HAL
open science

Représentation structurelle d'images fixes par transformée en ondelettes orientées

Guillaume Jeannic, Vincent Ricordel, Dominique Barba

To cite this version:

Guillaume Jeannic, Vincent Ricordel, Dominique Barba. Représentation structurelle d'images fixes par transformée en ondelettes orientées. GRETSI 2007, Sep 2007, Troyes, France. hal-00251797

HAL Id: hal-00251797

https://hal.science/hal-00251797

Submitted on 12 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Représentation structurelle d'images fixes par transformée en ondelettes orientées

Guillaume Jeannic, Vincent Ricordel, Dominique Barba

Institut de Recherche en Communications et en Cybernétique de Nantes
École Polytechnique de l'Université de Nantes, la Chantrerie : rue Christian Pauc, BP 50609, 44306 Nantes Cedex 03, France
prénom.nom@univ-nantes.fr

Résumé − En dépit de son succès, la transformée en ondelettes 2-D standard n'est pas complètement adaptée pour représenter des contours ou des textures orientées d'une image. Quelques transformées ont été proposées pour pallier ce manque soit en augmentant la famille génératrice de fonctions analysantes, soit en adaptant localement la transformée en ondelettes. Nous nous inscrivons dans cette dernière démarche pour proposer une représentation adaptative de l'information selon deux directions locales : la direction locale de régularité et sa direction orthogonale, donnant lieu à une représentation structurelle de l'information en fonction des sous-bandes, et une méthode pour déterminer les orientations privilégiées au sein d'une région de l'image.

Abstract − Despite its success, the standard 2-D wavelet transform is not fully efficient at representing edges or oriented textures in images. Several transforms have been proposed to overcome this lack by increasing the number of analysis functions or by adapting locally the wavelet transform. In this last manner we propose an adaptive representation that decomposes the information along two directions : the local direction of regularity, and its orthogonal, producing a structural decomposition of the information, and a method to detect the number of orientations in an area of the image.

1 Introduction

La transformée en ondelettes (TO) bidimensionnelle séparable a prouvé son efficacité à représenter des singularités horizontales, verticales ou ponctuelles mais éprouve des difficultés à représenter efficacement les contours qui ne sont pas strictement horizontaux ou verticaux. Ainsi les coefficients d'ondelettes de tels contours vont se retrouver de façon redondante dans les différentes sous-bandes d'une même résolution.

Ces dernières années, de nouvelles représentations d'images ont été proposées pour surmonter ce défaut des ondelettes conventionnelles. Certaines proposent une analyse suivant un nombre supérieur d'orientations que l'horizontale et la verticale au prix d'une certaine redondance. D'autres adoptent une analyse adaptative en ondelettes basée sur une représentation de la géométrie de l'image analysée. Cette géométrie peut être représentée de façon explicite par les contours extraits de l'image [1, 2], ou de façon implicite par une carte des orientations qui peut être construite selon des critères de minimisation de l'énergie de sous-bandes hautes fréquences. Cette information est ensuite utilisée pour filtrer localement l'image selon une direction donnée. Ce filtrage peut se faire selon des droites discrètes orientées[3], en utilisant des échantillons particuliers d'un voisinage[4, 5] ou en utilisant des valeurs interpolées selon une direction donnée[6, 7].

À part le cas des ondelettes de Chappelier[5] qui adopte une approche originale par sous-échantillonnage en quiconque, deux transformées en ondelettes unidimensionnelles sont appliquées sur l'image, accompagnées de deux sous-échantillonnages, l'un selon les colonnes, l'autre selon les lignes. L'ordre de ces sous-échantillonnages influe sur les orientations des filtrages possibles, toutefois avec ces dernières approches, l'ordre choisi est le même pour toute l'image. Ce qui conduit à ce que, même lorsque les contours sont bien considérés, ou avec la transformée en ondelettes classique lorsque les contours sont parfaitement horizontaux ou verticaux, les singularités orthogonales à ces contours soient réparties entre différentes sous-bandes. Par exemple dans le cas de la transformée en ondelettes classique, les détails horizontaux et verticaux se retrouvent dans des sous-bandes respectives. Nous proposons ici une nouvelle représentation pour laquelle les singularités orthogonales aux contours sont représentées au sein d'une seule et même sous-bande.

Notre représentation en ondelettes est implantée via un schéma de lifting orienté similaire à [6] mais où l'ordre des sous-échantillonnages varie en fonction du choix de la première orientation de filtrage. Pour chaque bloc de l'image, cette orientation de filtrage est déterminée à partir de la mesure de l'orientation du gradient en tous points de ce bloc. Cette mesure nous permet aussi de caractériser si une région de l'image présente une seule, plusieurs ou aucune orientation privilégiée. Les blocs de même orientation sont aussi fusionnés avec la contrainte d'une segmenta-

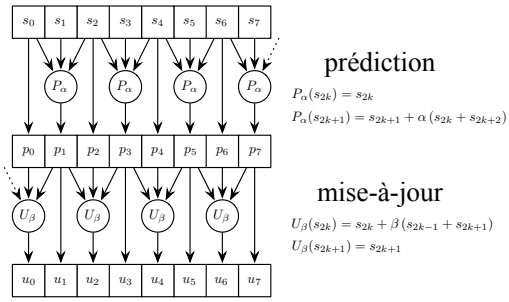


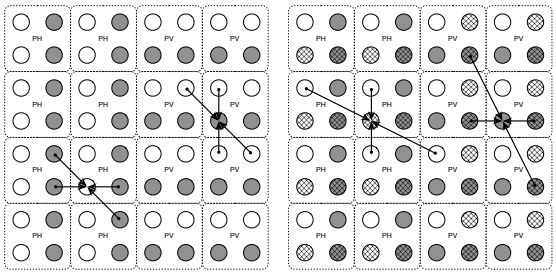
FIG. 1 – Schéma de lifting 1-D avec une étape de prédiction et une étape de mise-à-jour (avec typiquement $\alpha = -0,5$ et $\beta = 0,25$ pour l'ondelette 5/3).

tion organisée autour d'une structure d'arbre quaternaire. L'image est finalement filtrée localement le long de cette orientation, et également dans la direction orthogonale.

2 Schéma de lifting orienté

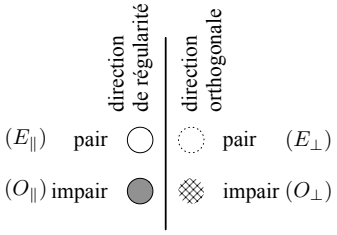
Le schéma de lifting[8] permet une implantation de la transformée en ondelettes 1-D de complexité plus faible qu'en utilisant des bancs de filtres. Les échantillons impairs sont prédits à partir des échantillons pairs, ces derniers sont ensuite mis-à-jour à partir des erreurs de prédiction précédemment calculées (voir figure 1). La transformée en ondelettes 2-D séparable standard est implantée par deux lifting successifs suivant les lignes et suivant les colonnes.

Le schéma de lifting offre aussi une certaine souplesse dans le choix des échantillons utilisés. Par exemple, pour une implantation 2-D de la transformation en ondelettes selon une orientation plus proche de l'horizontale que de la verticale, on peut choisir de ne pas se restreindre à utiliser des échantillons sur une même ligne pour prédire ou mettre à jour un échantillon. Le schéma de lifting est alors dit basé colonne car pour l'étape de prédiction, les éléments d'une colonne impaire sont prédits à partir d'une combinaison linéaire des éléments de colonnes paires. Pour une orientation plus proche de la verticale, un lifting basé ligne est appliqué. La parité des échantillons est alors redéfinie en fonction des indices des lignes et des colonnes et de l'ordre des schémas de lifting (basé colonne et basé ligne). Tant que la parité et la linéarité de l'interpolation sont respectées, la reconstruction reste exacte[2]. Ainsi de façon générale, illustrée sur la figure 3, pour le premier filtrage dans le sens de la régularité, un élément impair ($O_\parallel$) peut être prédit par toute combinaison linéaire d'éléments pairs ($E_\parallel$), et inversement pour l'étape de mise-à-jour. Pour le deuxième filtrage dans la direction orthogonale, un élément ($O_\parallel O_\perp$) (i.e. impair pour la direction de régularité et impair pour la direction orthogonale) peut



(a) Premier filtrage dans la direction de régularité.

(b) Deuxième filtrage dans la direction orthogonale.



(c) Légende de la parité des échantillons.

FIG. 2 – Redéfinition de la parité des échantillons pour le lifting. PH désigne les blocs où un sous-échantillonnage des colonnes est appliqué en premier, PV les blocs où un sous-échantillonnage des lignes est appliqué en premier. Les désignations PH et PV des blocs sont ici choisies qu'à titre d'illustration.

être prédit uniquement par des éléments ($O_\parallel E_\perp$), et les éléments ($E_\parallel O_\perp$) uniquement par des éléments ($E_\parallel E_\perp$). Un point clé préalable au filtrage, demeure donc la phase de classification des zones de l'image en fonction de leur orientation.

3 Segmentation d'images par détection implicite d'orientations privilégiées

La détermination implicite de l'orientation de filtrage se fait souvent par minimisation de l'énergie de la sous-bande de hautes fréquences (*i.e* les erreurs de prédiction). Par exemple, dans le cas de l'ondelettes 5/3, l'énergie dans la sous-bande des hautes fréquences après un premier filtrage orienté selon θ est définie en tout point (x, y) par (1) pour $-\frac{\pi}{4} < \theta < \frac{\pi}{4}$, par (2) sinon, avec $I_{x,y}$la valeur obtenue par interpolation linéaire au point (x, y).

$$\left[I_{x,y} - \frac{1}{2}\left(I_{x-1,y-\tan\theta} + I_{x+1,y+\tan\theta}\right)\right]^2 \qquad (1)$$

$$\left[I_{x,y} - \frac{1}{2}\left(I_{x-\tan(\frac{\pi}{2}-\theta),y-1} + I_{x+\tan(\frac{\pi}{2}-\theta),y+1}\right)\right]^2 \qquad (2)$$

Cette mesure n'est toutefois pas adéquate pour déterminer si une zone de l'image est plutôt uniforme ou plutôt texturée, si elle présente une seule direction privilégiée ou plusieurs. En effet, remarquons d'abord que deux directions assez proches mais correspondant à des contours d'orientations différentes donnent une réponse qui ne correspond à aucune des deux orientations mais à celle qui minimise l'énergie globale. On observe dans ce cas un seul minimum au lieu de deux comme sur la figure 3-b. De plus cette mesure est sensible à l'interpolation linéaire qui introduit deux biais. Le premier est l'effet d'atténuation du bruit qui varie en fonction de la pondération des échantillons utilisés dans l'interpolation. Cet effet d'atténuation est maximal pour la pondération $(\frac{1}{2}, \frac{1}{2})$ de l'interpolation linéaire et a tendance à diminuer l'erreur de prédiction. L'erreur d'interpolation introduit un second biais, tendant à augmenter l'erreur de prédiction et atteignant son effet maximal aussi pour la pondération $(\frac{1}{2}, \frac{1}{2})$. Ces deux biais produisent deux effets opposés qui ne se compensent pas mutuellement. Ces deux effets impliquent que la mesure d'énergie pour une orientation θ n'est pas comparable avec une mesure pour une orientation θ' dans le cas où θ et θ' ne correspondent pas à la même interpolation. L'effet atténuateur de bruit prédomine dans les zones uniformes ou texturées sans orientation privilégiée, et ne provoque donc pas une réponse uniforme (voir figure 3-h). L'erreur d'interpolation a pour effet pratique sur une image naturelle d'introduire de faux minima locaux, ou de déplacer des minima locaux correspondant réellement à une orientation privilégiée (voir figure 3-e).

Nous proposons une métrique permettant de détecter la présence d'orientations privilégiées et leur nombre, ou alors s'il s'agit d'une zone uniforme ou texturée. Cette mesure s'appuie sur la mesure du gradient de Sobel. Si le gradient ∇I au point (x, y) n'est pas le vecteur nul, la probabilité que la direction orthogonale de $\nabla I(x, y)$ soit θ est modélisée par la loi normale de moyenne $\theta_I + \frac{\pi}{2}$ et d'écart-type $\frac{\sigma}{\|\nabla I\|}$[9] :

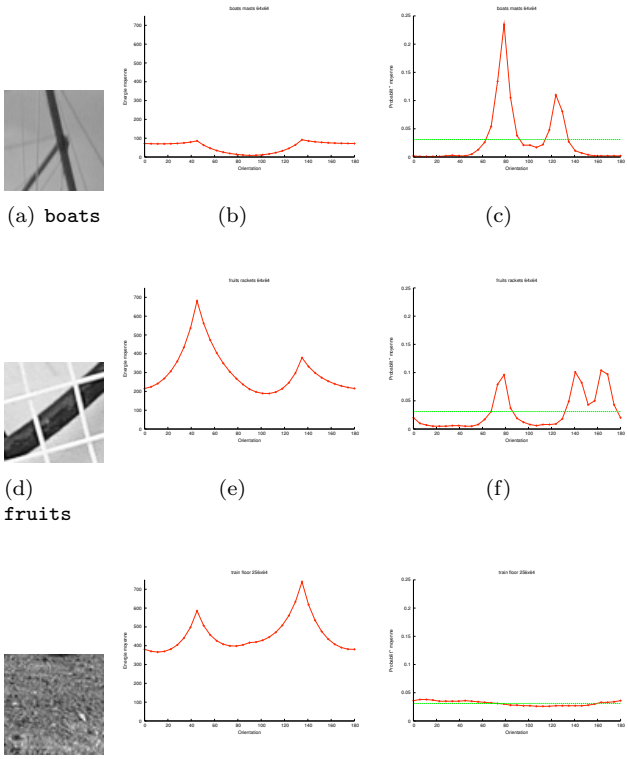
$$P(x, y, \cdot) \sim \mathcal{N}\left(\theta_I + \frac{\pi}{2}, \frac{\sigma}{\|\nabla I\|}\right),$$

où σ est l'écart-type de l'erreur sur le gradient, θ_I est l'orientation du gradient mesurée par $\arctan(\theta)$. Avec K une constante de normalisation, on a donc :

$$P(x, y, \theta) = \frac{1}{K} e^{-dist^2(\theta, \theta_I(x,y) + \frac{\pi}{2}) \cdot \frac{\|\nabla I\|^2}{\sigma^2}}.$$

où $dist(\theta_1, \theta_2)$ représente l'écart entre deux orientations définies modulo π.

Nous mesurons la moyenne des probabilités $M(\theta)$ pour tous les pixels d'une région de l'image, dont la norme du gradient est supérieure à un seuil τ.



(a) `boats` (b) (c)

(d) `fruits` (e) (f)

(g) `train` (h) (i)

FIG. 3 – Comparaison entre les méthodes de minimisation de l'énergie moyenne pour la détection des orientations privilégiées : les courbes (b, e, h) sont obtenues via les équations (1) et (2), celles (c, f, i) selon notre méthode. Sur les courbes, les orientations sont données en degré. Les blocs d'images utilisés sont donnés en (a, d, g).

Si le nombre de pixels de norme supérieure à ce seuil ne dépasse pas un quorum, par exemple égal à la hauteur pour une région carrée, la zone est considérée comme uniforme.

Soit d l'écart entre la répartition de notre mesure selon θ (en trait plein sur les figures 3-c, f, i) et une répartion uniforme (en pointillé) :

$$d = \sum_{\theta \in \Theta} \left(M(\theta) - \frac{1}{N}\right)^2,$$

où N est le nombre d'orientations testées. Si cet écart est inférieur à un seuil d_{min} alors la zone est classée en tant que texture sans orientation privilégiée. Sinon le nombre de maxima locaux où $M(\theta)$ est supérieur à $\frac{1}{N}$ donne le nombre de orientations privilégiées détectées.

4 Conclusion

Un arbre quaternaire des orientations est construit par notre méthode basée sur la modélisation stochastique de la mesure du gradient. La direction de régularité pour le

filtrage est choisie comme celle maximisant notre mesure si une orientation privilégiée est détectée, 0° sinon pour se replacer dans ce cas avec un filtrage par ondelettes classique.

L'image est filtrée d'abord le long de la direction de régularité puis dans la direction orthogonale en respectant les contraintes de parité des échantillons. Contrairement aux sous-bandes produites par la transformée en ondelettes 2-D standard où chaque sous-bande représente les singularités (horizontales, verticales ou à $\pm 45^\circ$) le long d'une orientation fixe, les sous-bandes de notre transformée représentent les singularités le long d'une orientation variable localement adaptée au contenu, ces singularités sont aussi de nature différente et se classent en 2 types des sous-bandes. Une sous-bande HFReg qui contient les hautes fréquences le long de la direction de régularité de l'image, elle est plus faible en énergie pour des contours ou des textures unidirectionnelles qui sont alors idéalement perçus comme un "bruit" le long des structures géométriques. Une autre sous-bande HFOrtho qui va contenir au contraire les hautes fréquences dans la direction orthogonale aux contours extraits, et on va y retrouver toute la structure géométrique de l'image.

Références

[1] Le Pennec (E.) et Mallat (S.), Sparse geometric image representations with bandelets, *IEEE Transactions On Image Processing*, avril 2005.

[2] Jeannic (G.), Ricordel (V.) et Barba (D.), The edge driven oriented wavelet transform : An anisotropic multidirectional representation with oriented lifting scheme, dans *Visual Communications and Image Processing*, janvier 2007.

[3] Velisavljević (V.), Beferull-Lozano (B.), Vetterli (M.) et Dragotti (P. L.), Directionlets : Anisotropic Multidirectional representation with separable filtering, *IEEE Transactions on Image Processing*, vol. 15, n° 7, 2006, p. 1916–1933.

[4] Chang (C.-L.) et Girod (B.), Direction-adaptive discrete wavelet transform for image compression, *IEEE Transactions on Image Processing*, vol. 16, n° 5, may 2007.

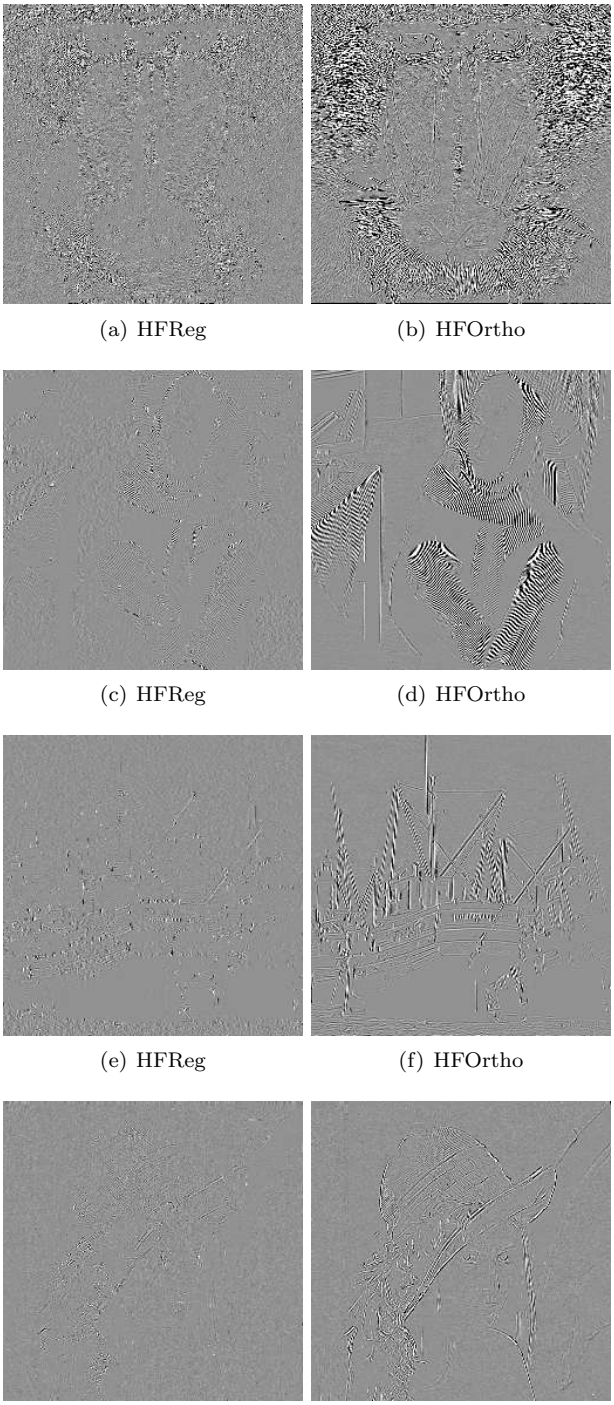
[5] Chappelier (V.) et Guillemot (C.), Oriented wavelet transform on a quincunx pyramid for image compression, dans *IEEE International Conference on Image Processing*, 2005.

[6] Ding (W.), Wu (F.), Wu (X.) *et al.*, Adaptive directional lifting-based wavelet transform for image coding, *IEEE Transactions on Image Processing*, vol. 16, n° 2, février 2007, p. 416–427.

[7] Wang (D.), Zhang (L.) et Sperenza (F.), Curved wavelet transform for image coding, *IEEE Transactions On Image Processing*, vol. 15, n° 8, août 2006, p. 2413–2421.

[8] Daubechies (I.) et Sweldens (W.), Factoring wavelet transforms into lifting steps, *J. Fourier Anal. Appl.*, vol. 4, n° 3, 1998, p. 247–269.

[9] Villéger (E.), *Constance de largeur et désoclusion dans les images digitales.* Thèse de doctorat, Université de Nice - Sophia Antipolis, décembre 2005.



(a) HFReg (b) HFOrtho

(c) HFReg (d) HFOrtho

(e) HFReg (f) HFOrtho

(g) HFReg (h) HFOrtho

Fig. 4 – Contenu des sous-bandes HFReg (à gauche) et HFOrtho (à droite) pour les images `baboon`, `barbara`, `boats` et `lena`.