de points non-compensés, de points compensés par propagation simple (\hat{d}^0), de points compensés suite au processus d'estimation itératif, de nombre moyen d'itération par pixel compensé et d'erreur quadratique moyenne (EQM) entre l'image originale et l'image compensée en mouvement.

3 Estimation de mouvement multigrille

3.1 Estimation multigrille sur une décomposition passe-bas

De nombreux travaux [8],[4] utilisant une décomposition passe-bas (Gaussienne bien souvent) ont déjà illustré l'efficacité de stratégies multigrilles pour un processus d'estimation de primitives -dont le mouvement- en termes d'amélioration de l'estimation mais aussi de réduction du coût calculatoire. Ces méthodes nécessitent essentiellement la définition

- d'une stratégie de parcours (ascendante, descendante, oscillatoire) au sein de la structure pyramidale. Nous avons expérimenté ici le cas purement descendant des résolutions grossières aux plus fines.
- d'une stratégie de coopération interniveaux: opérateurs de projection, d'interpolation. Dans notre cas, le mouvement projeté et interpolé d'un niveau à l'autre constituera une valeur initiale potentielle introduite dans l'algorithme d'estimation.

Le tableau 1.b illustre les performances d'une telle approche multigrille utilisant une pyramide passe-bas sur l'algorithme de Wiener modifié.

3.2 Estimation multigrille sur une décomposition passe-bande

Une décomposition dyadique ondelette telle que décrite dans [5] et illustrée en figure 3 peut être considérée comme composée de trois sous pyramides passe-bas orientées respectivement en bandes fréquentielles horizontales, verticales et diagonales. Une extension immédiate de l'approche multigrille passe-bas à cette décomposition passe-bande orthogonale consisterait à lancer trois processus d'estimation multigrille sur chacune des sous-pyramides orientées. Cette approche, que nous avons testée, se heurte à deux inconvénients majeurs:

- l'obtention de trois champs d'estimation de mouvement indépendants entre eux pour lesquels aucun opérateur de fusion a posteriori ne se révèle efficace
- une mauvaise performance en termes de coûts opératoires et d'amélioration des estimées.

Ces inconvénients peuvent être élégamment contournés par la définition d'une approche multicontraite telle que présentée ci-dessous.

4 Estimation de mouvement multicontraite

4.1 Les approches multicontraites

Les méthodes multicontraites consistent à rechercher N équations indépendantes pour la résolution d'un même processus d'estimation. Dans le cas de l'estimation de mouvement [7],[6] il s'agira de dériver sur plusieurs fonctions spectrales ou fonctions dérivées de l'image-observation initiale (moments, dérivées d'ordre supérieur...) une même équation de contrainte telle que l'équation de contrainte du mouvement (ECM) dans notre cas:

$$\begin{cases} I_x^{(1)} d_x + I_y^{(1)} d_y = -I_t^{(1)} \\ \vdots \\ I_x^{(N)} d_x + I_y^{(N)} d_y = -I_t^{(N)} \end{cases} \quad (4)$$

où $d = (d_x, d_y)$ et I_t désigne le gradient temporel.

Cette formulation multicontraite peut être aisément introduite dans le cas d'une décomposition ondelette; les N équations portent alors sur l'observation de N sous-bandes (par exemple 4 sous-bandes à un même niveau de décomposition dyadique) et supposent que le même vecteur d est identifiable dans chacune des sous-bandes -ce qui est vrai dans le cas de filtres passe-bandes idéaux-. La méthode classique retenue pour la résolution de ce système linéaire multicontraint est l'estimation de d aux moindres carrés. La résolution numérique de ce système se heurte dans le cas présent au mauvais conditionnement du déterminant de la pseudo-inverse et de la sensibilité des résultats d'estimation aux calculs des gradients. L'approche Wiener offre à l'inverse une solution beaucoup plus élégante et fiable pour l'estimation multicontraite-multibandes.

4.2 Estimation multicontraite par approche de Wiener

Si nous reprenons les équations introduites précédemment, il est immédiat de définir un voisinage de M pixels pris simultanément sur plusieurs bandes fréquentielles d'un même niveau de résolution lors de la décomposition dyadique ondelette. La figure 1 illustre différents choix possibles de voisinage multibandes (BB, BH, HB, HH désignent les bandes filtrées de manière séparable par le filtre B passe-bas et le filtre H passe-haut d'une paire de filtres ondelettes; les expérimentations réalisées utilisent les filtres de longueur 4 introduits par [3]). Seul le champ de mouvement sur la bande basse-fréquence (bande dénotée BB) est estimé. Les autres bandes sont compensées par ce même champ et la reconstruction après compensation se fait par le banc de filtres de reconstruction ondelettes.

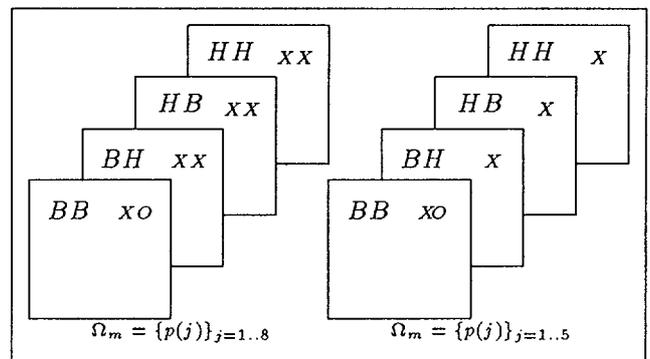


Figure 1 : Différentes configurations de voisinage multibandes Ω_m . O : pixel courant, X : pixel voisin.

5 Résultats expérimentaux et conclusion

Les tableaux 1.a, 1.b et 1.c résument les performances quantitatives et comparées des estimateurs de mouvement utilisant respectivement une approche monorésolution (base de référence), une approche multigrille sur pyramides passe-bas et une approche multicontraite par estimation de Wiener sur pyramides ondelette. Les résultats fournis portent sur la séquence "Interview" (séquence non entrelacée) dont une image est montrée en Fig. 5 ainsi que la différence temporelle entre deux images successives, la décomposition passe-bande ondelette est montrée dans la figure 3, la reconstruction par compensation de mouvement d'un champ de taille $S/4$ obtenu par la méthode multicontraite, S désignant la surface de l'image initiale, et l'image d'erreur de reconstruction sont respectivement montrées en fig. 5. Nous constatons une reconstruction parfaite visuelle pour l'estimation d'un champ de taille quatre fois moindre et une vitesse de convergence de l'algorithme d'estimation plus faible que dans le cas monorésolution ce qui tend à valider l'approche multibande-multicontraite. Ce travail se poursuivra par le test conjoint des approches multigrille et multicontraite ainsi que l'introduction de cet estimateur de mouvement dans un schéma complet de codage de séquences par compensation de mouvement.



a)

Paramètres	$imax = 5, SDFD = 2, \mu = 5$		
Algorithme	Walker	Wiener	Wiener mod.
% non comp.	18.48	13.64	3.85
%comp.(prop.)	64.92	70.08	71.20
%comp.(iter.)	16.60	16.29	24.96
Moy.(iter.)	1.93	1.76	3.55
\sqrt{EQM}	1.95	1.45	1.1

b)

Niveaux	2	1	0
% non comp.	3.24	2.99	3.02
%comp.(prop.)	76.67	78.92	77.50
%comp.(iter.)	20.09	18.10	19.48
Moy.(iter.)	11.42	5.59	3.47
\sqrt{EQM}	2.51	1.51	0.95

c)

Paramètres	$imax = 5, SDFD = 2, \mu = 5$		
%no. comp.	4.48	$\sqrt{EQM}(BB_1)$	1.46
%comp.(prop.)	72.98	$\sqrt{EQM}(HH_1)$	1.13
%comp.(iter.)	22.55	$\sqrt{EQM}(HB_1)$	1.28
moy.iter	3.08	$\sqrt{EQM}(BH_1)$	1.63
		$\sqrt{EQM}recons.$	2.43

Table 1 : a) comparaison des estimateurs : b) estimation multigrille sur pyramide passe-bas : c) estimation multicontraite sur pyramides ondelette.

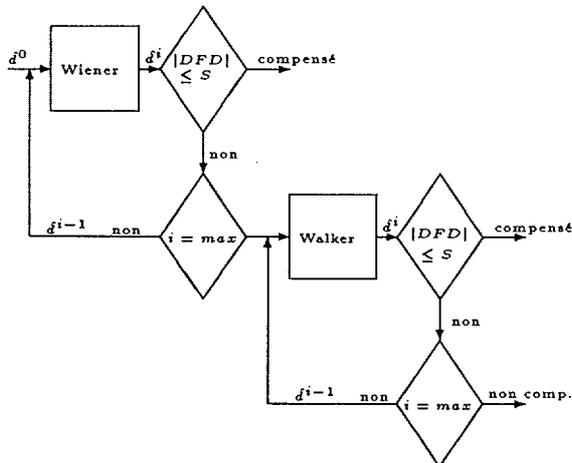


Figure 2 : Algorithme de Wiener modifié.

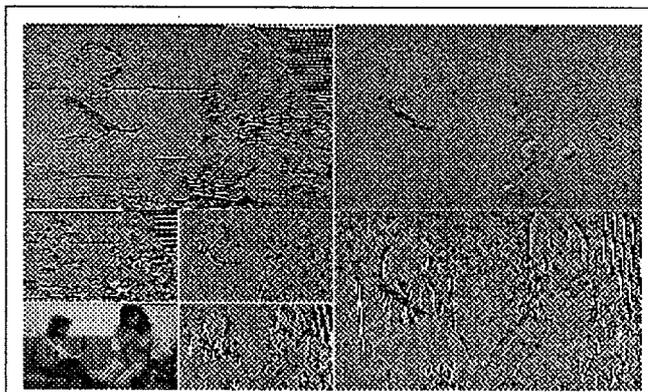


Figure 3 : Pyramides ondelette à 2 niveaux d'une image de "INTERVIEW" générée par des filtres à 4 coefficients.

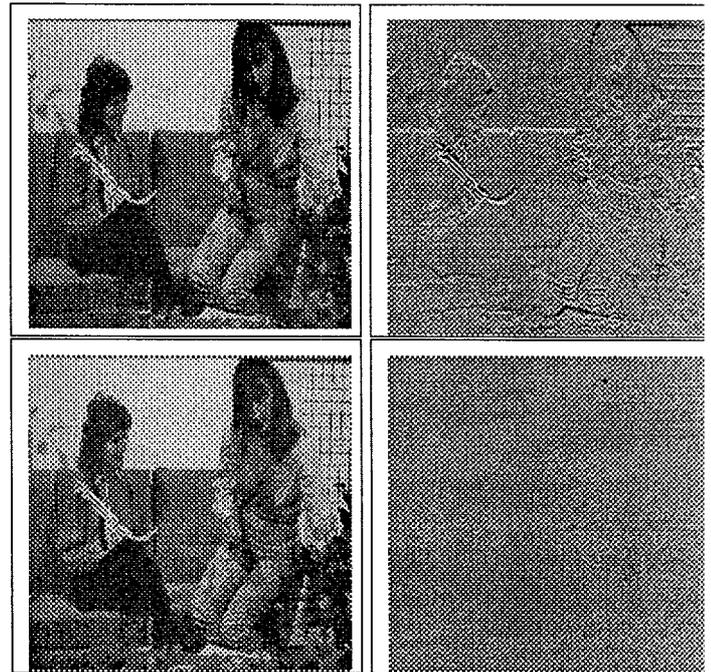


Figure 4 : En haut, une image de la séquence "INTERVIEW" et la différence temporelle avec l'image suivante, la racine de l'EQM = 16.80, en bas l'image compensée par les estimées multicontraite et image d'erreur correspondante.

Bibliographie

- [1] BAAZIZ N. and LABIT C. *Multigrid motion estimation on pyramidal representations for image sequence coding*. Technical Report 572, IRISA/INRIA, Rennes, February 1991.
- [2] BIEMOND J. , LOOIJENGA L. and BOEKKEE D. E. A pel-recursive wiener-based displacement estimation algorithm. *Signal Processing*, 13:399-412, 1987.
- [3] DAUBECHIES I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36:961-1005, September 1990.
- [4] ENKELMANN W. Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences. *CVGIP*, 43:150-177, 1988.
- [5] MALLAT S. G. A theorie for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674-693, July 1989.
- [6] MILLOUR C. *Contribution a la vision dynamique : une approche multi-resolutions et multi-traitements*. PhD thesis, Université de Paris-Sud, France, Mars 1989.
- [7] MITICHE A., WANG Y. F. and AGGARWAL J. K. Experiments in computing optical flow with the gradient-based, multiconstraint method. *Pattern Recognition*, 20(2):173-179, 1987.
- [8] ROSENFELD A. et al. *Multiresolution image processing and analysis*. Editor Springer-Verlag, 1984.
- [9] TERZOPOULOS D. Image analysis using multigrid relaxation methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):129-139, March 1986.
- [10] WALKER D. R. and RAO K. R. Improved pel-recursive motion compensation. *IEEE Transactions on Communications*, 32(10):1128October 1984.