

Bandelettes et représentation géométrique des images

Erwan LE PENNEC, Stéphane MALLAT

CMAP / École Polytechnique
91128 PALAISEAU - FRANCE
lepennec@cmapx.polytechnique.fr

Résumé – Les bases de bandelettes décomposent une image selon des vecteurs multiéchelles allongés le long d’un flot géométrique indiquant des directions de régularité de l’image. L’optimisation de cette géométrie est faite par un algorithme rapide de meilleur base qui permet d’obtenir des résultats de compression

Abstract – Bandelet bases decompose an image along multiscale vectors that are elongated in the direction of a geometric flow. This geometric flow indicates directions in which the image grey levels have regular variations. The geometric flow is optimized with a fast best basis algorithm, which leads to image compression results.

1 Géométrie et image

La recherche d’une bonne représentation est un problème centrale du traitement d’image. On cherche ici des bases permettant une représentation creuse des images, c’est-à-dire permettant de bien les approcher avec peu de coefficients. Cette propriété est en effet à la base de nombreux algorithmes de compression ou de débruitage.

Les bases d’ondelettes orthogonales permettent d’obtenir une telle représentation. Elles ne sont cependant pas optimales pour les images. Bien que très efficaces pour les zones régulières, les textures homogènes et les singularités ponctuelles, elles ne peuvent exploiter la régularité de nature géométrique des contours. Leur capacité à exploiter une régularité homogène fait qu’elles sont cependant à la base de la construction des bandelette présentées ici. On notera la base d’ondelettes bidimensionnelles du domaine Ω construite à partir de l’ondelette ψ et de la fonction d’échelle ϕ correspondante [2] :

$$\left\{ \begin{array}{l} \phi_{j,m_1}(x_1) \psi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1) \phi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1) \psi_{j,m_2}(x_2) \end{array} \right\}_{(j,m_1,m_2) \in I_\Omega} \quad (1)$$

L’utilisation de la nature géométrique des images et son exploitation pour améliorer les représentations ont déjà été proposées dans de nombreux travaux. On peut citer notamment Donoho et Candès [1] qui ont introduit les frames de curvelets, des structures anisotropes, permettant d’exploiter la régularité d’ordre 1 des contours, Cohen et Matei [3] qui proposent un schéma d’approximation non-linéaire basée sur les ondelettes ENO et Vetterli et al. [8] qui proposent un principe de découpage en carrés dyadiques des images avec une approximation polynômiale par morceaux sur chacun des morceaux. Toutes ces représentations permettent d’exploiter en partie la régularité géométrique mais ne permettent pas de l’exploiter complètement. La régularité des contours est limitée pour les deux premières représentations tandis que les deux dernières ne sont pas adaptées à des contours lissés. Déformés

pour suivre la régularité des images comme les bandelettes utilisées en 2001 [4], les bases de bandelettes présentées ici permettent une exploitation efficace de cette régularité géométrique.

2 Base de bandelettes

Au lieu de représenter la géométrie des images par des contours, qui sont bien souvent mal définis, celle-ci est définie par des flots géométriques. Les vecteurs d’un flot géométrique indiquent des directions où l’image a localement des variations régulières. Les bases orthonormées de bandelettes sont construites en divisant l’image en régions où ces flots restent parallèles. La détermination effective de ce flot géométrique est repoussé à la section suivante.

Les bandelettes sont construites à partir d’ondelettes bidimensionnelles déformées le long du flot géométrique. Dans une région Ω , le flot géométrique est un champ de vecteur $\vec{\tau}(x_1, x_2)$ qui donne une direction dans laquelle la fonction f est régulière dans un voisinage. Pour construire des bases orthonormées, une condition de parallélisme doit être imposée sur le flot. Celui-ci est choisi soit parallèle verticalement, $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1)$, soit parallèle horizontalement, $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_2)$. Pour permettre plus de flexibilité, cette condition de parallélisme n’est imposée que dans des sous-régions Ω_i de Ω . La région Ω est donc partitionnée en des régions Ω_i munies d’un flot parallèle soit verticalement soit horizontalement. Dans le cas où la fonction f est uniformément régulière sur une région Ω_i , le flot géométrique perd son sens et n’est pas défini. En pratique, on se restreindra à des partitions en carrés dyadiques.

On définit alors une base pour chaque région Ω . Si aucun flot n’est défini, on utilise une base d’ondelettes bidimensionnelles de $L^2(\Omega)$. Sinon, cette base est remplacée par une base de bandelettes. La construction de cette base est maintenant présentée dans le cas où le flot est parallèle verticalement, le cas horizontal étant similaire. Le flot s’écrit donc $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1)$ et quitte à le renormaliser

on peut l'écrire sous la forme $\vec{r}(x_1) = (1, c'(x_1))$. On pose alors $x_{\min} = \inf_{x_1} \{(x_1, x_2) \in \Omega\}$ et on définit la ligne de flot comme une courbe intégrale du flot. Celle-ci satisfait l'équation $x_2 = c(x_1) + c_0$ où

$$c(x) = \int_{x_{\min}}^x c'(u) du, \quad (2)$$

et c_0 est un paramètre de translation. Par construction, l'image a une variation régulière le long de ces lignes. Afin d'exploiter cette régularité, les ondelettes bidimensionnelles sont déformées pour suivre le flot. L'image déformée $Wf(x_1, x_2) = f(x_1, x_2 + c(x_1))$ est régulière le long des lignes horizontales (x_2 fixé). On utilise alors pour la région déformée

$$\Omega' = W\Omega = \{(x_1, x_2) : (x_1, x_2 + c(x_1)) \in \Omega\} \quad (3)$$

la base d'ondelettes bidimensionnelles de $L^2(\Omega')$:

$$\left\{ \begin{array}{l} \phi_{j,m_1}(x_1) \psi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1) \phi_{j,m_2}(x_2) \\ \psi_{j,m_1}(x_1) \psi_{j,m_2}(x_2) \end{array} \right\}_{(j,m_1,m_2) \in I_{\Omega'}} \quad (4)$$

L'opérateur de déformation W étant orthogonal, l'application de son inverse à ces ondelettes donne une base orthonormée de $L^2(\Omega)$ que l'on appelle base d'ondelettes déformées :

$$\left\{ \begin{array}{l} \phi_{j,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1)) \\ \psi_{j,m_1}(x_1) \phi_{j,m_2}(x_2 - c(x_1)) \\ \psi_{j,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1)) \end{array} \right\}_{(j,m_1,m_2) \in I_{\Omega'}} \quad (5)$$

Afin d'exploiter la régularité de la fonction f selon le flot, on remplace les ondelettes déformées de la forme

$$\{\phi_{j,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1))\}_{m_1,m_2} \quad (6)$$

par la famille de fonctions engendrant le même espace

$$\{\psi_{l,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1))\}_{l>j,m_1,m_2} \quad (7)$$

Cette opération est appelée *bandelettisation* et on vérifie qu'elle s'implémente par une simple transformée en ondelettes monodimensionnelles discrètes. Les fonctions

$$\psi_{l,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1)) \quad (8)$$

sont appelées bandelettes du fait de leur support allongés selon les lignes de flots. La base orthonormée de bandelettes du domaine Ω est donc définie par :

$$\left\{ \begin{array}{l} \psi_{l,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1)) \\ \psi_{j,m_1}(x_1) \phi_{j,m_2}(x_2 - c(x_1)) \\ \psi_{j,m_1}(x_1) \psi_{j,m_2}(x_2 - c(x_1)) \end{array} \right\}_{j,l>j,m_1,m_2} \quad (9)$$

Pour une partition donnée $\Omega = \cup_i \Omega_i$ et les flots géométriques correspondants, on définit ainsi une base orthonormée de bandelettes ou d'ondelettes (si il n'y a pas de flots) pour chaque Ω_i . L'union de ces bases forment une base par blocs de $L^2(\Omega)$. L'orthogonalité des bandelettes et des ondelettes peut-être relâchée : en utilisant une ondelette biorthogonale monodimensionnelle ψ et sa fonction d'échelle ϕ , la même construction donne une base biorthogonale pour chaque $L^2(\Omega_i)$ et donc une base biorthogonale par blocs de $L^2(\Omega)$.

La transformation en bandelettes et son inverse s'implémente par un algorithme rapide de type banc de filtres [6] de complexité $O(N^2)$ où N^2 est le nombre de pixels de l'image. Pour éviter les effets de blocs, on utilise un schéma de lifting adapté permettant de traverser les bords [6].

3 Optimisation de la géométrie

La difficulté principale est de déterminer une géométrie adaptée à l'application souhaitée. Dans le cas de la compression, il s'agit de minimiser l'erreur de reconstruction $\|f - f_R\|$ à budget de R bits fixés pour son code qui doit décrire à la fois les coefficients de bandelettes et la géométrie spécifiant la base.

Afin de diminuer le nombre de paramètres de la représentation et de pouvoir obtenir une partition optimale par un algorithme rapide, on se restreint à des partitions en carrés dyadiques. Pour chacun des carrés, on doit décider si il y a un flot géométrique, si celui-ci est parallèle horizontalement ou verticalement et enfin quel doit être ce flot. On étudie maintenant la détermination du flot à parallélisme fixée, l'optimisation des autres paramètres venant plus tard.

Si le flot est choisi parallèle verticalement, $\vec{r}(x_1, x_2) = (1, c'(x_1))$, sur un carré Ω , ce flot va être déterminé pour garantir que l'image a des variations régulières selon lui. Étant donné un noyau de lissage $\theta(x_1, x_2)$, on minimise pour cela une énergie de flot

$$\mathcal{E}(\vec{r}) = \int_{\Omega} \left| f \star \frac{\partial \theta}{\partial x_1}(x_1, x_2) + c'(x_1) f \star \frac{\partial \theta}{\partial x_2}(x_1, x_2) \right|^2 dx_1 dx_2 \quad (10)$$

Dans le cadre de la compression, le flot doit être représenté avec un petit nombre de paramètres quantifiés. La fonction $c'(t)$ est alors paramétrée dans une famille de *box spline* b dilatés d'un facteur 2^l :

$$c'(t) = \sum_n \alpha_n b(2^{-l}t - n) \quad (11)$$

Les paramètres α_n s'obtiennent alors par la résolution du système linéaire associé à la minimisation de l'énergie du flot et sont quantifiés a posteriori. Le facteur 2^l qui correspond à la régularité du signal est ajusté dans l'optimisation globale de la géométrie. Le cas du flot parallèle horizontalement se traite de manière similaire.

L'optimisation globale de la géométrie s'obtient par une approche de type Lagrangienne [7] : on recherche la base permettant de minimiser un coût de la forme

$$\|f - f_R\|^2 + \lambda R = D + \lambda R \quad (12)$$

dans le cas de la compression. En faisant varier λ , on est assuré de trouver ainsi une base minimisant l'erreur pour une contrainte sur R fixée. On montre par ailleurs [6] qu'à pas de quantification Δ fixé pour les coefficients de bandelettes, on peut déterminer la valeur λ optimale : $\lambda = (3\Delta^2)/(4\gamma_0)$ où γ_0 représente le coup de codage d'un coefficient de bandelettes non nul. Si l'on suppose que le pas de quantification de la géométrie est fixé, le seul paramètre restant est donc ce pas de quantification Δ .

L'optimisation globale de la géométrie est obtenue à l'aide d'un algorithme rapide de recherche de meilleure base utilisant le caractère additif du coût lagrangien :

$$D + \lambda R = \sum_i D_i + \lambda R_i \quad (13)$$

où l'indice i indique une restriction à la région Ω_i . Pour chaque carré dyadique Ω , on teste toutes les possibilités de

flot (absence, parallèle verticalement ou horizontalement) pour différents paramètres l dans l'estimation du flot. On conserve alors celle minimisant le coût lagrangien sur cette région. Ce calcul nécessite un total de $O((\log_2 N)^2 N^2)$ opérations [6]. La partition optimale est alors obtenue par un algorithme rapide de bas en haut le long des branches de la segmentation en arbre quaternaire. Pour tout carré Ω_i , une partition en carrés plus petits $\Omega_i = \cup_l \Omega_l$ est préférable si elle abaisse le coût lagrangien :

$$D_i + \lambda R_i \geq \sum_l D_l + \lambda R_l . \quad (14)$$

On part alors d'une partition de l'image en tous les carrés à la profondeur maximale J de l'arbre, ceux-ci étant typiquement de côté 4. A la profondeur $J - 1$, on compare le coût lagrangien des carrés de côté 8 avec la somme des coûts lagrangiens des 4 carrés de côté 4 les composant et on conserve la partition qui minimise ce coût. A la profondeur suivante $J - 2$, on compare le coût des carrés de taille 16 avec la somme des coûts des 4 partitions optimales des 4 sous-carrés les composant. Ce procédé d'agrégation est itéré en remontant dans l'arbre et conduit à une partition minimisant le coût lagrangien. Cette optimisation ne nécessite donc au total que $O((\log_2 N)^2 N^2)$ opérations.

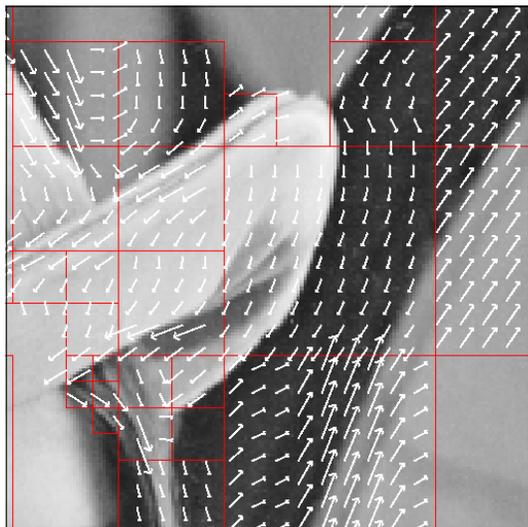


FIG. 1 – Détail de la segmentation optimale et du flot géométrique associé pour l'image de Lena et un pas de quantification Δ de 30. (Les carrés ne contenant pas de flot n'ont pas été représentés)

4 Compression

Les performances d'une base en compression dépendent de la capacité d'approximation de celle-ci. Pour une classe d'image régulière par morceaux, on montre [5] que les bandelettes ont de bien meilleures performances d'approximation non-linéaire que les ondelettes. Cette amélioration se traduit par une amélioration pour la compression y compris dans le cas d'images naturelles.

La figure 2 présente la courbe de performance pour les images de Lena et Barbara. Le codage en bandelettes est supérieur au codage en ondelettes en améliorant le PSNR

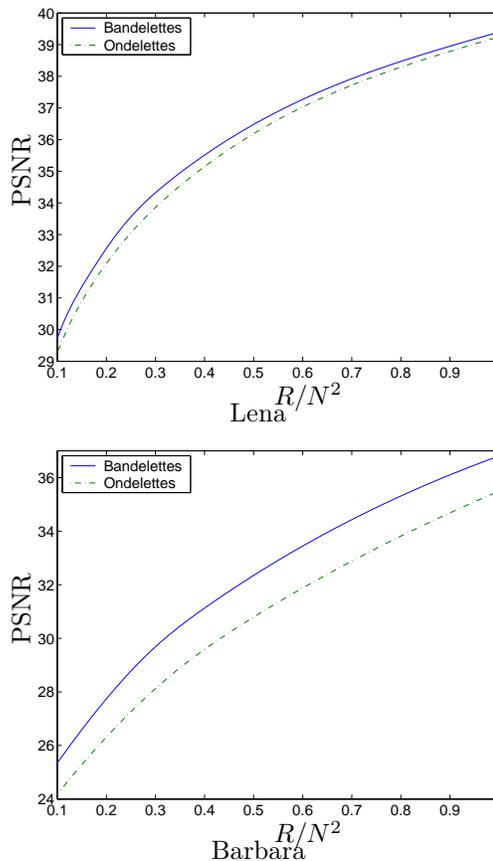


FIG. 2 – Courbe de distorsion en PSNR du codage en bandelettes (lignes pleines) et en ondelettes (lignes pointillées) pour les images de Lena et Barbara en fonction de la compression en bits par pixel.

pour un même taux de compression par 0,5 db pour Lena et 1,5 db pour Barbara pour une large gamme de débit. La figure 3 indique que le gain perceptuel est encore plus substantiel. Les images obtenues avec les bandelettes restent régulières le long de la géométrie utilisée tandis que celles obtenues avec les ondelettes présentent des effets oscillants associés à la grille carrée des ondelettes. Le résultat est une impression visuelle nettement meilleure avec les bandelettes. De plus grâce au lifting traversant les bords des carrés de la partition, les effets de blocs, que l'on aurait pu craindre, sont très faibles.

5 Conclusion

La géométrie utilisée dans la construction des bandelettes est basée non pas sur la notion de contours mais sur celle d'un champ de vecteur indiquant une direction de régularité locale des images. Ce flot géométrique est simplifié par une segmentation adaptative en carré et des conditions de parallélisme. L'optimisation de cette géométrie pour une application donnée s'obtient par un algorithme rapide de recherche de meilleure base et conduit à une amélioration dans le cadre du codage par rapport aux ondelettes.



FIG. 3 – Résultat de compression : (a) Lena , (b) détail de Lena et (c) détail de Barbara. L'image de Lena est comprimée à 0,22 bits par pixel en bandelettes et en ondelettes et les PSNR respectifs sont de 33,04 db et 32,55 db. Pour Barbara, l'image est comprimée à 0,39 bits par pixels et les PSNR respectifs sont de 31,22 db et 28,50 db.

Références

- [1] E. Candès et D. Donoho. Curvelets : A surprisingly effective nonadaptive representation of objects with edges. Dans L. L. Schumaker, A. Cohen, et C. Rabut, éditeurs, *Curves and Surfaces fitting*. Vanderbilt University Press, 1999.
- [2] A. Cohen, W. Dahmen, et R. DeVore. Multiscale decompositions on bounded domains. Rapport technique 113, IGPM, May 1995.
- [3] A. Cohen et B. Matei. Nonlinear subdivisions schemes : applications to image processing. Dans A. Iske, E. Quack, et M. Floater, éditeurs, *Tutorial on multiresolution in geometric modelling*. Springer, 2002.
- [4] E. Le Pennec et S. Mallat. Représentation d'images par bandelettes et application à la compression. Dans *GRETSI*, Toulouse, Septembre 2001.
- [5] E. Le Pennec et S. Mallat. Non linear image approximation with bandelets. rapport technique, CMAP / École Polytechnique, 2003.
- [6] E. Le Pennec et S. Mallat. Sparse geometrical image representation with bandelets. *IEEE Transactions on Image Processing*, 2003. soumis (<http://www.cmap.polytechnique.fr/~lepennec>).
- [7] K. Ramchandran et M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Transactions on Image Processing*, 2(2) :160–175, April 1993.
- [8] R. Shukla, P. L. Dragotti, M. N. Do, et M. Vetterli. Rate-distortion optimized tree structured compression algorithm for piecewise smooth images. *IEEE Transactions on Image Processing*, 2002. soumis.