

Extraction de contours multirésolution pour un codage d'images par bandelettes

G. Jeannic¹

V. Ricordel¹

D. Barba¹

¹ IRCCyN / Équipe Image et VidéoCommunication

École Polytechnique de l'Université de Nantes

La Chantrerie, rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3

{guillaume.jeannic, vincent.ricordel, dominique.barba}@univ-nantes.fr

Résumé

Depuis quelques années, un certain nombre de transformées orientées a été développé pour améliorer les inconvénients de la transformée en ondelettes bidimensionnelle conventionnelle. Parmi elles, la transformée en bandelettes s'adapte au contenu de l'image pour opérer un filtrage de type ondelettes le long des contours et, exactement, exploiter la régularité dans la direction de ces contours autour d'une bande étroite. Cette transformée requiert une extraction de contours adaptée. Nous proposons donc une méthode d'extraction de contours répondant à ces contraintes. Notre approche produit, à différents niveaux de résolution une carte des contours extraits.

Mots clefs

transformée en ondelettes, transformée en bandelettes, extraction de contours

1 Introduction

La transformée en ondelettes bidimensionnelle utilisée actuellement pour la compression d'image fixe dans la norme JPEG2000 résulte du produit tensoriel de deux transformées en ondelettes unidimensionnelles appliquées suivant les lignes et les colonnes de l'image. Ces ondelettes bidimensionnelles séparables ont prouvé leur capacité à détecter les singularités horizontales, verticales, ou ponctuelles (assimilées à des singularités diagonales). Toutefois, l'analyse de contours qui ne sont pas parfaitement horizontaux, verticaux ou diagonaux conduit à une représentation suboptimale de l'information : les coefficients d'ondelettes représentant de tels contours se retrouvent éparpillés dans les différentes sous-bandes au lieu d'être fortement présents dans une unique sous-bande. Dans un contexte de compression avec perte, la transformée en ondelettes couplée à des techniques de quantification classiques entraîne plus d'artéfacts visuels sur les contours d'orientations quelconques que sur les contours purement horizontaux, verticaux ou diagonaux (voir figure 1).

D'autre part, si la transformée en ondelettes est par exemple à même de détecter une singularité le long d'une colonne, dans le cas d'un contour horizontal, une singula-

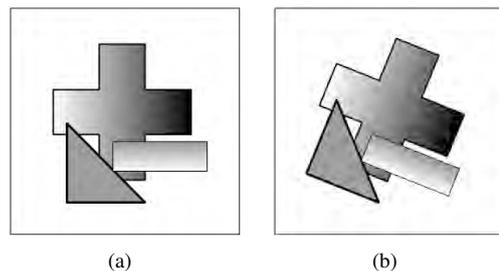


Figure 1 – À taux de compression égal avec JPEG2000, ici d'un ratio 1 : 31, on constate moins de défauts visuels au niveau des contours horizontaux, verticaux et diagonaux (a) que sur les contours orientés différemment (b).

rité est détectée pour chaque élément de ce contour. Les ondelettes n'exploitent donc pas complètement la régularité d'un contour pour le représenter.

Depuis quelques années, de nouvelles représentations d'images ont été développées de façon à pallier à ces défauts de la transformée en ondelettes tout en conservant ses avantages, à savoir sa multirésolution, sa bonne localisation en espace et en fréquence, ainsi que sa décimation critique. On peut ainsi trouver deux approches différentes dans la littérature : les transformées qui, sur le même modèle que la transformée en ondelettes, prennent en compte plus d'orientations que les simples horizontales et verticales, et d'autres transformées qui elles adaptent la transformée en ondelettes conventionnelle au contenu de l'image. Parmi ces transformées adaptatives, la transformée en bandelettes requiert une information sur le contenu de l'image, à savoir ses contours. Un schéma d'extraction adapté sera présenté ainsi que les résultats obtenus.

1.1 Transformées à bases fixes

Parmi les transformées où la base de projection est fixe, la transformée en "ridgelettes" [1] décompose une image en contours rectilignes traversant toute l'image. L'idée est de pouvoir représenter un tel long contour avec un unique coefficient en rapport avec son amplitude et situé dans une sous-bande caractérisant son épaisseur. En pratique une

telle décomposition est obtenue par transformée en ondelettes unidimensionnelle sur différentes projections de Radon de l'image. À part dans des conditions très restrictives [2], cette transformation introduit de la redondance, ce qui n'est pas souhaité dans un contexte de codage de source. D'autre part, les contours dans une image sont rarement rectilignes et aussi longs que les dimensions même de l'image. Pour résoudre le second point, la transformation est appliquée par blocs avec les mêmes problématiques d'effets de blocs que la transformée en cosinus discrète. L'idée est reprise par la transformée en "curvelettes" [3] qui introduit une analyse multiéchelle en appliquant une transformée en "ridgelettes" par blocs après une décomposition en sous-bandes. Ces deux transformées nécessitent l'implémentation d'un opérateur de rotation ce qui n'est pas simple dans le domaine discret. À cette fin a été développée la transformée en "contourlettes" [4] à base de filtres non séparables. Elle propose une décomposition multirésolution et multidirectionnelle qui s'approche de la décomposition en curvelettes. Outre la redondance introduite par cette transformée, la forte taille des filtres employés introduit des artefacts visuels.

1.2 Transformées à bases adaptatives

La transformée en ondelettes bidimensionnelle peut être implémentée selon un schéma de "lifting". La technique est appliquée en utilisant des échantillons de l'image le long des lignes (respectivement des colonnes). La seconde approche que l'on retrouve dans la littérature utilise cette technique mais en utilisant des échantillons qui ne sont pas tous situés sur une même ligne (respectivement une même colonne). Cela revient à appliquer une transformée en ondelettes le long d'une direction particulière. Le choix d'une telle direction est délicat. Dans la transformée en "curved wavelet" [5], une transformée en ondelettes est appliquée selon cinq orientations différentes pour chaque bloc d'une image. Une direction est ensuite retenue pour chaque bloc, celle minimisant l'énergie contenue dans la sous-bande des hautes fréquences. Le choix de la direction est alors implicite et peut correspondre à l'orientation globale des contours de tout un bloc. Dans la transformée en bandelettes [6], le choix de la direction est la conséquence d'une extraction de contours dans l'image. Exactement une bande autour d'un contour donné est réalignée par translation selon les lignes ou les colonnes. Une transformée en ondelettes bidimensionnelles est ensuite appliquée sur cette bande réalignée. Notre bande est donc filtrée selon deux directions : la direction du contour qui est caractérisée par une certaine régularité, et la direction orthogonale au contour redressé qui est caractérisée par une forte singularité. Cette régularité dans la direction du contour est ensuite exploitée par application d'une transformée en ondelettes unidimensionnelle sur les lignes (respectivement colonnes) des détails horizontaux (respectivement verticaux) pour un contour réaligné sur l'horizontale (respectivement la verticale, voir aussi la figure 2.

Ces deux transformées induisent un supplément de codage que ce soit pour indiquer le choix d'orientation pour chaque bloc de l'image, ou pour représenter les contours dans l'image.

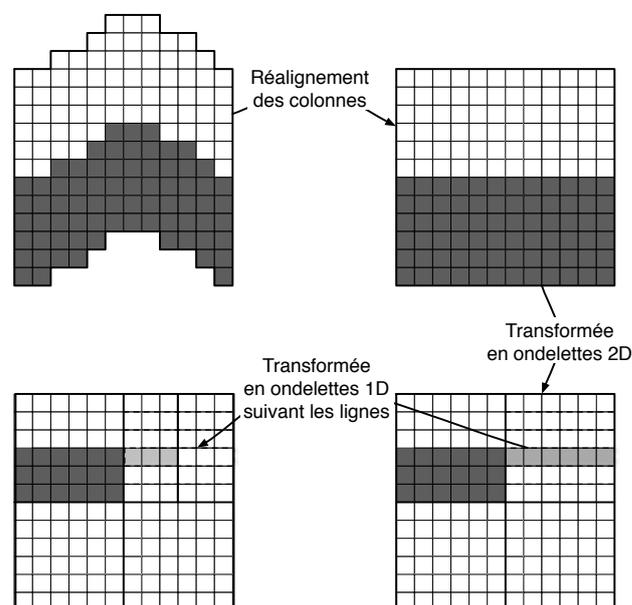


Figure 2 – Exemple de décomposition en bandelettes d'un contour réaligné sur l'horizontale.

2 Une extraction de contour adaptée aux bandelettes

Notre attention se porte sur les transformations adaptatives qui n'introduisent pas de redondance mais un supplément de codage et dont l'implémentation peut se faire selon un schéma de lifting comme la transformée en ondelettes bidimensionnelle conventionnelle. L'adaptation de la transformée en "curved wavelet" au contenu de l'image dépend de la taille des blocs utilisés. La transformée en bandelettes permet une adaptation plus flexible pour filtrer des contours quelconques, mais ne permet toutefois pas de traiter des zones texturées orientées. On peut dès lors imaginer filtrer les bandes autour des contours forts d'une images par la transformée en bandelettes et utiliser la transformée en ondelettes courbées sur les autres zones, dites appartenant au "fond", qui peuvent contenir des textures orientées.

La transformée en bandelettes requiert l'extraction des contours de l'image. Ces contours guident le filtrage de l'image, ils doivent donc être transmis au décodeur. Pour cela Le Pennec [6], les représente par des courbes splines. Dans ce contexte de codage, il est plus intéressant d'extraire des contours longs, plutôt que des contours morcelés. Il serait simple d'éliminer les contours de longueurs insuffisantes par un post-traitement après extraction, mais il est encore plus intéressant d'arriver à représenter plusieurs segments de contours en les reliant en un seul et

unique contour, et donc une seule et unique spline.

Le Pennec propose dans sa thèse une extraction *ad hoc* de contour basé sur une extraction de points de contours forts à différents niveaux de résolution. Les points sont reliés entre eux selon des contraintes basées entre autre sur l'orientation du gradient en ces points, de telle sorte qu'une continuité existe suivant les orientations. Dans le cas où un point a plusieurs successeurs possibles, la liaison entre ces points est rompue. Ceci afin de ne pas créer de coins et ainsi éviter la superposition de bandes.

Notre schéma d'extraction de contours fait aussi intervenir une analyse multi-échelle afin de pouvoir extraire des contours plus longs et être moins sensible au bruit. Notre approche permet en plus de produire pour chaque niveau de résolution une carte complète des contours extraits. Une transformation en bandelettes peut être donc réalisée à chacun des niveaux. Enfin nous souhaitons gérer la détection des noeuds et des coins dans l'image afin de pouvoir les traiter de façon adaptée et non comme le serait le fond de l'image.

3 Extraction de contours multirésolution

Le schéma que nous proposons combine de trois mécanismes (voir figure 3) : la construction d'une pyramide multi-échelle de l'image [7], la prolongation de contours par abaissement progressif du seuil du gradient, et l'affinement en position de contours obtenus à résolution inférieure par la méthode des contours actifs. Ce schéma multirésolution permet de trouver le compromis entre contours suffisamment longs et suffisamment précis. L'affinement en position de contours permet de conserver les structures détectées à basse résolution. Des contours qui auraient été détectés de façon morcelés à cause de chutes locales du gradient ou à cause du bruit, conservent grâce à cette étape leur unité. La prolongation de contours par abaissement progressif du seuil du gradient permet à une résolution donnée de faire apparaître de nouveaux germes ou de prolonger des contours déjà détectés à plus basse résolution en s'appuyant sur les points de contours forts.

3.1 Prolongation de contours à une résolution donnée

Un simple détecteur des extrema locaux du gradient produit un ensemble de points correspondant aux contours de l'image mais aussi aux textures et au bruit. Nous voulons donc retenir et relier entre eux, uniquement les points appartenant aux contours forts de l'image tout en restant robuste au bruit (voir figure 4).

Notre approche est itérative et vise à chaque boucle, à prolonger les extrémités des contours déjà formés ou à faire émerger de nouveaux germes. Exactement pour une itération donnée, le seuil du gradient est abaissé faisant apparaître de nouveaux points. Aux extrémités des contours déjà formés, apparaissent d'éventuels points successeurs qui sont reliés s'ils minimisent une fonction de coût. Cette

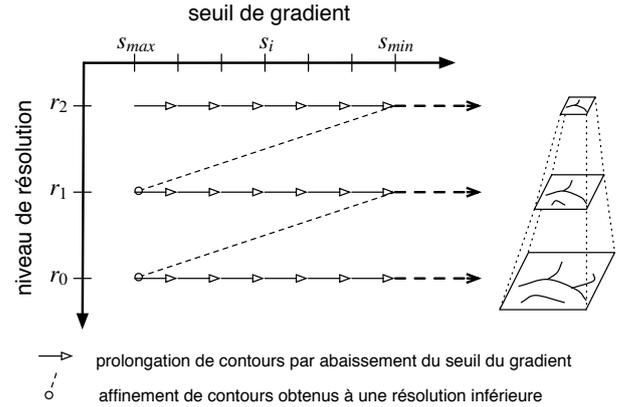


Figure 3 – Schéma d'extraction multirésolution des contours.

étape d'agrégation de points au contour est poursuivi tant que des successeurs sont trouvés. Lorsque l'on ne peut plus prolonger de contour, le seuil du niveau de gradient est abaissé et on réitère le processus.

Si l'on considère un contour à prolonger d'extrémité P_e , la fonction de coût à minimiser :

$$\alpha E_{continuité} + \beta E_{orientation} + \gamma E_{image} \quad (1)$$

fait intervenir trois termes pondérés par des paramètres α , β et γ :

$E_{continuité}$: une contrainte de continuité qui tend à prolonger la chaîne dans la direction donnée par P_e et son prédécesseur ;

$E_{orientation}$: une contrainte lié à la mesure de l'orientation du gradient au point P_e ;

E_{image} : l'opposé de la norme du gradient au point candidat pour privilégier les points de contours forts.

On recherche parmi les trois points 8-connexe qui font varier l'orientation de la chaîne de 0° ou $\pm 45^\circ$ celui qui minimise la fonction de coût. Si ce dernier à un niveau de gradient suffisant, il prolonge le contour d'un élément.

Un contour prolongé peut rencontrer une autre chaîne pour former des coins ou des noeuds. S'il rencontre l'extrémité d'une autre chaîne, on réalise une fusion si un angle droit ne se forme pas entre ces deux structures, sinon on forme un coin. Si le contour rencontre un point interne d'un autre contour, on aura alors formation d'un noeud.

Lorsque l'on a atteint un seuil minimal du gradient, la prolongation s'arrête. On a donc alors extrait un ensemble de contours à une résolution donnée. Ces structures extraites peuvent maintenant être transposées à la résolution supérieure pour être prolongées. À la résolution supérieure, de nouvelles structures et détails peuvent aussi émerger.



(a)



(b)

Figure 4 – Une image de la séquence *Foreman* (a) et les extrema locaux du gradient de norme supérieure à 0,05% de la norme maximale théorique (b). Attention, ici aucun contour n'est formé, les points sont à relier.

3.2 Affinement d'un contour à la résolution directement supérieure

Les éléments 8-connexes d'un contour sont projetés à la résolution juste supérieure. Ces points, qui se retrouvent distants de 2 ou $2\sqrt{2}$ pixels, sont considérés comme l'initialisation d'un contour actif dont l'énergie est modélisée par :

$$\alpha' E_{\text{élasticité}} + \beta' E_{\text{courbure}} + \gamma' E_{\text{image}} \quad (2)$$

où :

$E_{\text{élasticité}}$: représente l'élasticité entre les éléments de contours. Ce terme est minimal quand ces points restent à une distance constante les uns des autres ;

E_{courbure} : représente la courbure du contour, cette énergie est minimale pour une courbure nulle ;

E_{image} : est égale à l'opposé de la norme du gradient afin d'attirer les éléments du contours actif vers les points de forte norme du gradient.

Cette énergie est minimisée par un algorithme glouton [8]. Les points du contour actif sont déplacés à tour de rôle pour minimiser la fonction de coût. Lorsqu'un minimum local est atteint ou qu'un certain nombre d'itérations est dépassé, les éléments du contour sont reliés entre eux par l'algorithme de Bresenham.

Nous attirons l'attention sur le fait qu'un noeud qui se forme à une résolution donnée, ne correspond pas nécessairement à un noeud à la résolution supérieure. La norme du gradient décroît à proximité des noeuds, par conséquent la position des noeuds en est moins précise. Pour cette raison, les extrémité de contours de type noeud ne sont pas projetés à la résolution supérieure. Si deux contours forment un contours à une résolution donnée, ils seront projetés à la résolution supérieure sans être connecté, mais pourront l'être par la suite par prolongation de ces contours.

4 Résultats

La prolongation des contours à pleine résolution ne parvient pas toujours à relier des structures qui appartiennent vraisemblablement au même contour. Dans certains cas une chute locale du gradient ne permet pas un lien entre ces structures, dans d'autres cas un coin se forme à la connexion entre deux structures qui empêche leur fusion. En procédant par exemple à une extraction de contours multirésolution sur deux niveaux de résolution, on permet une meilleure fusion entre les contours (voir figure 5).

Le nombre de contours ainsi détectés reste très important. Cette sur-détection n'est pas souhaitable dans un contexte de codage par bandelettes. Un post-traitement pour éliminer les contours de petite taille ne permettraient pas de supprimer systématiquement des contours proches. Par contre, une extraction de contours initiale à basse résolution, puis affiner ces contours de résolution en résolution sans faire apparaître de nouvelles structures à partir d'un certain niveau, permet d'obtenir des contours se situant à une certaine distance les uns des autres (voir figure 6).

5 Conclusion

Notre approche d'extraction de contours multirésolution permet une description à plusieurs niveaux de résolution de l'image. Dans le contexte d'une application de codage par transformée en bandelettes, il est nécessaire d'extraire des contours présentant une structure suffisamment longue afin de réduire leur coût de codage. L'extraction de contours à basse résolution, suivi d'un affinement des structures extraites de résolution en résolution jusqu'à la pleine résolution permet de répondre à cette contrainte et d'obtenir des contours suffisamment éloignés.

Références

- [1] Emmanuel J. Candès et David L. Donoho. Ridgelets : a key to higher-dimensional intermittency? *Philosophical Transactions Royal Society London A*, (357) :2495–2509, 1999.

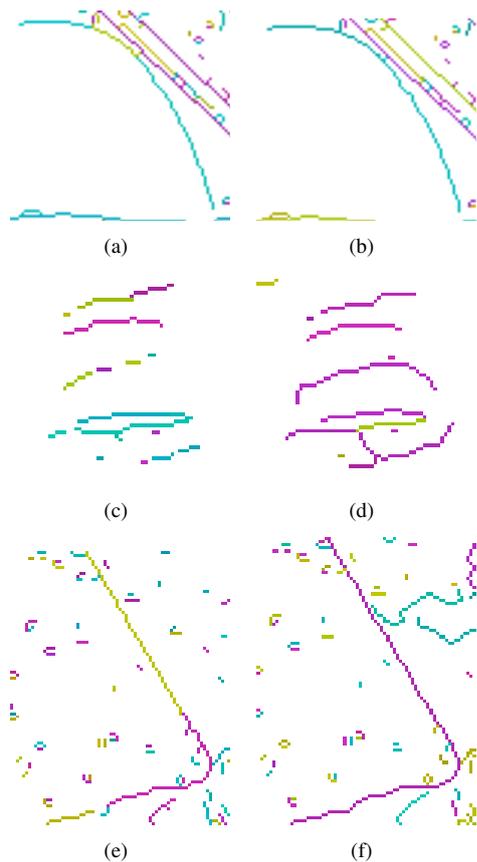


Figure 5 – Des exemples de contours extraits directement à pleine résolution (a, c, e), et avec à deux niveaux de résolution (b, d, f).

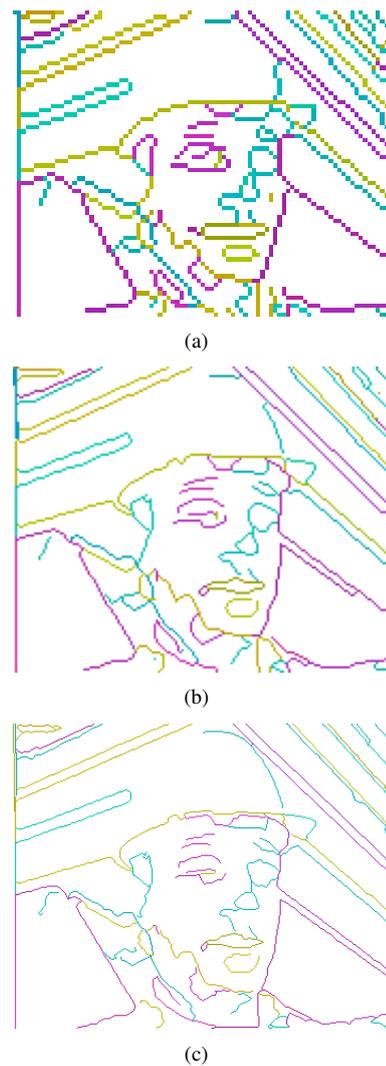


Figure 6 – Extraction de contours sur le deuxième niveau d'une décomposition multirésolution (a) puis affinage des structures extraites à la résolution de niveau un (b), puis à la pleine résolution (c).

- [2] Minh N. Do et Martin Vetterli. The finite ridgelet transform for image representation. *IEEE Transactions On Image Processing*, 12(1), janvier 2003.
- [3] Jean-Luc Starck, Emmanuel J. Candès, et David L. Donoho. The curvelet transform for image denoising. *IEEE Transactions On Image Processing*, 11 :670–684, novembre 2000.
- [4] Minh N. Do et Martin Vetterli. The contourlet transform : An efficient directional multiresolution image representation. *IEEE Transactions On Image Processing*, 14(12) :2091–2106, december 2005.
- [5] Demin Wang, Liang Zhang, et André Vincent. Curved wavelet transform and overlapped extension for image coding. Dans *ICIP 2004*, Singapour, 2004.
- [6] Erwan Le Pennec. *Bandelettes et représentation géométrique des images*. Thèse de doctorat, École Polytechnique, decembre 2002.
- [7] Peter J. Burt et Edward H. Adelson. The laplacian pyramid as a compact image code. Dans *IEEE Transactions On Communications*, volume 31, pages 532–540, 1983.

- [8] Jean-Jacques Rousselle. *Les contours actifs, une méthode de segmentation. Application à l'imagerie médicale*. Thèse de doctorat, Université François Rabelais de Tours, juillet 2003.