

Étude comparative de techniques de lifting orientées sur des contenus synthétiques représentant des structures rectilignes

G. Jeannic

V. Ricordel

D. Barba

IRCCyN - Équipe IVC

Université de Nantes

Polytech’Nantes, rue Christian Pauc, F-44306 NANTES

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

Résumé

Les limites de la transformées en ondelettes pour représenter des contours d’orientation quelconques sont connues. De nouvelles représentations orientées ont été définies à partir de différentes techniques de lifting. Cet article présente une comparaison de trois d’entre elles, et plus précisément de leur capacité à minimiser l’énergie de la sous-bande reconstruite des hautes fréquences pour des contenus synthétiques représentant des structures rectilignes orientées.

Mots clefs

représentation d’image, transformée en ondelettes, lifting, représentation directionnelle.

1 Introduction

La transformée en ondelettes 1D permet de représenter efficacement des singularités ponctuelles. Étendue à deux dimensions par le produit tensoriel de deux transformées en ondelettes 1D, elle permet de représenter les singularités appartenant à des contours horizontaux, verticaux ou diagonaux. De plus un contour idéal orienté selon l’horizontale (respectivement la verticale) n’est pas détecté par le filtre d’analyse correspondant aux singularités verticales (respectivement horizontales), permettant ainsi de concentrer l’information dans une seule sous-bande des hautes fréquences. Il s’agit cependant d’une limitation de la transformée en ondelettes. En effet si elle permet de représenter efficacement les contours horizontaux et verticaux, tout contour orienté différemment se retrouve représenté par des coefficients dans chacune des sous-bandes des hautes fréquences.

La transformée en ondelettes 1D peut être implantée par des techniques de *lifting* [1]. Ces techniques se décomposent en trois parties. Une décomposition polyphase qui sépare en deux ensembles les échantillons notés pairs et les échantillons notés impairs. Ensuite une succession d’étapes de prédiction et de mise à jour qui modifient l’un de ces ensembles par rapport à l’autre, et vice versa. Et enfin une étape de mise à l’échelle de ces éléments. La transformée

en ondelettes 2D standard est alors obtenue en appliquant successivement une transformée en ondelettes 1D implantée par lifting selon les lignes puis selon les colonnes (ou inversement).

La flexibilité de ces techniques de lifting a permis de développer de nouvelles transformées en ondelettes 2D orientées qui s’adaptent au contenu de l’image à représenter. Contrairement à la transformée en ondelettes 2D standard qui opère un filtrage selon deux directions fixées : l’horizontale et la verticale, ces transformées orientées procède selon deux directions adaptatives. Pour cela, le principe général de ces techniques réside dans le choix des échantillons, i.e. leurs coordonnées, d’un ensemble qui sont utilisés pour prédire et mettre à jour la valeur des échantillons de l’autre ensemble.

Ce choix est donc entre autre dépendant de la décomposition polyphase utilisée. Dans [2] par exemple, la décomposition utilisée est quinconce. Nous nous restreindrons ici au décomposition polyphase régulière.

Différentes implantations de techniques de lifting orientées ont été proposées. Certaines introduisent de l’interpolation dans les étapes de prédiction et de mise à jour : linéaire [3, 4] ou spline [5]. D’autres utilisent des échantillons situés sur des droites discrètes au lieu des échantillons simplement voisin de gauche et de droite [4, 6, 2].

Dans un contexte de codage d’image, les transformées sont généralement comparées à travers une approche débit-distorsion. Afin de produire un train binaire, l’image transformée doit donc être quantifiée, et les indices de quantification doivent être codés. Des images reconstruites à différents débits sont comparés avec une métrique de distorsion ou de qualité. Pour pouvoir comparer strictement deux transformées, les lois de quantification et les règles de codage doivent être identiques. Toutefois une telle loi ou une telle règle adaptée à la transformée en ondelettes 2D n’est pas nécessairement la plus adaptée à une autre transformée. L’objectif de cet article sera de comparer différentes implantations de techniques de lifting orientées d’un point de vue statistique uniquement sans l’influence de la quantification et du codage.

Dans une première section, trois techniques de lifting orientées, qui vont être comparées, seront détaillées. Ces transformées orientées seront appliquées sur divers contenus synthétiques représentant des structures rectilignes selon un protocole décrit dans une deuxième section. Finalement dans la troisième section, la capacité de ces transformées à minimiser l'énergie de la sous-bande reconstruite des hautes fréquences sera analysée.

2 Implantations de lifting orientées

Pour implanter une technique de lifting orienté, les étapes de prédiction et de mise à jour peuvent être obtenues en considérant des échantillons qui ne sont pas sur la même ligne pour un filtrage plutôt horizontal ou sur la même colonne pour un filtrage plutôt vertical.

Pour implanter une telle transformée, une solution consiste à prédire les éléments des colonnes impairs (respectivement lignes) à partir d'une combinaison linéaire des éléments des colonnes paires (respectivement lignes) et de mettre ensuite à jour les éléments pairs à partir des erreurs de prédiction précédemment calculées. Comme la transformée en ondelettes 2D standard, la reconstruction reste exacte et la décimation critique [4].

À partir de cette définition, différentes implantations pratiques sont possibles. Parmi celles-ci trois techniques de lifting orienté sont étudiées.

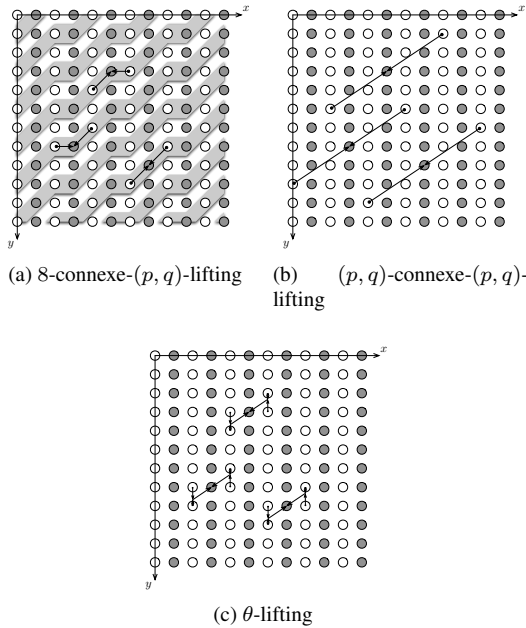


Figure 1 – Les trois techniques de lifting orienté étudiées.

2.1 8-connexe-(p, q)-lifting

Dans cette première variante [4], illustrée sur la figure 1-a, l'image ou une région de l'image est partitionnée en droites discrètes 8-connexes dont l'une est définie par :

$$0 \leq -q \cdot (x - x_0) + p \cdot (y - y_0) < \max(|p|, |q|) \quad (1)$$

avec (p, q) un vecteur représentant l'orientation de cette droite, et (x_0, y_0) les coordonnées d'un pixel de cette droite. Chaque pixel est associé à une droite discrète en calculant l'indice $\left\lfloor \frac{-q \cdot (x - x_0) + p \cdot (y - y_0)}{\max(|p|, |q|)} \right\rfloor$.

Cette implantation est équivalente à une transformée en ondelettes standard appliquée le long de ces droites discrètes.

2.2 (p, q)-connexe-(p, q)-lifting scheme

L'orientation de filtrage est là aussi définie par un vecteur (p, q) , mais dans ce cas les droites discrètes sont (p, q) -connexes au lieu d'être 8-connexes (voir figure 1-b). Cette approche est présentée dans [7, 6]. Cette technique de lifting est équivalente à une transformée en ondelettes appliquée selon une lattice correspondante.

2.3 θ -lifting scheme

Au contraire des deux implantations précédentes, pour le θ -lifting, illustré sur la figure 1-c, les valeurs utilisées dans la prédiction et la mise à jour peuvent être interpolées à partir d'échantillons de parité appropriée. La reconstruction reste exacte tant que l'interpolation est une opération linéaire. Dans la suite de cet article, l'interpolation linéaire sera considérée.

Cette technique de lifting ne correspond pas à une transformée en ondelettes d'éléments alignés selon une orientation définie par l'angle θ dont les valeurs de ces éléments sont obtenues par interpolation linéaire. En effet les interpolations des étapes successives de prédiction et de mise à jour induisent un effet ricochet, qui correspond à étaler les coefficients des filtres passe-bas et passe-haut équivalents à la technique de lifting, comme illustré sur la figure 2. Une conséquence notable est que l'énergie de la réponse impulsionnelle de ces deux filtres varie en fonction de l'orientation de filtrage.

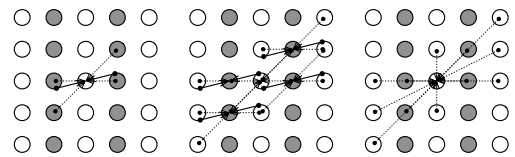


Figure 2 – Illustration de l'effet ricochet dans l'implantation du θ -lifting basé sur les filtres d'ondelettes 5/3. Sur la figure (a), l'échantillon central (pair) est mis à jour avec des valeurs interpolées à partir de quatre échantillons impairs. Ces échantillons impairs correspondent alors à des erreurs de prédiction calculés à partir de valeurs interpolées d'échantillons pairs (b). La figure (c) représente finalement à partir de quels échantillons l'échantillon pair central est réellement filtré.

3 Protocole d'étude

L'objectif de cet article est de comparer la capacité de ces trois implantations de lifting orientées à minimiser l'énergie de la sous-bande des hautes fréquences le long de l'orientation de filtrage. Pour cela, chaque technique de lifting orienté est appliquée selon toutes les orientations d'un ensemble restreint. La meilleure orientation pourrait être choisie comme celle minimisant l'énergie de la sous-bande des hautes fréquences. Toutefois, pour le θ -lifting, comme l'énergie de la réponse impulsionnelle des filtres d'ondelettes équivalents varie en fonction de l'orientation de filtrage considérée, la dynamique du signal de la sous-bande des hautes fréquences selon l'orientation de filtrage varie elle aussi. Afin de pallier à ce problème de normalisation, l'image est reconstruite à partir uniquement de la sous-bande des hautes fréquences, et il s'agit donc de minimiser l'énergie de cette image reconstruite pour la comparaison des différentes techniques de lifting.

Pour les techniques de (p, q) -lifting, l'ensemble des orientations de filtrage est défini par la série de Farey d'ordre 3 [8] (i.e. $\{(1, 0); (3, 1); (2, 1); (3, 2); (1, 1)\}$ pour les orientations entre 0° et 45°). Pour la technique de θ -lifting, cet ensemble est construit par un échantillonnage des orientations. Cet échantillonnage peut par exemple être uniforme ou basé sur une interpolation sub-pixelique. Un total de 16 orientations est ici choisi. En combinant les deux méthodes d'échantillonnage, les orientations qui seront testées entre 0° et 45° sont donc approximativement : $\{0^\circ; 11, 25^\circ; 14, 04^\circ; 22, 5^\circ; 26, 57^\circ; 33, 75^\circ; 36, 87^\circ; 45^\circ\}$.

La capacité de ces différentes techniques de lifting à minimiser le terme d'énergie considéré est étudiée pour des contenus synthétiques représentant des structures rectilignes, favorables pour ce type de transformées. Des images de test 64×64 sont générées pour trois types de contenus :

1. la droite discrète d'orientation (p_{ech}, q_{ech}) définie par l'équation (1) où $(x_0, y_0) = (\frac{N-1}{2}, \frac{N-1}{2})$ est la frontière de deux régions uniformes. Les pixels de coordonnées (x, y) tels que $\frac{-q_{ech} \cdot (x-x_0) + p_{ech} \cdot (y-y_0)}{\max(|p_{ech}|, |q_{ech}|)} \geq 0$ valent 64, les autres 192 ;
2. un échelon 2D idéal, orienté selon θ_{ech} , dont la valeur passe de 64 à 192 est filtré par un passe-bas dont la réponse impulsionnelle est unitaire sur le support d'un pixel, puis idéalement échantillonné ;
3. un signal sinusoïdal 2D, orienté selon θ_{ech} , de période $T = 5$, et d'amplitude crête à crête 128, est idéalement échantillonné.

Pour le premier contenu, un ensemble d'images de test est construit avec (p_{ech}, q_{ech}) appartenant à la série de Farey d'ordre 6, i.e. $\{(1, 0); (6, 1); (5, 1); (4, 1); (3, 1); (5, 2); (2, 1); (5, 3); (3, 2); (4, 3); (5, 4); (6, 5); (1, 1)\}$. Pour les deux autres, θ_{ech} varie de 0° à 45° par pas de 5° . La

distribution de ces orientations est choisi volontairement plus fine que celle des orientation de filtrage pour pouvoir étudier les performances de chaque transformée pour des orientations de contenu qui ne correspondent pas exactement à une orientation de filtrage.

4 Résultats

Les résultats présentés ont été calculés avec les filtres d'ondelettes 5/3.

Contenu 1 Le premier contenu, basé sur des droites discrètes 8-connexes, a une définition adaptée à celle de la technique du 8-connexe- (p, q) -lifting. En pratique, une différence de phase peut exister entre la définition des droites discrètes du contenu et les droites discrètes de filtrage.

Ici, volontairement, la droite d'indice 0 définissant le contenu passe par le centre de l'image de test, alors que celle définissant le filtrage passe par le point de coordonnées $(0, 0)$. Ainsi même pour une orientation de filtrage parfaitement adaptée à l'orientation du contenu, la capacité à minimiser le terme d'énergie étudié est extrêmement dépendante du décalage entre les droites discrètes du contenu et celles du filtrage.

Dans les faits on observe, sur la figure 4-a, pour une même orientation du contenu et du filtrage (i.e même (p, q)), une énergie nulle pour une orientation de $18, 43^\circ$ ($(p, q) = (3, 1)$), une énergie non nulle pour une orientation de $26, 57^\circ$ ($(p, q) = (2, 1)$) mais inférieure aux énergies obtenues avec d'autres orientations de filtrage, et une énergie non nulle et non minimale, comparée aux autres filtrages, pour $33, 69^\circ$ ($(p, q) = (3, 2)$).

Ce dernier cas est illustré sur la figure 3 par une image de test de dimension 16×16 afin d'expliquer le fait qu'un filtrage selon une orientation qui semble être la « bonne » peut être moins performant qu'un filtrage selon 0° . Pour un filtrage horizontal de l'image de test considérée, un élément impair proche d'un contour de valeur 64 ou 192 est prédit au pire par un pixel de valeur 64 et un pixel de valeur 192 produisant ainsi une erreur de prédiction maximale de $\frac{1}{\sqrt{2}} (64 - \frac{64+192}{2}) = -\frac{64}{\sqrt{2}}$ ou $\frac{1}{\sqrt{2}} (192 - \frac{64+192}{2}) = \frac{64}{\sqrt{2}}$ (voir figure 3-a). Pour un filtrage selon l'orientation définie par $(p, q) = (5, 2)$ qui correspond à l'orientation définissant l'échelon de l'image de test, certains éléments de contour de valeur 64 (respectivement 192) sont prédits par deux pixels de valeur 192 (respectivement 64) induisant une erreur de prédiction pouvant atteindre $\frac{1}{\sqrt{2}} (64 - \frac{192+192}{2}) = -\frac{128}{\sqrt{2}}$ (respectivement $\frac{1}{\sqrt{2}} (192 - \frac{64+64}{2}) = \frac{128}{\sqrt{2}}$) soit le double (voir figure 3-b). De plus, non seulement l'erreur de prédiction peut être plus élevée pour une orientation de filtrage plus proche de l'orientation de l'échelon que 0° , mais comme on l'observe sur les figures 3-a et 3-b, le nombre de pixels « mal prédits » est plus élevé pour l'orientation de filtrage $(5, 2)$ que pour $(1, 0)$.

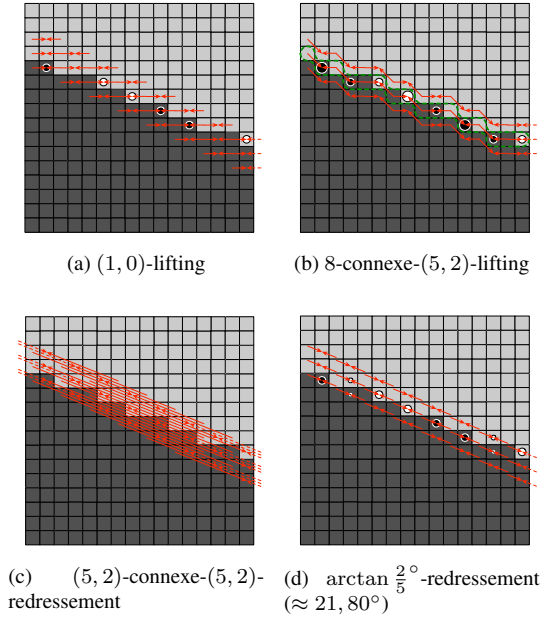


FIGURE 3 – Illustration de l'étape de prédiction d'une image de test correspondant au contenu 1 où $(p_{ech}, q_{ech}) = (5, 2)$. Les disques sur les échantillons des colonnes impaires correspondent à la valeur de l'erreur de prédiction pour les ondelettes 5/3. La couleur représente le signe de l'erreur (blanc-positif, noir-négatif), l'aire du disque la valeur absolue.

La deuxième technique, le (p, q) -connexe- (p, q) -redressement, profite pour ce contenu des avantages de la précédente technique sans ses inconvénients (exemple sur la figure 3-c). On observe donc bien sur la figure 4-b une énergie nulle lorsque l'orientation de filtrage coïncide avec l'orientation du contenu. De plus les résultats restent stables pour des orientations du contenu seulement voisins de l'orientation de filtrage.

Les résultats pour la dernière technique, le θ -lifting, sont donnés figure 4-c. On observe que quelle que soit l'orientation du contenu, n'importe quelle technique de lifting nécessitant une interpolation ne va pas réduire l'énergie plus qu'un filtrage à 0° , 45° , 90° ou 135° . L'interpolation est ici inutile car l'image de test est binaire (64 ou 192) et par conséquent aucune interpolation à partir de ces valeurs ne peut produire une erreur de prédiction plus faible que ne le fait une prédiction qui ne nécessite pas d'interpolation. En l'occurrence, comme on peut l'observer sur les figures 3-a et 3-d, toutes les erreurs de prédiction non nulles pour un filtrage à 0° se retrouvent aussi pour un filtrage à environ $21, 80^\circ$. Par contre d'autres erreurs apparaissent pour cette orientation qui n'existait pas pour un filtrage strictement horizontal.

Contenu 2 Le deuxième contenu rectiligne testé a été conçu pour représenter de façon très simplifiée un échelon analogique orienté qui serait acquis par une matrice de capteurs idéale. La valeur d'un pixel correspond alors à la valeur de l'échelon intégrée sur la surface de ce pixel. Les

résultats présentés sur la figure 5 sont similaires à ceux de la figure précédente. L'effet lisseur introduit dans la modélisation de l'échelon n'améliore pas les performances de la technique de θ -lifting.

Contenu 3 Les résultats présentés sur la figure 6 pour le troisième contenu représentant un signal sinusoïdal 2D orienté montre que l'interpolation de la technique de θ -lifting ne permet pas d'obtenir un terme d'énergie plus faible qu'avec un filtrage horizontal ou diagonal. On constate que, pour une sinusoïde 2D dont l'orientation est plus proche de l'horizontale que de la diagonale par exemple, que la moyenne des voisins de gauche et de droite d'une valeur à prédire, qui sont la plupart du temps plus faible pour l'une et plus élevée pour l'autre, offre une meilleure prédiction que la moyenne de deux valeurs interpolées selon la direction de la sinusoïde.

5 Conclusion

Trois implantations de techniques de lifting orienté de l'état de l'art ont été présentées. Ces techniques ont été appliquées sur différents contenus synthétiques représentant des structures rectilignes orientées afin de comparer leur performance à minimiser l'énergie de l'image reconstruite en ne conservant que les hautes fréquences selon la direction de filtrage.

Pour les contenus considérés, la technique de lifting faisant intervenir de l'interpolation linéaire ne permet pas de minimiser le terme d'énergie considéré pour des orientations de filtrage autres que l'horizontale, la verticale et les diagonales. Pour la technique de lifting équivalente à une transformée en ondelettes le long de droite discrètes 8-connexes, la conclusion est la même à part lorsqu'exceptionnellement le contenu est parfaitement en phase avec l'orientation de filtrage. Seule la technique de (p, q) -connexe- (p, q) -lifting arrive à tirer parti d'orientations de filtrage qui ne sont pas l'horizontale, la verticale ou la diagonale.

Toutefois les performances de cette dernière implantation à minimiser l'énergie de la sous-bande reconstruite des hautes fréquences selon la direction de filtrage diminue fortement avec la diminution du contraste des structures rectilignes considérées. De plus cette étude sur des contenus synthétiques ne prend pas en compte la notion de cohérence spatiale. Dans des images naturelles, la corrélation entre deux valeurs de pixels diminue avec la distance entre ces pixels, ce qui est préjudiciable pour le (p, q) -lifting qui utilise des échantillons pour la prédiction et la mise à jour situés à une distance $\sqrt{p^2 + q^2}$ de l'échantillon à prédire ou mettre à jour.

References

- [1] Ingrid Daubechies et Wim Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):247–269, 1998.
- [2] Vivien Chappelier et Christine Guillemot. Oriented wavelet transform for image compression and de-

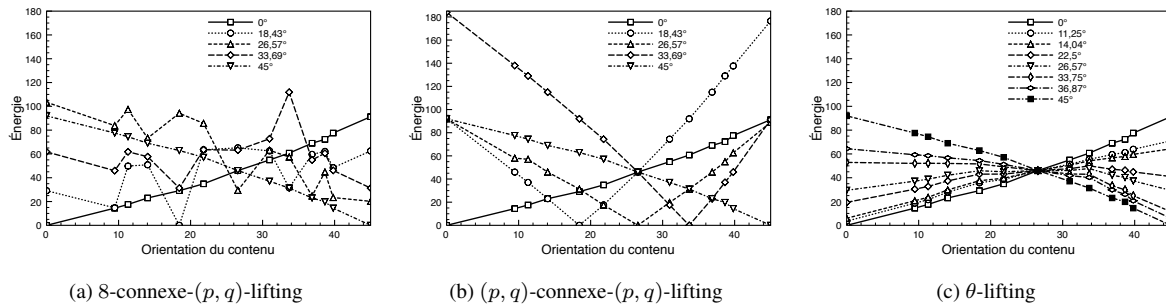


Figure 4 – Énergie des images représentant le contenu 1, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de lifting.

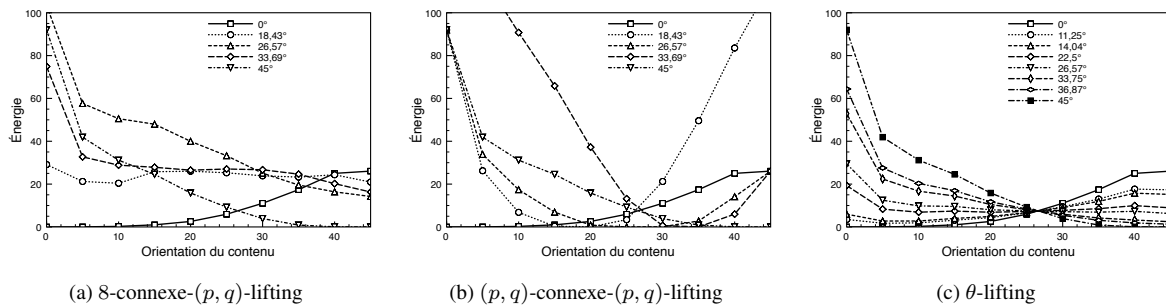


Figure 5 – Énergie des images représentant le contenu 2, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de lifting.

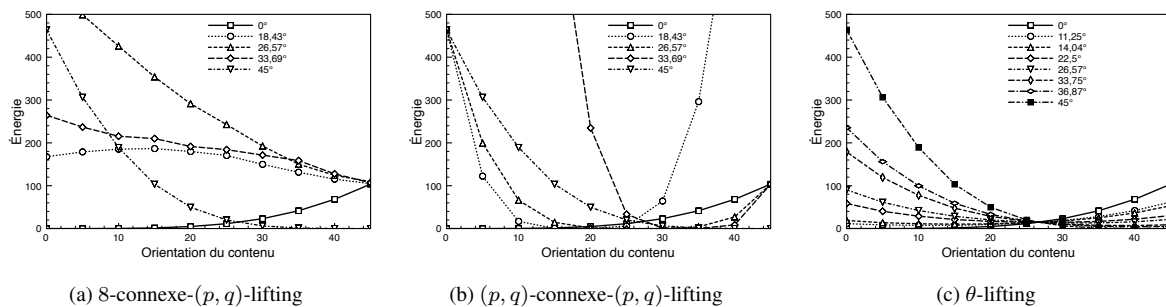


Figure 6 – Énergie des images représentant le contenu 3, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de lifting.

noising. *IEEE Transactions On Image Processing*, 15(10):2892–2903, octobre 2006.

[3] Demin Wang, Liang Zhang, et Filippo Sperenza. Curved wavelet transform for image coding. *IEEE Transactions On Image Processing*, 15(8):2413–2421, août 2006.

[4] Guillaume Jeannic, Vincent Ricordel, et Dominique Barba. The edge driven oriented wavelet transform: An anisotropic multidirectional representation with oriented lifting scheme. Dans *Visual Communications and Image Processing*, San Jose, USA, janvier 2007.

[5] Erwan Le Pennec et Stéphane Mallat. Sparse geometric image representations with bandelets. *IEEE Transactions On Image Processing*, 14(4):423–438, avril 2005.

[6] Chuo-Ling Chang et Bernd Girod. Direction-adaptive discrete wavelet transform for image compression. *IEEE Transactions on Image Processing*, 16(5), mai 2007.

[7] Vladan Velisavljević, Baltasar Beferull-Lozano, Martin Vetterli, et Pier Luigi Dragotti. Directionlets: Anisotropic Multidirectional representation with separable filtering. *IEEE Transactions on Image Processing*, 15(7):1916–1933, 2006.

[8] G.H. Hardy et E.M. Wright. *An Introduction to the Theory of Numbers*. Oxford Science Publications, 1979.