

# Codage d'image par ondelettes unidimensionnelles orientées

Vivien Chappelier

Christine Guillemot

Slavica Marinković

IRISA/Université de Rennes 1, Campus de Beaulieu, 35042 RENNES

prenom.nom@irisa.fr

## Résumé

Nous présentons une technique de codage d'image basée sur une décomposition multiéchelle sur grille quinconce. Une transformée reposant sur l'application d'une ondelette unidimensionnelle dans une direction définie localement par une carte d'orientation est proposée. Une modélisation markovienne de la carte est utilisée afin de réaliser une optimisation débit-distorsion conjointe du coût de codage de la carte et des coefficients d'ondelettes. Un gain pouvant atteindre 1 dB par rapport aux techniques basées sur les ondelettes séparables est observé sur des images comportant de fortes caractéristiques directionnelles.

## Mots clefs

compression d'images, ondelettes

## 1 Introduction

La compression à bas débit repose sur la construction d'opérateurs d'approximations non-linéaires qui permettent d'approcher efficacement le signal à l'aide d'un faible nombre de coefficients. Dans ce cadre, les ondelettes sont connues pour représenter de manière optimale les signaux unidimensionnels possédant un nombre fini de discontinuités [1].

Cependant dans le cas d'images composées de régions délimitées par des contours, les discontinuités apparaissent sous la forme de courbes. Les ondelettes séparables ne sont pas la manière optimale de les représenter. Ainsi, dans les schémas de compression d'image basés sur les ondelettes séparables [2], on observe une forte concentration de coefficients d'énergie importante le long des contours de l'image. De nombreuses approches récentes tentent de résoudre ce problème en filtrant l'image le long des contours. Les contourlettes [3] reposent sur l'utilisation d'un banc de filtre directionnel. L'image est décomposée sous forme d'une pyramide multiéchelle, dont chaque niveau est redécomposé en sous-bandes orientées correspondant à une direction bien définie. Toutefois, cette transformée a le désavantage d'être redondante et peu efficace à haut débit.

La compression par bandelettes [4] se base sur un modèle géométrique explicite de l'image décrivant l'orientation des contours. Le signal d'image est alors représenté par une somme d'ondelettes séparables appliquées le long des contours et un résidu codé par ondelettes séparables.

Une technique similaire se base sur une modélisation par quad-tree. Ces approches nécessitent de coder le modèle géométrique. L'optimisation de la répartition du débit entre le codage du modèle et le codage des coefficients d'ondelette devient alors difficile.

Nous proposons ici une transformée adaptative qui consiste à appliquer une ondelette unidimensionnelle dans une direction définie par une carte d'orientation sur une pyramide d'échantillonnage quinconce. Cette carte est régularisée de manière à optimiser la répartition du débit entre les coefficients d'ondelette et la carte d'orientation. L'ondelette orientée et la décomposition multirésolution sont présentées dans la section 2. La section 3 décrit la technique de régularisation de la carte d'orientation et son lien avec l'optimisation débit-distorsion. Les résultats expérimentaux sont présentés en section 4.

## 2 Ondelette orientée

### 2.1 Ondelettes et lifting

Etant donné un signal  $S(t)$ , la transformée en ondelette dyadique consiste à représenter  $S$  par sa décomposition sur une base de fonctions d'ondelettes  $(\Psi_{k,n})_{k \in \mathbb{N}, n \in \mathbb{Z}}$ :

$$\Psi_{k,n}(t) = \frac{1}{\sqrt{2^k}} \Psi\left(\frac{t-n}{2^k}\right)$$

$$S(t) = \sum_{k,n} c_{k,n} \Psi_{k,n}(t)$$

où  $\Psi(t)$  est la fonction d'ondelette mère et  $\Psi_{k,n}(t)$  représentent des versions dilatées et translatées de cette ondelette de base. Afin de conserver les propriétés de phase linéaire et de reconstruction parfaite, une base d'ondelette biorthogonale est considérée. La transformée est alors réalisable au moyen d'un banc de filtre à réponse impulsionnelle finie à 2 canaux, itéré sur la bande basse.

Dans [5], Daubechies *et al* montrent qu'il est possible de factoriser la matrice polyphase d'un tel banc de filtre en un certain nombre d'étapes de prédiction ( $P_i$ ) et de mise à jour ( $U_i$ ), suivies d'une étape finale de mise à l'échelle ( $M$ ). Cette décomposition, appelée *lifting*, a l'avantage de généraliser la notion d'ondelettes permettant la conception d'ondelettes de seconde génération sur des espaces où la notion de fréquence n'est plus définie. En outre, elle per-

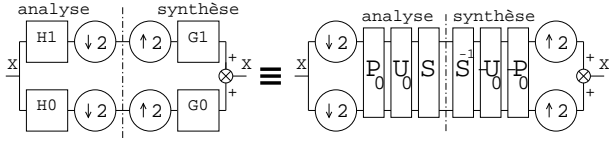


Figure 1 – Banc de filtre d'ondelettes à deux canaux et décomposition équivalente en étapes de lifting.

met une mise en oeuvre efficace du filtrage et garantit la reconstruction parfaite (Fig. 1).

L'analyse s'effectue donc de la manière suivante. La grille d'échantillonnage est tout d'abord séparée en deux cosets :

$$C : \begin{cases} S_0[n] := S[2n], \\ S_1[n] := S[2n + 1]. \end{cases}$$

L'étape de prédiction consiste à prédire chaque échantillon de  $S_1$  à partir de ses voisins dans  $S_0$  :

$$P(\alpha_i) : S_1[n] := S_1[n] + \alpha_i(S_0[n] + S_0[n + 1]).$$

De même, l'étape de mise à jour consiste à ajuster la valeur de chaque échantillon de  $S_0$  à partir de ses voisins dans  $S_1$  :

$$U(\beta_i) : S_0[n] := S_0[n] + \beta_i(S_1[n] + S_1[n - 1]).$$

Ces étapes sont réalisées l'une après l'autre alternativement. Une étape finale de mise à l'échelle permet d'ajuster l'amplitude des coefficients obtenus :

$$R(\chi) : \begin{cases} S_0[n] := \chi S_0[n], \\ S_1[n] := \frac{1}{\chi} S_1[n]. \end{cases}$$

On obtient alors, respectivement, l'approximation de  $S$  dans  $S_0$  correspondant au filtrage par  $H_0$  suivi d'une décimation, et l'erreur de prédiction dans  $S_1$  correspondant au filtrage par  $H_1$  suivi d'une décimation. Le signal original  $S$  est reconstruit en inversant les étapes de lifting, c'est à dire en appliquant successivement  $R(\frac{1}{\chi})$ ,  $U(-\alpha_{n-1})$ ,  $P(-\beta_{n-1})$ , ...,  $U(-\alpha_0)$ ,  $P(-\beta_0)$  puis en regroupant  $S_0$  et  $S_1$  :

$$C^{-1} : \begin{cases} \hat{S}[2n] := S_0[n], \\ \hat{S}[2n + 1] := S_1[n]. \end{cases}$$

## 2.2 Ondelette unidimensionnelle sur grille quinconce

De même que dans le cas unidimensionnel, il est possible de définir pour le cas bidimensionnel une structure multi-résolution issue de l'itération d'un banc de filtre à 2 canaux en s'appuyant sur la théorie des lattices. On définit la lattice issue de la matrice entière  $M$  de la manière suivante :  $L(M) = \{m \in \mathbb{Z}^2, m = Mn, n \in \mathbb{Z}^2\}$ .

En choisissant  $M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$  et  $e = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , il est possible de partitionner la grille d'échantillonnage carrée  $\mathbb{Z}^2$  en deux grilles quinconces complémentaires  $L_0^0 =$

$L(M)$  et  $L_1^0 = L(M) + e$  (Fig. 2). Cette opération peut être répétée récursivement sur  $L_0^0$  de manière à obtenir une structure multirésolution :

$$\begin{cases} L_0^k = L(M^k), \\ L_1^k = L(M^k) + M^{k-1}e. \end{cases}$$

On obtient alternativement une grille d'échantillonnage quinconce ou carrée, en remarquant que  $L(M^2) = L(2I) = 2\mathbb{Z}^2$ . De plus, une lattice quinconce  $L(M^k)$  (pour  $k$  impair) peut être vue comme une lattice carrée tournée de  $45^\circ$  dont l'écart entre chaque point vaut  $\sqrt{2}^k$ . Ainsi, ce partitionnement récursif définit une pyramide *quinconce* d'échantillonnage multirésolution  $(L_1^0, \dots, L_1^{n-1}, L_0^{n-1})$  sur  $n$  niveaux dont le facteur de sous-échantillonnage global entre chaque échelle est de  $|\det(M)| = 2$ .

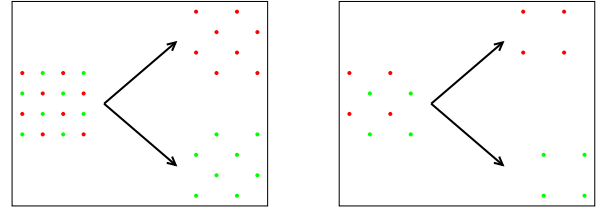


Figure 2 – Séparation de la grille d'échantillonnage en deux grilles complémentaires.

Une base d'ondelettes quinconces peut alors être établie sur une telle structure d'échantillonnage multirésolution [6]. Contrairement au cas bidimensionnel séparable (à 4 canaux), une seule fonction ondelette mère (contre 3) est nécessaire pour représenter le signal.

Plutôt que d'utiliser des ondelettes quinconces, nous proposons d'appliquer une ondelette unidimensionnelle dans une direction de filtrage choisie de manière adaptative pour chaque coefficient d'ondelette. Ces coefficients, représentant l'erreur de prédiction, sont calculés sur la grille  $L_1^k$ , tandis que l'image d'approximation est calculée sur la grille  $L_0^k$  sur laquelle la décomposition est ensuite itérée. Chaque point de la grille  $L_1^k$  possède 4 voisins. Il est donc possible de prédire un point  $p \in L_1^k$  par ses voisins dans une même ligne ou dans une même colonne (Fig. 3). Ce choix parmi les deux directions de prédiction est réalisé de manière à minimiser l'erreur de prédiction. Il définit sur  $L_1^k$  une carte binaire  $M_k$  indiquant la direction locale de filtrage, dont la connaissance est nécessaire lors de la reconstruction. Afin de calculer le coefficient d'ondelette au point  $p$ , les étapes de prédiction du lifting pour une ondelette biorthogonale unidimensionnelle sont appliquées dans la direction choisie en  $p$ . Les étapes de mise à jour nécessitent un traitement particulier. Un point  $p'$  de la grille  $L_0^k$  peut être utilisé de 0 à 4 fois pour la prédiction de ses voisins dans  $L_1^k$ , contrairement à 2 fois exactement dans le cas unidimensionnel classique. Afin de conserver la moyenne du signal d'approximation, les coefficients  $\beta_i^*$  utilisés pour

les étapes de mise à jour sont déduits des coefficients  $\beta_i$  du cas unidimensionnel en fonction du nombre  $v$  de voisins ayant utilisé  $p'$  comme prédicteur :

$$\beta_i^* = \begin{cases} \frac{2}{v}\beta_i & \text{si } v \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

L'étape de mise à jour est alors définie comme suit :

$$U(\beta_i) : S_0[n] := S_0[n] + \beta_i^* \sum_{n^* \in V(n)} S_1[n^*]$$

où  $V(n) \subset L_1^k$  est l'ensemble des voisins de  $n$  utilisant  $n$  comme prédicteur. Lorsque la direction de prédiction est identique pour tout les points de  $L_1^k$ , on retrouve exactement les mêmes étapes de lifting que pour l'ondelette unidimensionnelle. La reconstruction parfaite est assurée par l'inversibilité de chaque étape de lifting, moyennant la connaissance de la carte  $M_k$  à chaque niveau de décomposition. De plus, en supposant que la grille  $L_1^k$  est partitionnable en un nombre restreint de régions ( $R_i$ ) sur lesquelles la direction de filtrage est constante, la propriété de quasi-orthogonalité de la décomposition en ondelette est conservée.

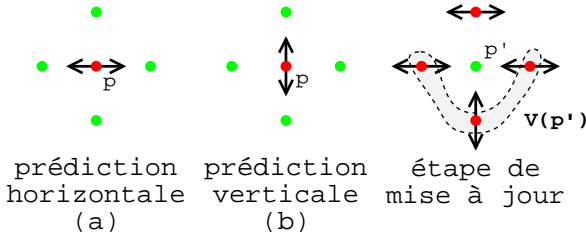


Figure 3 – La prédiction est effectuée selon l'une ou l'autre des directions (a) ou (b). La mise à jour dépend des directions choisies pour les quatre pixels voisins.

### 3 Optimisation débit-distorsion

La carte d'orientation est choisie de manière à minimiser l'erreur résiduelle de prédiction à chaque niveau, afin de compacter au maximum l'énergie du signal dans la bande d'approximation finale. Sans codage entropique, la carte d'orientation coûte environ 1 bit par pixel à coder, ce qui est prohibitif. Cependant, l'information binaire sur la direction de filtrage n'est pas toujours pertinente. En effet, lorsque la distorsion obtenue par prédiction dans une direction ou l'autre est similaire, le choix de l'orientation n'induit qu'une faible variation de la distorsion. Ceci se produit surtout sur les zones uniformes ou les zones texturées, qui constituent la majeure partie d'une image naturelle. Il est donc possible de propager l'information d'orientation à partir des contours, de manière à homogénéiser la carte d'orientation, moyennant une perte minime en distorsion. La carte peut alors être codée efficacement au moyen d'un codeur arithmétique contextuel exploitant les dépendances conditionnelles causales entre les orientations.

La carte d'orientation est modélisée sous la forme d'un graphe de Markov sur  $(L_1^k)$ . Tout noeud du graphe se trouve dans un état binaire  $m_i$  représentant l'orientation

de filtrage en ce point  $i$ . Chaque noeud est relié à ses voisins dans la même échelle  $k$  par un 8-voisinage. Ceci permet de modéliser la cohérence intra-échelle des orientations dans un voisinage. De plus, un noeud est relié à son noeud parent à l'échelle  $k - 2$  et à ses 4 noeuds fils à l'échelle  $k + 2$  lorsqu'ils existent. Ceci traduit la dépendance inter-échelle entre les orientations. Les fonctions de potentiel sont définies de manière identique sur l'ensemble des cliques  $\mathcal{C}(V_{m_i})$  du voisinage  $V_{m_i}$  de  $m_i$  par

$$\forall c \in \mathcal{C}(V_{m_i}), \phi(m_i, c) = \begin{cases} -\alpha & \text{si } \forall m_j \in c, m_j = m_i \\ +\alpha & \text{sinon} \end{cases}$$

Le théorème d'Hammersley-Clifford [7] donne la loi conditionnelle de  $m_i$  sachant son voisinage  $V_{m_i}$  à une constante  $\gamma$  près :

$$Pr(m_i | V_{m_i}) = \gamma e^{-\sum_{c \in \mathcal{C}(V_{m_i})} \phi(m_i, c)} \quad (1)$$

et son coût de codage vaut alors :

$$\begin{aligned} R(m_i | V_{m_i}) &= -\log_2(Pr(m_i | V_{m_i})) \\ &= \frac{1}{\ln(2)} \left( \sum_{c \in \mathcal{C}(V_{m_i})} \phi(m_i, c) - \ln(\gamma) \right). \end{aligned}$$

L'optimisation du débit utilisé par la carte par rapport au débit utilisé pour coder les coefficients s'effectue alors en minimisant une énergie de Gibbs globale. Celle-ci intègre à la fois l'énergie de l'erreur résiduelle de prédiction  $y_i(m_i)$  dépendant de l'orientation  $m_i$  choisie en  $i$ , et le coût de codage  $R(m_i | V_{m_i})$  de cette orientation  $m_i$  fonction des décisions prises dans le voisinage  $V_{m_i}$  du point  $i$ .

La recherche de la carte optimale  $M^*$  s'effectue alors en minimisant par échantillonnage de Gibbs du graphe de Markov l'énergie suivante :

$$M^* = \underset{M}{\operatorname{argmax}} \left( \gamma e^{-\sum_i \|y_i(m_i)\|^2 - \lambda \sum_i R(m_i, V_{m_i})} \right).$$

Ceci revient à minimiser par rapport aux cartes  $M$  le lagrangien  $J(\lambda, M) = D(M) + \lambda R(M)$  avec  $D(M) = \sum_i \|y_i(m_i)\|^2$  et  $R(M) = \sum_i R(m_i, V_{m_i})$  et  $\lambda$  fixé. La carte  $M^*$  obtenue définit alors la meilleure configuration

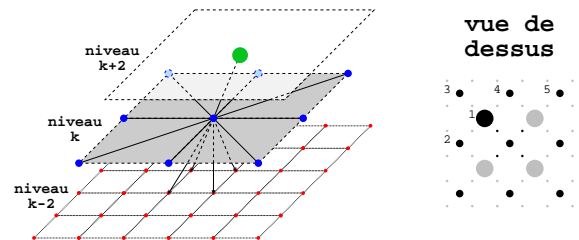


Figure 4 – Graphe markovien des dépendances. La figure 3D représente en traits forts le voisinage du point central au niveau  $k$ . La figure 2D représente en noir la localisation de ce même voisinage dans  $\mathbb{Z}^2$ . Les points les plus petits correspondent à l'échelle la plus fine. Les points numérotés de 1 à 5 indiquent la sous-partie causale du voisinage.

de filtrage sous contrainte de débit lorsqu'aucun coefficient d'ondelette n'est codé, c'est à dire la configuration maximisant la concentration d'énergie sous contrainte de débit. Cette carte est ensuite utilisée pour le codage progressif de l'image et permet d'obtenir des sous-bandes de hautes fréquences contenant moins d'énergie que l'ondelette séparable, ce qui compense le surcoût de codage de la carte.

## 4 Résultats

Nous comparons la méthode proposée à l'ondelette séparable en terme de performance débit-distorsion. Dans les deux cas, une ondelette biorthogonale 9/7 est utilisée. Le paramètre d'homogénéité  $\alpha$  du graphe de Markov est fixé de manière empirique à 0,85. La carte d'orientation est codée par un codeur arithmétique contextuel estimant la probabilité conditionnelle de chaque orientation à partir des orientations choisies sur son sous-voisinage causal (Fig. 4) et de l'équation 1. Le débit est estimé par l'entropie stationnaire des sous-bandes, à laquelle est ajoutée le coût de codage de la carte dans le cas de l'ondelette orientée.

La figure 5 illustre le gain obtenu en appliquant la méthode proposée à l'image *barbara* à un débit de 0.3 bpp. De manière subjective, les contours sont mieux restitués visuellement. La figure 6 présente les courbes débit-distorsion pour les deux méthodes appliquées à l'image *barbara*. On observe un gain objectif de 0.9 dB par rapport à l'ondelette séparable sur une large plage de débit. Cependant, à bas débit ( $< 0.1$ bpp), l'ondelette séparable présente de meilleures performances par rapport à l'ondelette orientée, qui est pénalisée par le coût fixe de la carte d'orientation. Le tableau 1 présente les résultats obtenus pour un codage sur 3 niveaux (resp. 6 niveaux) par ondelette séparable (resp. ondelette orientée) sur diverses images. Le gain obtenu par la méthode proposée varie grandement en fonction de l'image. Il est d'autant plus important que l'image comprend des caractéristiques directionnelles.

image	lena	barbara	bike	goldhill
séparable (dB)	33.98	26.93	23.66	30.51
orientée (dB)	34.29	27.87	23.95	30.55

Tableau 1 – Comparaison de l'ondelette séparable et l'ondelette orientée en terme de PSNR à 0.3bpp pour une décomposition sur 3 niveaux (resp. 6 niveaux).

## 5 Conclusion

Nous proposons une nouvelle méthode de compression d'image basée utilisant des ondelettes unidimensionnelles orientées sur pyramide d'échantillonnage quinconce multirésolution. La direction d'application de l'ondelette est définie par une carte d'orientation optimisée par un critère débit-distorsion. Nous observons jusqu'à 1dB de gain par rapport aux méthodes basées sur les ondelettes séparables. La méthode d'optimisation utilisée pour le moment est coûteuse en temps de calcul. En outre, le coût de codage de la carte est fixe et pénalise les performances de la méthode



Figure 5 – [gauche] Codage de l'image *barbara* 512x512 par ondelette 9/7 séparable sur 3 niveaux à 0.3 bits/pixel, PSNR = 26.93 dB. [droite] Codage par ondelette 9/7 orientée sur 6 niveaux à 0.3 bits/pixels, PSNR = 27.87 dB. La bande basse est à la même résolution dans les deux cas.

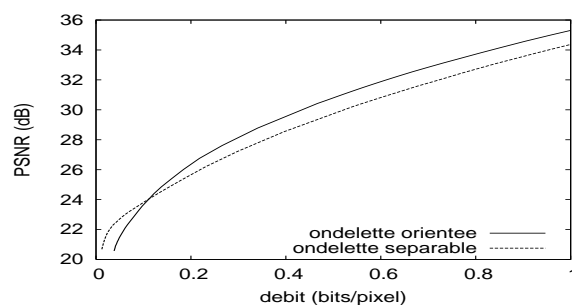


Figure 6 – Courbe débit-distorsion pour l'image *barbara* codée sur 5 niveaux d'ondelette séparable et 10 niveaux d'ondelette orientée.

présentée à bas débit. Nous orientons actuellement nos travaux vers la recherche de meilleurs algorithmes d'optimisation et d'une méthode de codage progressif de la carte.

## Références

- [1] M. Vetterli, "Wavelets, approximation and compression," Sept. 2001.
- [2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," in *IEEE Trans. Image Process.*, Apr. 1992, vol. 1, pp. 205–220.
- [3] M. Do and M. Vetterli, "Pyramidal directional filter banks and curvelets," in *IEEE Proc. Int. Conf. Image Proc.*, 2001.
- [4] Erwan LePennec and Stephane Mallat, "Sparse geometric image representation with bandelets," in *submitted to IEEE Trans. on Image Process.*, 2003.
- [5] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," in *J. Fourier Anal. Appl.*, 1998, vol. 4(3), pp. 247–269.
- [6] A. Gouze, M. Antonini, and M. Barlaud, "Quincunx lifting scheme for lossy image coding," Sept. 2000.
- [7] P. Clifford, "Markov random fields in statistics," in *Disorder in physical systems, A volume in honour of J. M. Hammersley*, G. R. Grimmett and D. J. A. Welsh, Eds. Clarendon Press, Oxford, 1990.