Compression des images par ondelettes : intérêt d'une approche géométrique

Jacques Froment

PRISME, UFR Math. et Info., Université Paris 5 R. Descartes
45 rue des Saints-Pères, 75270 Paris cedex 06, France
et
CMLA, Ecole Normale Supérieure de Cachan,
61 avenue du Président Wilson, 94235 Cachan cedex, France.

Jacques.Froment@math-info.univ-paris5.fr

Résumé L'avènement du nouveau standard international de compression d'images fixes JPEG 2000 concrétise une décennie de recherches qui a vu s'imposer les représentations par bases d'ondelettes. Les algorithmes de compression par ondelettes ou par paquets d'ondelettes souffrent cependant d'un défaut connu sous le nom de pseudo-phénomène de Gibbs et qui affecte la qualité des contours de l'image décodée. Nous rappelons comment une représentation géométrique des contours, à l'aide de lignes de niveau, permet d'éviter ce phénomène tout en conservant pour les zones texturées les avantages d'une approche ondelettes. Ce schéma est suffisamment souple pour être adapté à des images particulières, comme nous l'illustrons sur une image d'angiographie : sur ce type d'images le critère de sélection de l'information géométrique doit être différent de celui appliqué aux images naturelles.

Mots clés Compression d'images, paquets d'ondelettes, lignes de niveau, angiographie.

1 Introduction

En utilisant une base d'ondelettes biorthogonales à la place d'une base en cosinus, le nouveau standard international de compression d'images fixes JPEG 2000 [10] améliore considérablement l'ancien standard JPEG. L'avantage d'une représentation en ondelettes sur une représentation purement fréquencielle, comme celle en cosinus, tient à la localisation des atomes dans le plan espace-fréquence : au voisinage des lignes de contour d'une image, la décroissance des coefficients en ondelettes est plus rapide. Au prix d'un accroissement de complexité, l'utilisation d'une base de paquets d'ondelettes adaptée aux caractéristiques de l'image améliore encore le pavage du plan espace-fréquence. Pour autant, toute représentation espace-fréquence reste assujettie au principe d'incertitude d'Heisenberg qui interdit une localisation parfaite en espace des atomes conservant une localisation fréquencielle. Un contour idéal, défini comme la trace sur l'image de l'arête d'un objet, est une structure purement géométrique qui échappe ainsi à l'analyse harmonique classique. Alors que sur une image correctement échantillonnée un contour serait légèrement flou, dans la pratique on observe qu'un contour est suffisamment singulier pour engendrer sur son voisinage des oscillations, lors de l'opération de seuillage et de quantification des coefficients. Bien que d'intensité moindre ce phénomène est semblable à celui de Gibbs, lequel apparaît lors de l'opération de seuillage des coefficients de Fourier d'une fonction non régulière. Aux taux de compression élevés, les structures oscillantes forment des artefacts très génants pour la perception visuelle. Aux taux plus faibles, bien qu'apparemment imperceptibles, elles peuvent apparaître lors de post-traitements tels qu'une déconvolution, une détection de contour ou un rehaussement de contraste.

Afin d'éviter l'apparition du pseudo-phénomène de Gibbs, des recherches s'orientent vers une modification de la notion d'ondelettes qui puisse doter ces dernières d'une meilleure localisation géométrique. La géométrie de ces nouveaux atomes peut être fixe, comme dans le cas des *ridgelets* [5] et des *curvelets* [1], ou elle peut s'adapter aux contours de l'image, ainsi en est-il des *bandelets* [13]. Dans chaque cas, le concepteur doit transiger avec le principe d'incertitude. L'approche que nous présentons est différente au sens où elle propose une représentation purement géométrique des lignes de contour. Elle consiste d'abord à remarquer que dans toute image u l'information pertinente peut être grossièrement séparée entre un *sketch* v, c'est-à-dire une esquisse semblable à une image de bandes dessinées, régulière à l'intérieur de chaque région dont les bords correspondent aux contours perceptuels, et entre les microtextures w qui sont ici définies comme la différence sur chaque région de l'image initiale avec celle de la bande dessinée [9, 6]. L'intérêt principal du modèle $u = v+w$ pour la compression est l'utilisation de représentations différentes, adaptées aux structures particulières des images u et w. Dans [6], nous avons proposé de représenter le *sketch* par des régions dont les bords sont formés par les lignes de niveau perceptuellement les plus significatives [3], tandis que les microtextures bien localisées en fréquence sont quantifiées dans une base d'ondelettes biorthogonales.

Dans cet article, nous illustrons la souplesse du modèle $u = v + w$ en proposant une variante de la méthode de compression structurée. Tout d'abord, nous considérons une image très particulière qui ne rentre pas dans le cadre antérieur des images "naturelles" : une angiographie. Il s'agit d'une radiographie des vaisseaux sanguins d'un organe comme le cerveau, le cœur ou le rein, obtenue après injection d'un produit iodé opaque au rayon X. Le produit de contraste permet de visualiser la circulation du sang dans les veines. L'information portée par le *sketch* est alors formée par la géométrie des principales veines et artères, tandis que les tissus forment les zones texturées. Pour ce type d'images, le critère de sélection des lignes de niveau les plus pertinentes doit être modifié. Enfin, nous expérimentons une représentation des textures dans une base optimale de paquets d'ondelettes. Sur ce dernier point, le lecteur pourra se reporter au manuscript [8] pour obtenir une présentation plus détaillée.

2 Les formes définies par des lignes de niveau

Notre approche géométrique passe par la définition de régions P_i, que nous appellerons "formes" ("*shapes*" suivant la terminologie introduite dans [12]), telles que les lignes de contour perceptuel correspondent le mieux possible au bord de ces régions. Dans le cas où l'image originale u est "naturelle", au sens où elle correspond à une photographie numérisée d'une scène du monde réel, les lignes de contour peuvent être représentées à l'aide de portions de lignes de niveau, lesquelles joueraient le rôle d'atomes de la perception [3]. On définit une forme comme la région à l'intérieur de la frontière donnée par une ligne de niveau, sachant qu'une ligne de niveau est une courbe de Jordan fermée ou rencontrant le bord de l'image. Si l'on considère la décomposition de l'image en ses ensembles de niveau inférieur $[u < \lambda] = \{x \in \Omega / u(x) < \lambda\}$, pour λ parcourant l'ensemble des niveaux de gris et $\Omega \subset \mathbb{R}^2$ le domaine de l'image u, une forme est donc une composante connexe d'un de ces ensembles de niveau qui est "remplie" pour "boucher les trous" dans le cas où la composante n'est pas simplement connexe. Nous noterons une telle forme $[u < \lambda]^k$. Une décomposition de l'image en formes s'écrit donc

$$\mathcal{P} = (P_i)_i = \{[u < \lambda]^k\}_{k,\lambda}, \tag{1}$$

les valeurs λ et les indices k devant être correctement choisis de manière à ce que les bords ∂P_i coïncident au mieux avec les contours perceptuels. Une sélection des formes les plus significatives peut être opérée en utilisant deux critères complémentaires [8] : le rapport isopérimétrique de la forme [7] et le nombre de jonctions en T le long de sa frontière [3]. Ces critères tiennent compte de la physique de la formation des images naturelles : les contours des objets physiques, en général peu oscillants, génèrent des formes de faible rapport isopérimétrique ; les phénomènes d'occlusion ou de transparence le long des contours génèrent un nombre important de jonctions en T.

En imagerie médicale cependant, la plupart des images ne sont pas "naturelles" et les critères précédents ne permettent pas une sélection efficace des formes. Un critère de contraste entre l'intérieur et l'extérieur de la forme peut cependant souvent s'appliquer. Il est particulièrement pertinent dans le cas de l'angiographie, puisque le liquide de contraste circulant dans les veines crée justement un contraste important. Notons que si le contraste d'une forme est peu souvent utilisé, c'est qu'il ne permet pas d'obtenir une description de l'image invariante par changement de contraste. Dans le contexte de la compression, la perte de l'invariance ne porte pas à conséquence. Nous définissons le contraste d'une forme S par

$$\mathcal{C}(S) = \frac{\sqrt{\sum_{x \in \partial S} \left(u^+(x) - u^-(x)\right)^2}}{\mathrm{per}(S)}, \tag{2}$$

$u^+(x)$ désignant le niveau de gris à l'extérieur de la frontière au point du bord x, $u^-(x)$ le niveau de gris à l'intérieur et $\mathrm{per}(S)$ la longueur de la frontière en nombre de points. L'algorithme de décomposition de l'image en ses formes significatives consiste simplement à calculer l'ensemble de toutes les formes (à l'aide de la *FLST* [12]), puis à ne retenir que les plus significatives au sens de la fonction de contraste, la quantité de formes enregistrées étant donnée par le débit allouée pour le codage du *sketch*.

3 Codage et reconstruction du *sketch*

Le *sketch* est défini à partir de la géométrie des formes, à laquelle on ajoute des informations sur les niveaux de gris de l'image u. De manière à préserver la géométrie, le bord des formes est codé sans perte, en utilisant un codage de chaîne classique (Freeman). La structure particulière des ensembles de niveau dont sont issus ces bords permet cependant quelques optimisations, notamment dans le cas de courbes se recouvrant partiellement [8]. Afin de minimiser la quantité de niveaux de gris à coder, seules deux valeurs de gris sont conservées pour chaque forme : la valeur moyenne sur le côté intérieur de la forme et celle sur le côté extérieur. Ces valeurs sont quantifiées puis codées en utilisant un codage arithmétique prédictif. Un point $x \in \Omega$ peut appartenir au côté (intérieur ou extérieur) de plusieurs formes : il peut donc y avoir, durant la phase de décodage, ambiguïté quant au niveau à attribuer à x. Grâce à la propriété d'inclusion des ensembles de niveau, il est cependant toujours possible de retrouver la valeur la plus proche de $u(x)$, que nous noterons $g(x)$. Lorsque x n'appartient à aucun côté d'aucune forme, on pose $g(x) = 0$.

La reconstruction d'une image v à partir des bords des formes $(P_i)_i$ et des valeurs en ces bords données par la graine g, est un problème d'interpolation. Nous utilisons le modèle AMLE (*Absolute Minimizing Lipschitz Extension)* qui permet une interpolation régulière à l'intérieur des formes, quelque soient les conditions aux bords [2, 4]. L'interpolation est réalisée en résolvant une équation d'évolution par un schéma d'Euler implicite.

4 Codage du résidu

Si un nombre suffisant de formes pouvait être codé, le résidu serait uniquement composé de microtextures. Cette situation n'est cependant pas réaliste, car le débit du *sketch* serait trop important. Dans la pratique on observe donc dans w une information géométrique, due à la représentation imparfaite des lignes de contours, ainsi qu'une information basse fréquence, due à la différence moyenne de niveaux de gris entre les régions originales et celles reconstruites. Notre approche nous permet cependant de considérer ces informations comme non significatives. Comme les microtextures consituent la composante essentielle du résidu, nous devons utiliser une représentation prioritairement bien adaptée aux motifs très oscillants. Les bases de paquets d'ondelettes offrent une localisation dans les hautes fréquences bien meilleures que les bases d'ondelettes, puisque dans ces dernières la hauteur des rectangles d'Heisenberg est nécessairement croissante avec la fréquence. Nous avons ainsi choisi de représenter le résidu dans une base de paquets d'ondelettes qui est optimale dans le contexte de la compression. L'algorithme de choix de la meilleure base et le codage des coefficients correspondent au codeur défini par F. Meyer, A. Averbuch et J-O. Strömberg dans [11].

5 Illustration sur une image d'angiographie

L'algorithme de compression structurée est illustré à la figure 1. Nous l'avons comparée avec la compression par paquets d'ondelettes seul (image G). Bien que la compression structurée ne permette pas d'atteindre le même rapport signal à bruit, la qualité de l'image reconstruite est visuellement plus satisfaisante (comparer F à G). Pour l'essentiel, cela tient à l'absence d'oscillations au voisinage des contours des veines, grâce à leur représentation géométrique par des lignes de niveau. Cet absence d'artefact est rendu encore plus visible si on applique à l'image décodée certains post-traitement comme un filtre passe-haut, ici un filtre "plus net" (comparer H à I).

6 Conclusion

Si une simple modification du critère de sélection des formes significatives permet d'appliquer avec succès la compression structurée sur des images d'angiographie, la spécificité de ce type d'image conjuguée à l'information géométrique fournit par le *sketch* offre des perspectives de compression supérieure. Par exemple, la redondance des bords des formes générée par le parallélisme local des contours des veines pourrait être exploité.

Références

1. E. Candès and D. Donoho. Curvelets : a surprisingly effective nonadaptive representation of objects with edges. In L. Schumaker A. Cohen, C. Rabut, editor, *Curve and Surface Fitting : Saint-Malo 1999*. Vanderbilt Univ. Press, 2000.
2. F. Cao. Absolutely minimizing lipschitz extension with discontinuous boundary data. *Note aux C.R. Acad. Sci. Paris*, t.327(I) :563–568, 1998.
3. V. Caselles, B. Coll, and J-M. Morel. Topographic maps and local contrast changes in natural images. *Int. J. Comp. Vision*, 33(1) :5–27, 1999.
4. V. Caselles, J-M. Morel, and C. Sbert. An axiomatic approach to image interpolation. *IEEE Trans. on Image Proc.*, 7(3) :376–386, 1998.
5. M. N. Do and M. Vetterli. Orthonormal finite ridgelet transform for image compression. In *Proc. of ICIP'2000*, volume 2, pages 367–370, 2000.
6. J. Froment. A functional analysis model for natural images permitting structured compression. *ESAIM :COCV Control, Opt. and Cal. of Var.*, 4 :473–495, 1999.
7. J. Froment. Perceptible level lines and isoperimetric ratio. In *Proc. of ICIP'2000*, volume 2, pages 112–115, 2000.
8. J. Froment. Image compression through level lines and wavelets packets. Manuscript. `http://www.math-info.univ-paris5.fr/~jf/publi.html`, 2001.
9. J. Froment and S. Mallat. Second generation compact image coding with wavelets. In C.K. Chui, editor, *Wavelets - A Tutorial in Theory and Applications*, pages 655–678. Academic Press, 1992.
10. M.J. Gormish, D. Lee, and M.W. Marcellin. Jpeg 2000 : overview, architecture and applications. In *Proc. of ICIP'2000*, volume 2, pages 29–32, 2000.
11. F. Meyer, A. Averbuch, and J-O. Stromberg. Fast adaptive wavelet packet image compression. *IEEE Trans. on Image Proc.*, 9(5), 2000.
12. P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Trans. on Image Proc.*, 9(5) :860–872, 2000.
13. E. Le Pennec and S. Mallat. Image compression with geometrical wavelets. In *Proc. of ICIP'2000*, volume 1, pages 661–664, 2000.

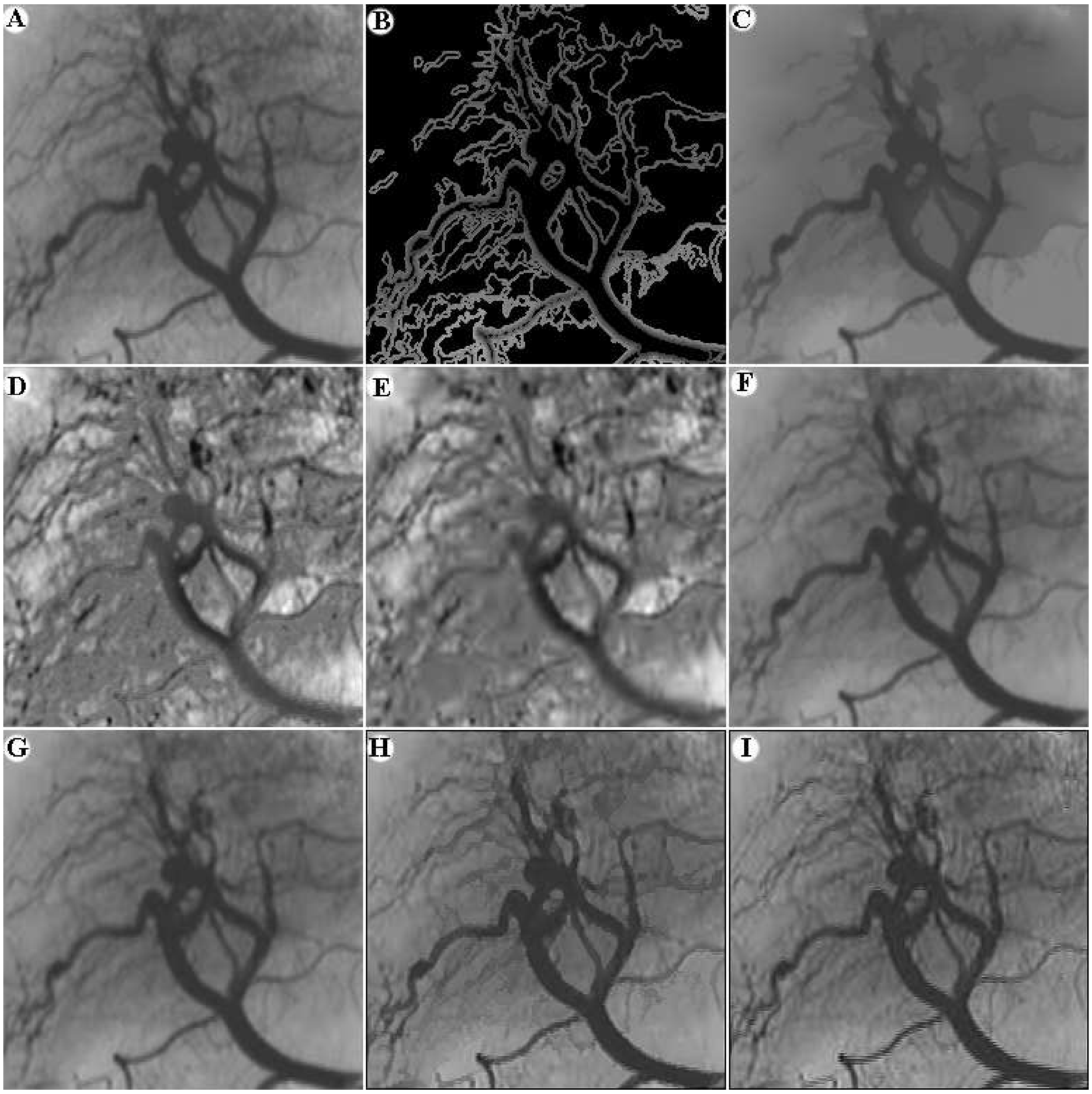

Fig. 1. Compression d'une image d'angiographie, au taux de 20 (débit de 0.4 bpp). **A** : image originale 256×256, 8 bpp. **B** : carte des valeurs aux bords utilisées pour reconstruire le *sketch* (c'est la graine g) ; codage de la géométrie des lignes de niveau sur 0.193 bpp et codage des niveaux de gris sur 0.007 bpp, soit un total de 0.2 bpp. **C** : sketch v obtenu à partir de B par AMLE ; les principales veines sont bien reconstruites tandis que les tissus sont remplacés par des régions homogènes. **D** : résidu A-C ; les "microtextures" sont ici essentiellement composées des tissus. **E** : résidu décodé après compression par le codeur en paquets d'ondelettes de F. Meyer, A. Averbuch et J-O. Strömberg ; débit de 0.2 bpp ; noter la bonne restitution des tissus. **F** : C+E décodage de l'image A compressée par notre méthode ; débit de $0.2 + 0.2 = 0.4$ bpp, PSNR $= 40.1$ db. **G** : pour comparaison, image A décodée après compression par le codeur de F. Meyer, A. Averbuch et J-O. Strömberg ; débit de 0.4 bpp, PSNR $= 41.0$ db. Si le rapport signal à bruit est supérieur, la qualité visuelle est inférieure à cause des oscillations au voisinage des contours (pseudo-phénomène de Gibbs). **H** : image F filtrée par un filtre passe-haut "plus net", pour mettre en évidence d'éventuels artefacts. **I** : image G filtrée par le même filtre, pour mettre en évidence le pseudo-phénomène de Gibbs.