

**Université du Québec en Outaouais
Département d'Informatique et d'Ingénierie**

**Extraction de primitives d'images pour la recherche par le
contenu : la transformée en contourlet**

Rapport technique
(Partie I)

Par

Nadia Baaziz

Année 2004

Table des Matières

- Table des matières
- I- Introduction
- II- Invariance des primitives aux transformations
- III- Approches d'extraction des primitives
- IV- La transformée en Contourlet d'images
 - 1- La transformée en Contourlet redondante
 - 2- Mise en œuvre de la pyramide redondante
 - 3- Mise en œuvre de la Contourlet redondante
 - 4- Génération des filtres directionnels
 - 5- La Contourlet diagonale redondante
- V- Conclusion
- VI- Références bibliographiques

I- Introduction

L'indexation d'informations visuelles pour la recherche par le contenu est un thème qui connaît, de nos jours, un intérêt grandissant dans les communautés de vision par ordinateur, de traitement d'images et vidéo, et de gestion de bases de données. Cet intérêt est notamment motivé par les avancées récentes en performance et en standardisation qu'ont connues les techniques de compression de données, ainsi que par l'utilisation généralisée de la technologie de l'Internet et les réseaux à large bande.

En effet, on assiste de nos jours à une explosion d'applications multimédia qui intègrent de manière croissante des quantités de données textes, audio, graphiques, images et vidéo. Dans ce contexte de volume constamment croissant de ces données, la technologie de l'information vise à en faciliter la représentation, la recherche, l'accès et à améliorer ainsi l'utilisabilité. Parmi ces applications figurent les bibliothèques digitales, la vidéo sur demande, la télémédecine, les systèmes d'information géographiques, la télésurveillance.

L'accès à de gros volumes de données visuelles comprend deux étapes essentielles:

a- le stockage et l'indexation: les images numériques (fixes ou en séquence) sont stockées dans un format de représentation pouvant être compressé ou non compressé. Les données sont traitées en vue de l'identification et de l'extraction de traits caractéristiques spécifiques à l'information contenue. Ces primitives extraites sont organisées et structurées pour former un index d'images. Il contient l'ensemble des primitives et des attributs qui décrivent la sémantique ou le contenu de chaque image. Son rôle est d'aider à l'efficacité du processus de recherche et d'accès aux données originales. Il est important de noter que l'indexation ne vise pas la description de l'image toute entière. L'essentiel est de pouvoir identifier des primitives d'images sur lesquelles on peut se baser pour comparer et déclarer qu'il y a ressemblance entre images.

b- la recherche et l'accès: l'utilisateur soumet une requête constituée d'un ou plusieurs attributs de primitives qui l'intéressent. Une recherche, utilisant une mise en correspondance et une mesure de similarité, est opérée sur l'index pour trouver des objets (indices) similaires aux objets recherchés. Les images associées à ces indices sont alors identifiées et présentées par ordre de degré de pertinence.

L'indexation d'images est donc cruciale pour l'efficacité de l'accès aux informations visuelles. L'approche manuelle est la plus efficace mais elle est aussi la plus fastidieuse et coûteuse. Les techniques d'indexation automatique de données textuelles sont plus anciennes et bien avancées, elles peuvent partiellement aider à cette tâche, mais leur utilisation complète dans ce contexte serait restrictive et inadaptée vue la richesse supérieure et variée des données images "on dit bien : une image vaut mille mots". En effet, des descripteurs textuels peuvent être associés à une image pour décrire le titre, l'auteur, la date, et peut être un résumé du contenu global, par contre la recherche de ressemblances pertinentes entre des images en se basant uniquement sur des descriptions textuelles pourrait laisser passer beaucoup de traits caractéristiques reliés à la structure du

signal image. C'est pour cette raison qu'on a recours au développement d'approches automatiques de recherche par contenu, qui se basent sur des techniques d'analyse et de traitement des images, pour extraire des descripteurs pertinents du contenu spatial de l'information visuelle.

II- Invariance des primitives aux transformation

L'extraction automatique de primitives pertinentes d'images est à la base de la recherche par le contenu. Ce processus vise à mettre en évidence et à améliorer les aspects de l'image qui sont pertinents aux requêtes des usagers. Les primitives visuelles sont en général de bas-niveau comme les couleurs, les textures (herbe, nuage, eau, tissu...), les régions d'objets ou les contours d'objets. Elles peuvent aussi évoluer en entités sémantiques de haut niveau comme les faces humaines ou les empreintes digitales.

Une des caractéristiques désirées des primitives bas-niveau est l'invariance à des transformations de base souvent dues à des distorsions accidentelles reliées aux conditions des capteurs d'acquisition (translation, rotation, miroir, bruits additifs...) [Sme 00]. Une image peut exister sous diverses variantes. Ces dernières peuvent être représentées comme le résultat d'une transformation d'une image originale par rotation, translation faible, zoom léger ou encore effet de miroir. À l'œil nu, l'utilisateur n'aura aucune difficulté, dans ce cas là, à détecter la ressemblance entre les images transformées et l'image originale. Mais les méthodes de comparaison automatiques, qui se basent sur des mesures, peuvent faillir dans cette tâche si on ne prend pas soin de choisir d'extraire des primitives qui sont invariantes à ces transformations. Par exemple, la couleur d'un objet ne change pas si l'objet subit une translation ou une rotation. Par contre, les conditions de luminosité peuvent changer sa couleur. On dit qu'une propriété est invariante à un ensemble de transformations donné, si et seulement si, cette propriété reste inchangée malgré l'existence de conditions indésirables exprimées par ces transformations. L'avantage d'une primitive invariante est son pouvoir à discriminer un objet indépendamment de la façon dont il est observé et de l'endroit où il est observé, moyennant une certaine perte d'information sur le contenu. Ainsi, les primitives invariantes sont davantage porteuses d'informations spécifiques à l'objet que le sont d'autres primitives. Cependant, le degré d'invariance croît au détriment du pouvoir discriminatoire. En effet, si une primitive est invariante sous un grand nombre de transformations, le nombre d'images ayant subi ces transformations ne peut être que plus grand et la recherche s'opère donc dans un espace plus étendu. Pour rendre la recherche plus sélective, on est amené à augmenter le nombre de primitives et donc à effectuer une recherche plus complexe. Pour conserver ou assurer le pouvoir discriminatoire, il serait souhaitable de pouvoir déterminer un minimum de primitives invariantes appropriées à un ensemble de transformations de dimension raisonnable (conditions indésirables), et de spécifier la requête de recherche avec un minimum d'invariants.

III- Approches d'extraction des primitives

L'extraction de primitives bas-niveau de l'image peut se faire aussi bien dans le domaine spatial que dans un domaine transformé. Les méthodes opérant dans le domaine fréquentiel ou bien dans le domaine spatio-fréquentiel ont reçu beaucoup d'attention. Traditionnellement, la représentation multi-échelles des données visuelles (images, vidéo) dans le domaine du codage des signaux tient compte de deux facteurs très importants, à savoir, l'échantillonnage critique qui garantit la compacité en un nombre minimal d'échantillons de représentation, ainsi que la reconstruction parfaite synonyme de réversibilité et de conservation totale de l'information contenue. La transformée en ondelettes largement exploitée dans diverses applications, et même récemment intégrée dans le standard de codage d'images JPEG2000, constitue un exemple typique pour ces représentations.

Par ailleurs, les communautés de chercheurs en robotique et en vision par ordinateur ont plutôt tendance à privilégier les représentations d'images exhibant une sensibilité directionnelle au spectre fréquentiel. Cette caractéristique tant désirée trouve sa justification dans les aspects orientationnels du système de perception visuelle humain. Ces représentations directionnelles sont souvent exploitées pour détecter, extraire, et restaurer des primitives pertinentes de l'image. Les transformées de Hough et de Radon en sont des exemples typiques.

En ce qui a trait à la détection de primitives de type contour d'objets, la catégorie des méthodes purement fréquentielles, parmi tant d'autres méthodes, a été intensivement appliquée. Les contours sont extraits et représentés par des coefficients de Fourier. Diverses variantes visent à améliorer la robustesse au bruit, la qualité de reconstruction et la réduction de la complexité d'implémentation. Pour tenir compte des avantages de l'orientation fréquentielle, de la localisation spatio-fréquentielle et de la multirésolution, les filtres de Gabor ainsi que les transformées en Ondelettes ont été utilisés [Sme 00][Rui 99].

Cependant, malgré le succès des Ondelettes dans divers domaines de traitement d'images tel que le codage [Baa 92], des faiblesses ont été constatées quant à leur utilisation pour la détection et la représentation des contours d'objets. En effet, l'orthogonalité de la décomposition et l'échantillonnage critique font apparaître des effets d'aliasing visibles autour des contours au moindre traitement non linéaire. De plus, le nombre d'orientations est limité et fixe, et les contours sont redondants d'un niveau de résolution à un autre, ce qui requiert un grand nombre de coefficients d'Ondelettes pour les représenter.

Les transformées en Ondelettes et les autres décompositions multi résolutions classiques semblent former une catégorie restreinte et limitée des possibilités de représentations multi échelles de signaux multidimensionnels. Des travaux récents montrent qu'il est possible de définir des cadres théoriques de représentations multi échelles plus vastes donnant naissance à de nouvelles transformées plus intéressantes et plus adaptées à l'extraction de structures géométriques lisses et continues telles que les contours d'objets [Can 02]. C'est le cas de la nouvelle famille des transformées en Ridgelet, en Wedgelet, en Curvelet et en Contourlet [Sta 02][Do 02]. Ce sont des décompositions multi échelles, qui opèrent selon une multitude d'orientations fréquentielles et qui offrent un bon compromis entre la représentation éparse (ou compacte) de traits caractéristiques et la

qualité perceptuelle de l'image reconstruite. En particulier, la transformée en Contourlet est une version discrète, donc adaptée aux images numériques, qui se base sur l'utilisation de bancs de filtres pyramidaux directionnels. L'image résultant de cette transformation est un ensemble de segments de contours lisses et localisés sur lesquels peuvent s'appliquer des traitements additionnels pour l'extraction de primitives contours d'objets [Do 01] [Kim 95] [Do 02].

IV- La transformée en Contourlet d'images

La transformée en Contourlet telle qu'elle a été introduite par Minh Do et Martin Vetterli [Do 01] [Do 02], est un nouveau schéma de décomposition d'images, qui fournit une représentation éparse des données aussi bien aux résolutions spatiales que fréquentielles. Une décomposition d'image en Contourlet est construite en combinant successivement deux étages de décomposition distincts : une décomposition multi échelle suivie d'une décomposition directionnelle. Le premier étage utilise une pyramide Laplacienne pour transformer l'image en une suite de niveaux LP passe-bande et un niveau passe-bas (approximation en basse fréquence de l'image). Le deuxième étage applique de manière appropriée des filtres bidimensionnels de type quinconce et un échantillonnage critique pour décomposer chaque niveau LP passe-bande en un nombre de bandes directionnelles, capturant ainsi des informations directionnelles ou orientées. Finalement, l'image se trouve représentée par un ensemble de sous bandes multi échelles et orientées, comme cela est montré dans les mises en Œuvres et les expérimentations qui vont suivre.

La transformée en Contourlet est à reconstruction parfaite. Son degré de redondance est relativement faible puisque le nombre total de coefficients Contourlet obtenu à l'issue d'une décomposition approche les $4/3$ du nombre d'échantillons de l'image originale. Ceci étant dû uniquement au surplus d'échantillon inhérent à la pyramide Laplacienne, puisque le reste de la décomposition est effectué avec échantillonnage critique. Comparée à la transformée en Ondelette discrète, la transformée en Contourlet avec sa propriété de sélectivité directionnelle (ou orientée) conduit à des améliorations et à de nouveaux potentiels pour les applications de traitement d'image. En effet, les contours fins sont mieux représentés puisque des expérimentations ont déjà clairement montré que les contours lisses sont représentés de manière efficace par quelques coefficients Contourlet localisés dans la bande à orientation appropriée [Do 01][Do 02].

IV-1 La transformée en Contourlet redondante

Habituellement, les transformées d'images qui sont sous échantillonnées à cadence critique sont adoptées, notamment pour des applications de codage et de compression de données. Cependant, pour l'analyse d'image et l'extraction de primitives, les transformées redondantes sont largement préférées même si un surplus de coefficients transformées et, par conséquent, plus de traitement sont engendrés. En général, les traitements hiérarchiques qui requièrent une coopération ou une mise en correspondance entre différents niveaux de résolution sont rendus plus faciles lorsque tous ces niveaux de résolution présentent la même taille (en nombre d'échantillons). C'est la raison pour

laquelle nous introduisons, dans ce qui suit la transformée en contourlet redondante ayant des sous bandes directionnelles toutes de même taille.

La redondance de la transformée en contourlet est atteinte en éliminant toute opération de sous échantillonnage dans le schéma de construction de la pyramide Laplacienne. En utilisant L filtres passe-bas avec les caractéristiques fréquentielles appropriées, L approximations passe-bas de l'image originale sont créées. La différence entre chaque approximation et sa version passe-bas subséquente est une image passe-bande. Le résultat final est une pyramide Laplacienne redondante avec $L+1$ niveaux de tailles égales; une image d'approximation et L images de type passe-bande (voir le diagramme de

décomposition de la Figure 1). Puisque chaque niveau de résolution a la taille de l'image originale et puisque la pyramide a $L+1$ niveaux, on en déduit que le facteur de redondance est égal à L . En appliquant une même décomposition directionnelle à D directions (avec échantillonnage critique) sur chaque niveau Laplacien redondant, on obtient une décomposition en contourlet redondante ayant LD sous bandes directionnelles toutes de même taille, en plus de l'approximation passe-bas de l'image ayant la taille de l'image originale. Chaque sous bande Contourlet est représentée par une sous image $C_{l,d}(M,N)$ où l est son niveau de résolution et d sa direction fréquentielle (ou orientation fréquentielle), et MN son nombre de coefficients Contourlet.

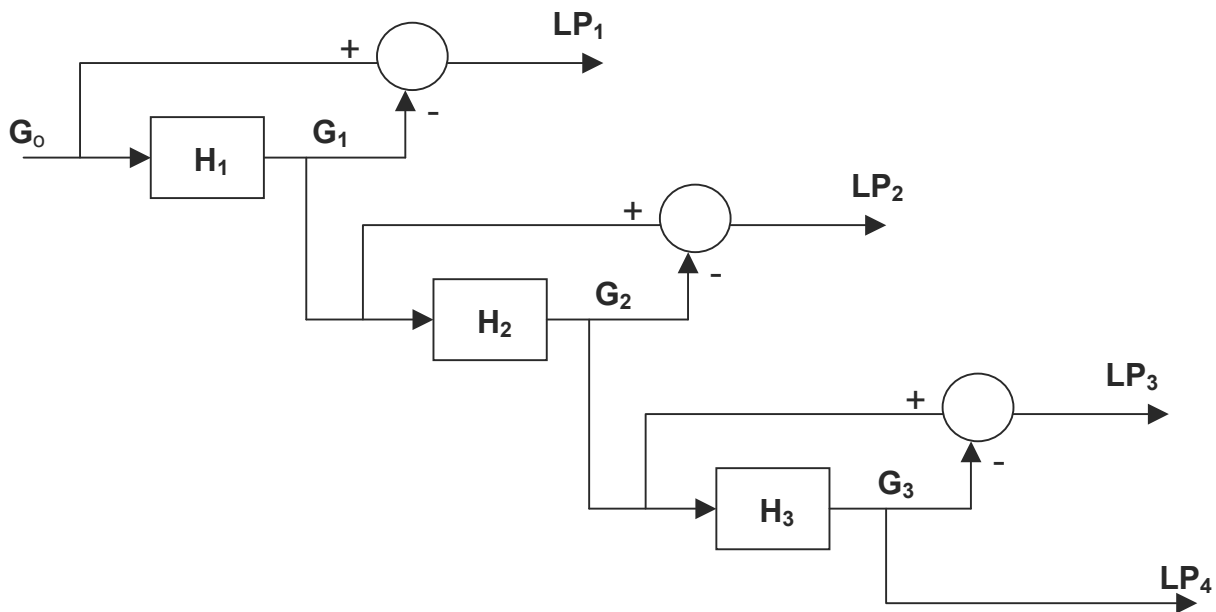


Figure 1 : Diagramme de décomposition en pyramide Laplacienne redondante à 4 niveaux ($L=3$). G_0 est supposée être l'image originale à décomposer. LP_4 est l'approximation passe-bas de l'image. Les filtrages passe-bas H_i sont bidimensionnels séparables, appliqués horizontalement et verticalement sur l'image.

IV-2 Mise en œuvre de la pyramide redondante

Une pyramide Laplacienne redondante à 4 niveaux a été générée ($L=3$). Ce schéma de décomposition a requis l'utilisation de 3 filtres monodimensionnels passe-bas pseudo Gaussiens H_1 , H_2 et H_3 dont les réponses impulsionnelles sont symétriques et de taille finie à 13, 27 et 53 coefficients respectivement. La Figure 3 montre les réponses fréquentielles de ces filtres. On peut ainsi observer la sélectivité fréquentielle qui augmente avec la taille du filtre, comme on peut aussi observer l'aspect lisse de la bande passante et de la bande d'arrêt due à la propriété Gaussienne (absence d'ondulations ou de lobes dans ces bandes). Chaque filtre est appliqué deux fois sur l'image à filtrer, horizontalement et verticalement, permettant ainsi de sélectionner les basses fréquences horizontales et verticales de l'image. Étant donné la symétrie des filtres, l'extension des bords de l'image utilise l'effet miroir ou la réflexion. L'application de l'opérateur de différence entre deux versions passe-bas successives de l'image produit un niveau passe-bande, tel que décrit par les équations suivantes :

$$G_0 = \text{Image originale}$$

$$L = 3$$

$$G_i(f_1, f_2) = G_{i-1}(f_1, f_2) \cdot H_i(f_1, f_2), \quad i=1,2,\dots,L$$

$$LP_i(m,n) = G_{i-1}(m,n) - G_i(m,n), \quad i=1,2,\dots,L$$

$$LP_{L+1}(m,n) = G_L(m,n)$$

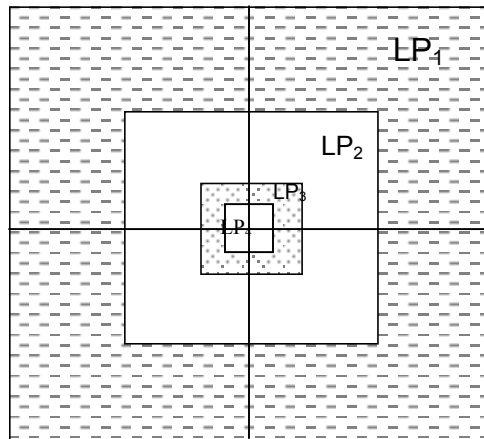


Figure 2 : Représentation fréquentielle d'une décomposition en pyramide Laplacienne redondante à 4 niveaux de résolution ($L=3$). L'approximation passe-bas de l'image est représentée par le niveau LP_4 .

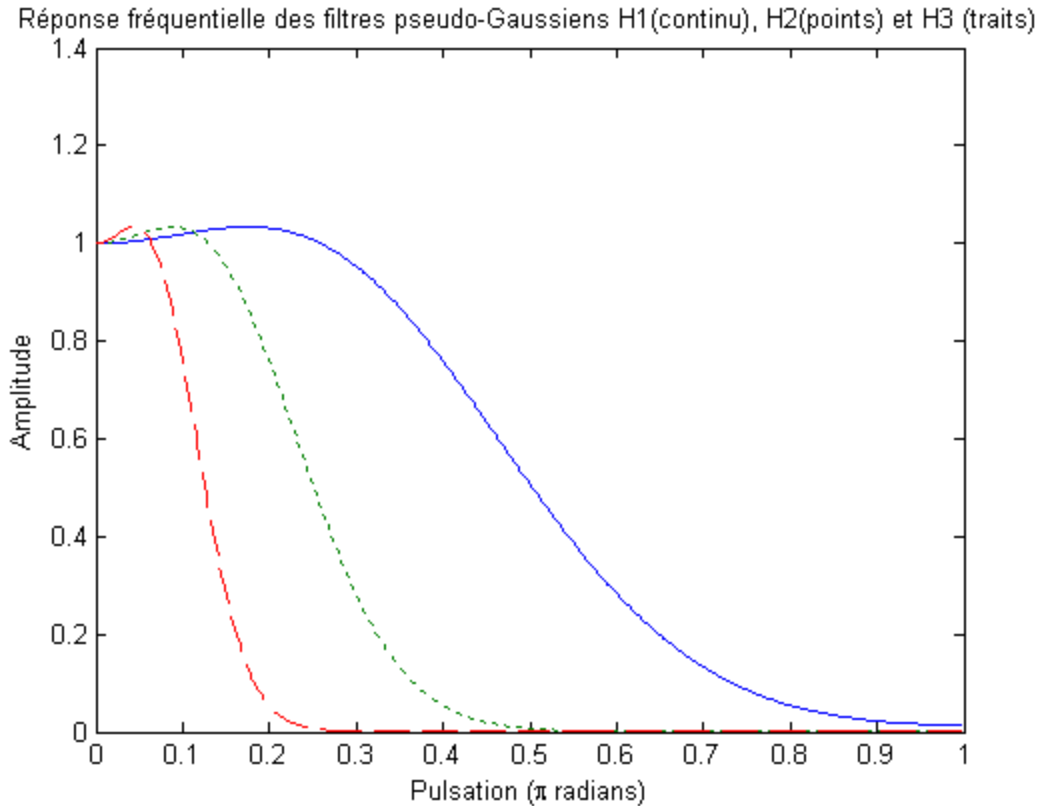


Figure 3 : Amplitude de la réponse fréquentielle des 3 filtres passe-bas pseudo-Gaussiens utilisés pour construire la pyramide Laplacienne redondante.

Noter que bien que tout le traitement soit effectué dans le domaine spatial (opérations de convolutions,...), pour des raisons pratiques, les équations ci-dessus expriment le filtrage dans le domaine fréquentiel bidimensionnel alors que la différence est exprimée dans le domaine spatial. Le résultat obtenu est une pyramide d'images à 4 niveaux ($L=3$). Si l'image originale est une luminance de taille M lignes et N pixels par ligne $G_0(M,N)$, alors chaque niveau LP est de taille similaire $LP_i(M,N)$. Le nombre total de coefficients obtenu est $4MN$, ce qui donne un facteur de redondance (surplus d'échantillon) égal à 3 (ou L) fois la taille originale. La Figure 2 montre la partition fréquentielle du spectre d'amplitude de l'image originale en sous bandes représentant les niveaux LP_i . La construction d'une pyramide Laplacienne redondante avec un nombre de niveaux de résolution supérieur à 4 requiert la définition de filtres additionnels passe-bas encore plus sélectifs. Vu les caractéristiques des réponses impulsionnelles dont les longueurs sont croissantes avec la sélectivité fréquentielle, le problème devient alors celui de trouver les bons filtres avec les bonnes longueurs pour que le filtrage ne soit pas trop coûteux en nombre d'opérations.

IV-3 Mise en œuvre de la Contourlet redondante

Pour implémenter cet étage, un banc de filtres bidimensionnels directionnels tel que montré dans la Figure 4, est utilisé. Chaque image entrée dans ce système est décomposée en 4 sous-bandes orientées en fréquence selon le schéma de la Figure 5. La structure de ce banc de filtres est figée sur deux niveaux de décomposition (donnant lieu à 4 branches de sortie) pour permettre de répondre aux contraintes suivantes :

- Les 4 branches de sortie opèrent avec la même cadence de sous échantillonnage, ce qui permet des sous bandes à tailles égales à la sortie.
- Les filtres et les grilles d'échantillonnage sont choisis de telle sorte que les sous-bandes en sortie soient des sous images définies sur des grilles spatiales rectangulaires. Ce qui va faciliter tout traitement ultérieur des sous-bandes Contourlet.

Comme conséquence de la structure figée, la seule flexibilité que l'on a pour l'instant quant au nombre de sous-bandes de la Contourlet redondante générée réside dans le nombre de niveaux Laplaciens redondants construits.

Les filtres utilisés sont bidimensionnels et requièrent des variantes de sous échantillonnage en quinconce. Tous les filtres utilisés dérivent d'une paire de filtres miroirs en quadrature (H_0 et H_1) en losange et à reconstruction parfaite. La section qui suit détaille la méthode de génération des filtres directionnels de décomposition et de reconstruction.

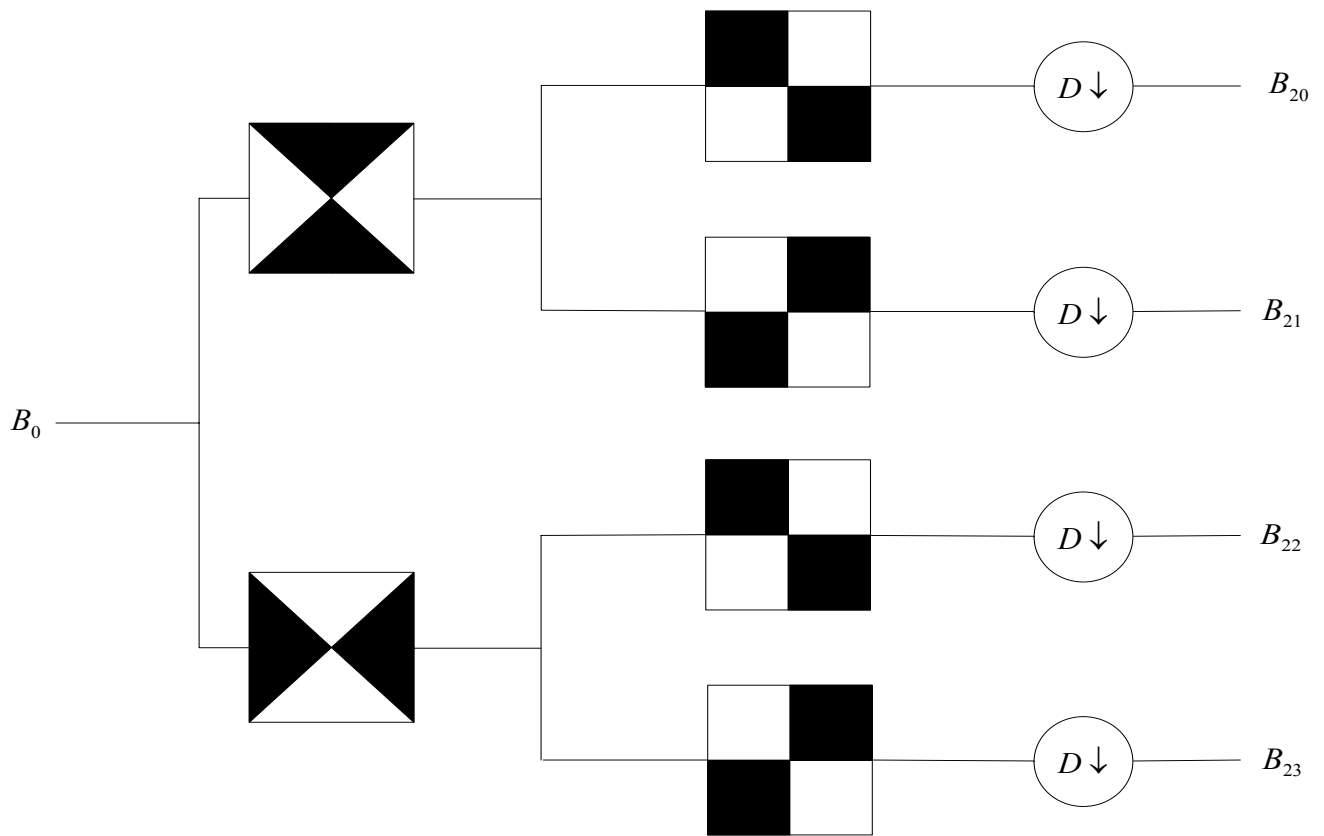


Figure 4 : Diagramme du banc de filtres directionnels utilisé pour générer la décomposition en Contourlet redondante. Chaque image B_0 à l'entrée est décomposée en 4 sous-bandes directionnelles. Les filtres utilisés sont bidimensionnels en forme d'éventail et de damier. L'échantillonnage D est rectangulaire de cadence 4. La taille des sous images B_{ij} est le $\frac{1}{4}$ de la taille de B_0 .

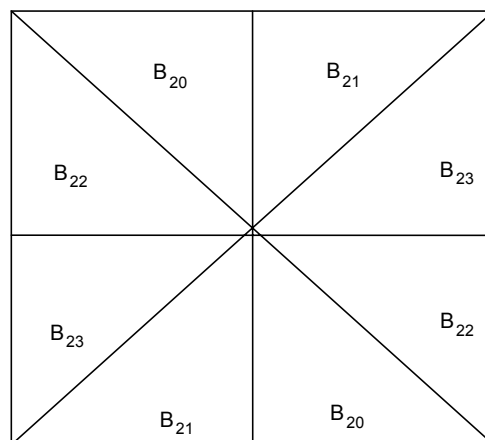


Figure 5 : Représentation fréquentielle d'une décomposition en Contourlet avec 4 directions fréquentielles ($D=4$), appliquée à une seule image B_0 . Les 4 sous-bandes obtenues sont orientées en fréquence et présentent des tailles de $\frac{1}{4}$ par rapport à la taille de B_0 .

Noter que les sous-bandes obtenues en sortie du banc de filtre directionnel sont numérotées dans les Figures 4 et 5 selon le schéma original proposé par l'auteur de la Contourlet. Cette numérotation tient compte du nombre de niveaux de décomposition (2) dans le banc de filtres directionnels, ainsi que du numéro de la branche de sortie (0,1,2,3). Cependant, pour la Contourlet redondante, nous proposons une autre numérotation qui tient compte du niveau de résolution l et de la direction fréquentielle d .

Pour générer la décomposition en Contourlet redondante d'une image, chacun de ses 3 niveaux Laplaciens LP_1 , LP_2 et LP_3 préalablement générés est passé dans le banc de filtres directionnels décrit ci-dessus. Ainsi, 12 sous bandes Contourlet sont générées et à cela se rajoute l'approximation passe-bas de l'image, tel que c'est montré dans les Figures 6 et 7.

Il est possible de réduire la taille totale des coefficients de la Contourlet redondante et de la réorganiser en 13 sous-bandes de tailles toutes égales à $\frac{1}{4} MN$ en appliquant un sous-échantillonnage de cadence 4 à l'image d'approximation passe-bas. Ceci peut être réalisé en apportant quelques modifications au schéma de décomposition en Laplacienne redondante. Dans le cas du diagramme de décomposition de la Figure 1, cette modification consiste à générer G_3 puis à décimer G_3 par une cadence 4 pour donner \hat{G}_3 , puis à interpoler \hat{G}_3 avec une cadence 4 et à la soustraire de G_2 pour obtenir LP_3 . Le niveau LP_4 sera alors l'approximation passe-bas décimée \hat{G}_3 qui est de taille $\frac{1}{4}MN$. Le nombre total des coefficients de la Contourlet redondante sera de $\frac{13}{4} MN$ au lieu de $4MN$. Dans la Figure 8, on peut voir un exemple de décomposition de l'image Lena (512x512) en Contourlet redondante. Les sous-bandes c_{21} , c_{22} , c_{23} et c_{24} sont montrées. On peut à l'occasion voir l'orientation des contours exhibée par chacune des sous-bandes.

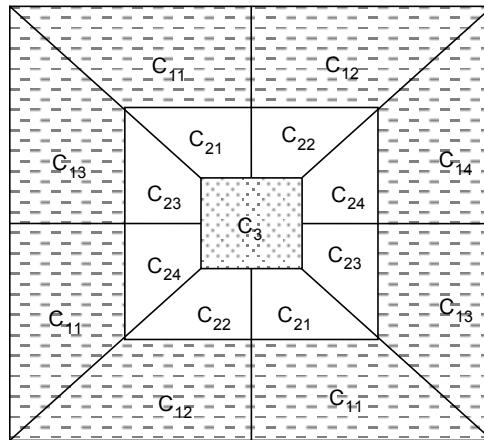


Figure 6 : Représentation fréquentielle d'une décomposition en Contourlet redondante avec 3 niveaux de résolution ($L=2$) et 4 directions fréquentielles ($D=4$). L'approximation passe-bas de l'image (C_3) n'est pas décomposée par le banc de filtres directionnels.

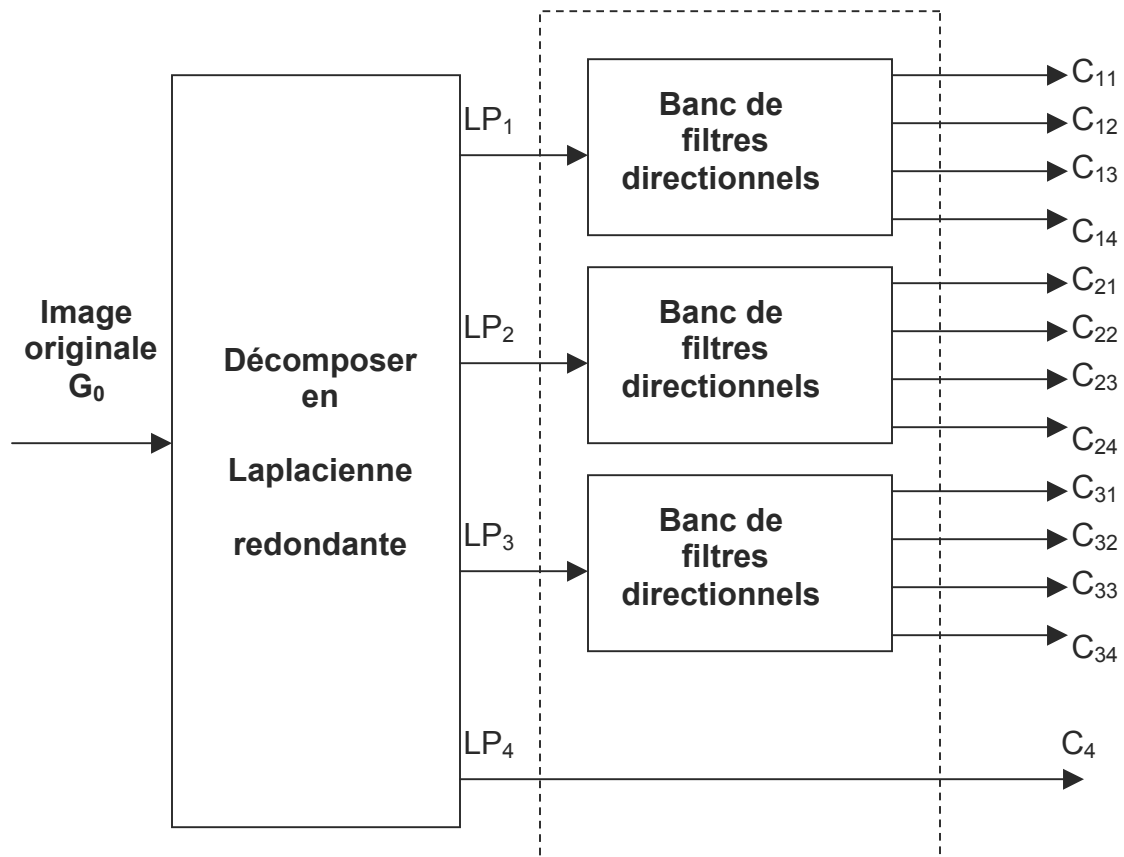


Figure 7 : Diagramme de décomposition en Contourlet redondante à 4 niveaux de résolution ($L=3$) et à 4 directions fréquentielles ($D=4$). La décomposition en Laplacienne redondante est détaillée dans la Figure 1 alors que le banc de filtres directionnels utilisé est détaillé dans le Figure 4. L'image originale de taille MN est ainsi transformée en 12 sous-images de tailles $\frac{1}{4}$ MN et une approximation basse-fréquence de taille MN (C_4).

IV-4 Génération des filtres directionnels

Selon [Phoong et al 95], la paire de filtres miroir en quadrature et à reconstruction parfaite peut être générée à partir des équations ci-dessous en supposant une valeur de départ pour $N=6$:

$$v_k = [0.630 - 0.913 \ 0.0972 - 0.0526 \ 0.0272 - 0.0144], \text{ avec } N = 6$$

$$\beta(z) = \sum_{k=1}^N v_k (z^{-N+k} + z^{-N-k+1})$$

$$H_0(z_0, z_1) = \frac{z_0^{-2N} + z_0^{-1} \beta(z_0 z_1^{-1}) \beta(z_0 z_1)}{2}$$

$$H_1(z_0, z_1) = -\beta(z_0 z_1^{-1}) \beta(z_0 z_1) \cdot H_0(z_0, z_1) + z_0^{-4N+1}$$

$$F_0(z_0, z_1) = -H_1(-z_0, -z_1)$$

$$F_1(z_0, z_1) = H_0(-z_0, -z_1)$$

Le calcul de ces équations a abouti aux filtres H_0 et F_1 dont les réponses impulsionnelles sont de taille 23x23 coefficients, et aux filtres H_1 et F_0 dont les réponses impulsionnelles sont de taille 45x45 coefficients. Pour notre application, seul le filtre H_0 est utilisé pour générer les filtres directionnels en éventail et en damier pour la décomposition en Contourlet, et seul le filtre H_1 est utilisé pour générer les filtres directionnels en éventail et en damier pour la reconstruction.

La génération des filtres en éventail et en damier est obtenue en appliquant des transformations sur un filtre en losange.

Le filtre en éventail horizontal (ou vertical) de décomposition est obtenu en décalant de $\frac{1}{2}$ le filtre H_0 horizontalement vers la droite (ou verticalement vers le haut). Sachant que tout décalage fréquentiel est une modulation temporelle ou spatiale, cette transformation revient à effectuer les opérations suivantes :

$$h_{\text{ver}}(m, n) = h(m, n) \cdot \cos(2\pi n f_0), \quad \text{avec } f_0 = 1/2$$

$$h_{hor}(m, n) = h(m, n) \cdot \cos(2\pi n f_0), \quad \text{avec } f_0 = 1/2$$

Le filtre damier gauche (ou droite) de décomposition est obtenu en appliquant une rotation fréquentielle de 45 degrés dans le sens des aiguilles d'une montre au filtre en éventail horizontal (au filtre en éventail vertical). Notons qu'une rotation de 45 degré dans le domaine fréquentiel équivaut à une rotation de -45 degré dans le domaine spatial. Concernant les filtres directionnels de reconstruction (en éventail et en damier) sont générés de manière similaire mais à partir du filtre passe-bas en losange H_1 .

Il est possible de construire un banc de filtres directionnel à partir de toute paire de filtres bidimensionnels ayant des propriétés d'orthogonalité ou de biorthogonalité et de reconstruction parfaite. L'expérience en a été faite avec une paire de filtres orthogonaux de taille 7x7 [Gau 03].

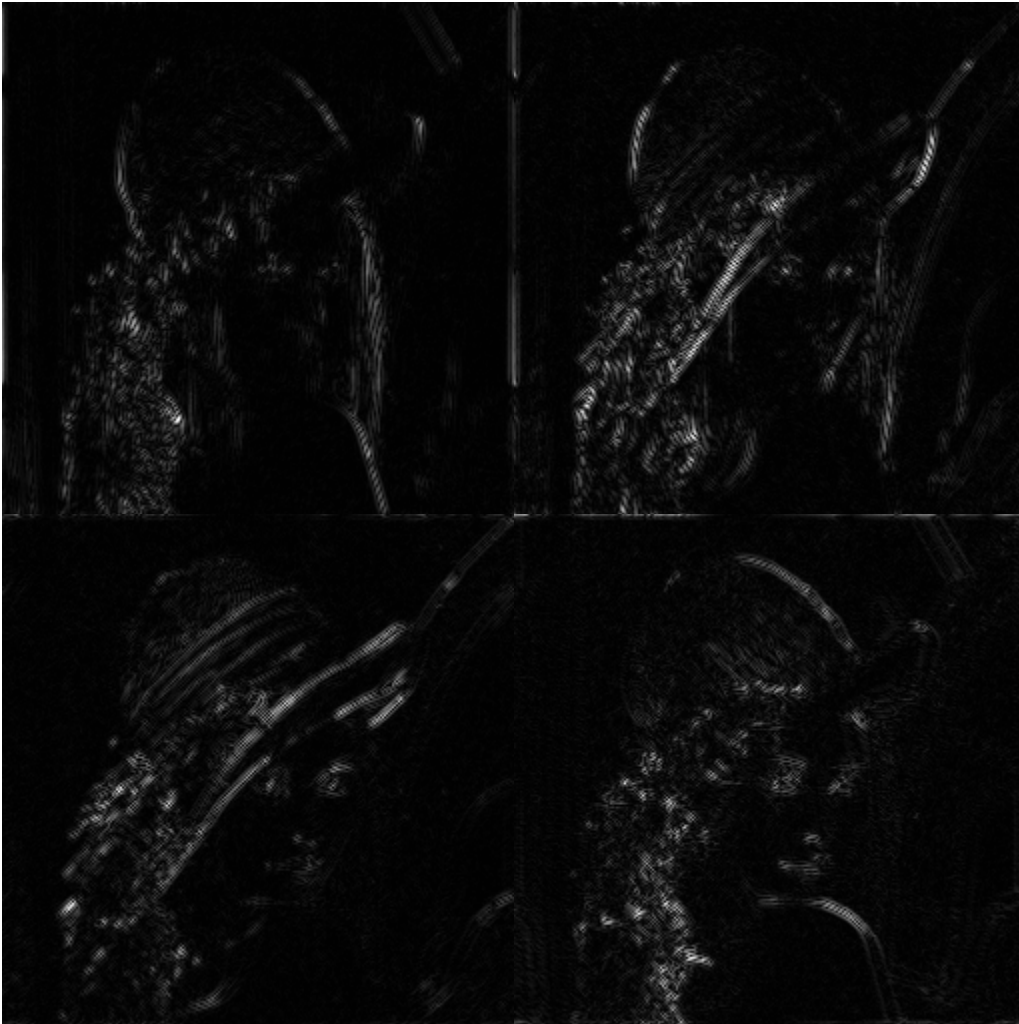


Figure 8 : Décomposition de l'image Lena (512x512) en contourlet redondante à $L=3$ et $D=4$. Les coefficients Contourlet ci-dessus constituent les sous-bandes C21, C22, C23 et C24 , de haut en bas et de gauche vers la droite respectivement. L'orientation des contours selon les résolutions directionnelles est apparente dans les différentes sous-bandes.

IV-5 La Contourlet diagonale redondante

Il s'agit d'une autre variante de la transformée en Contourlet redondante dont l'ensemble des sous-bandes réalisent un autre pavage du domaine fréquentiel tel que montré dans la Figure 9. Ce nouveau pavage est réalisé au niveau de la construction de la pyramide Laplacienne redondante. Au lieu d'effectuer le filtrage passe-bas séparable selon les directions horizontales et verticales (lignes et colonnes de l'image), il faut maintenant les effectuer selon les directions 45 degrés et -45 degrés, ce qui revient à extraire et à filtrer

chaque diagonale de l'image dans la direction de 45 degrés et chaque diagonale de l'image dans la direction -45 degrés. Ce changement de direction de filtrage est équivalent à une rotation du filtre bidimensionnel de 45 degrés pour donner un filtre en losange. Mais comme le filtre est séparable, il est préférable et plus avantageux d'opérer le filtrage en une dimension. Cependant, étant donné que les diagonales d'une image ne sont pas de même taille (contrairement aux lignes et colonnes), le filtrage requiert une bonne gestion de ces tailles ainsi qu'une extension adéquate des bords. Les premières expérimentations d'implémentation de ce filtrage ont révélé quelques problèmes aux bords, notamment aux quatre coins de l'image filtrée. Ceci mérite certainement une attention particulière et va être considéré dans un avenir proche dans la perspective de trouver un remède ou de proposer une amélioration.

Tel qu'on peut le constater dans sa représentation fréquentielle, cette nouvelle variante de la Contourlet redondante offre un autre choix de directions fréquentielles. Nous pensons que les sous-bandes ainsi construites vont rendre l'extraction dans l'image de primitives pertinentes et orientées beaucoup plus aisée. Des expérimentations prévues dans un travail futur vont être menées afin de montrer cette propriété. En se concentrant sur l'extraction de contours d'image orientés, il s'agit de montrer que cette extraction est plus efficace et plus éparsée lorsque la Contourlet redondante est diagonale. Des critères de performance seront établis à cette fin.

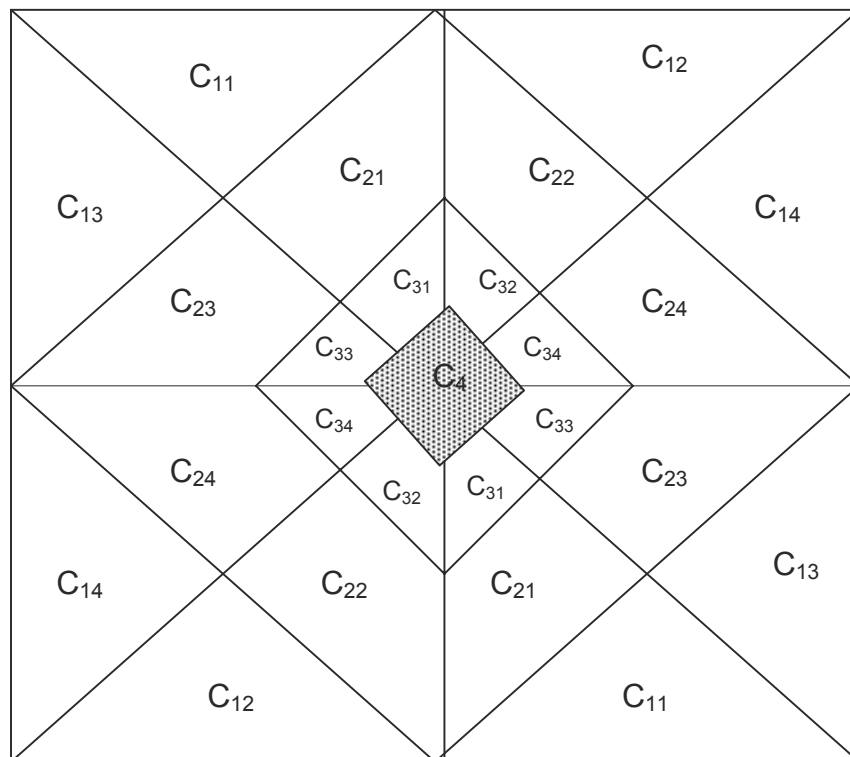


Figure 9 : Représentation fréquentielle d'une décomposition en Contourlet redondante diagonale avec 4 niveaux de résolution ($L=3$) et 4 directions fréquentielles ($D=4$). L'approximation passe-bas de l'image (C_4) n'est pas décomposée par le banc de filtres directionnels.

V- Méthodes d'extraction de primitives

Nous survolons rapidement l'essentiel des méthodes opérant dans le domaine fréquentiel ou spatio-fréquentiel. Nous commencerons par la transformée DCT qui est une transformée linéaire qui permet de représenter le contenu fréquentiel d'un signal en termes d'amplitude ou d'énergie. C'est une transformée d'images qui est largement utilisée comme étape de codage dans les standards de compression connus (JPEG, MJPEG, MPEG). Son application se fait usuellement sur des blocs d'images de luminance ou de chrominance de taille 8x8 pixels. L'extraction de primitives sur des images en format compressé ont donc eu recours à des méthodes opérant sur des blocs DCT. Ainsi on peut sauver la complexité d'avoir à décompresser l'image totalement (ou à la compresser totalement).

Dans [Lay 01], les primitives extraites à partir des blocs DCT sont des histogrammes d'énergie. Pour cela, n patrons sont considérés pour chaque bloc DCT (le coefficient DC, les 4 premiers coefficients, les 8 premiers coefficients sans le DC ou avec le DC, ... etc.). Pour chaque patron l'histogramme de toute l'image DCT est construit. Ainsi chaque image se trouve caractérisée par n primitives représentées par n vecteurs de type histogramme d'énergie. Lors de la recherche par le contenu, la requête ou l'image recherchée doit à son tour être exprimée sous forme de n vecteurs histogrammes. La mesure de similarité consiste à trouver la distance minimale (Euclidienne ou autre) entre vecteurs correspondants. L'image sélectionnée comme étant pertinente par rapport aux critères de recherche est celle qui exhibe le plus grand nombre de distances minimales. Les tests effectués sur plusieurs images ont montré que les primitives extraites sont invariantes par rapport à des transformations telles que l'image miroir, la transposition légère ou la rotation.

Dans [Sut 01], l'extraction de primitives pertinentes revient à segmenter l'image en régions, i.e., chercher des régions qui répondent à une certaine mesure d'homogénéité interne ou bien qui répondent à une certaine mesure de contraste avec d'autres objets ou leurs bords. L'approche adoptée combine l'amélioration des structures de l'image (image enhancement) et la détection de ses contours. Les blocs DCT sont étudiés pour établir le lien entre la distribution fréquentielle et la composition structurelle de l'image. Pour cela, les blocs sont traités en prenant en compte des critères basés sur la perception visuelle. Un bloc DCT contient des informations contours de basses fréquences, de fréquences moyennes et de hautes fréquences. Chaque gamme exhibe une sensibilité particulière. L'algorithme proposé tente de séparer ces gammes en introduisant sur les coefficients DCT une matrice d'amélioration des contours L et un facteur d'ajustement d'énergie α . Grâce à cette matrice, les coefficients DCT sont traités et pondérés différemment. Beaucoup de souplesse est offerte quant au choix des valeurs de L . Dans la partie II de ce rapport, seront abordées d'autres méthodes d'extractions opérant sur d'autres transformées spatio-fréquentielles.

VI- Conclusion

L'extraction de primitives pertinentes est une thématique pour laquelle beaucoup d'effort a été investi par l'ensemble de la communauté des chercheurs. Avec la diversité des applications que connaît de nos jours la technologie de l'information, on réalise combien est nécessaire et profitable l'extraction de primitives pertinentes et leur classification. En effet, ces dernières peuvent être vues comme l'échelon de base d'une hiérarchie de traitement menant à une modélisation sémantique de haut niveau du contenu de l'information visuelle. Notre approche se veut l'exploitation des présentations spatio-fréquentielles et directionnelles pour l'extraction de primitives pertinentes telles que les contours d'images. Notre choix s'est arrêté sur la transformée en contourlet, qui présente a priori, un potentiel intéressant pour l'analyse. Nous avons commencé à proposer des variantes redondantes pour préparer un champ adéquat à l'extraction des primitives, un travail qui va être entamé dans un avenir proche.

VII- Références bibliographiques

- [**Sme 00**] Smeulders A., Worring W., Santini S., Gupta A., and Jain R., ``Content-based image retrieval at the end of the early years``, IEEE transactions PAMI, Vol. 22, No. 12, pp1349-1380, Dec. 2000.
- [**Rui 99**] Rui Y., Huang T. S., and Chang S., ``Image retrieval: current techniques, promising directions, and open issues``, Journal of Visual Communications and Image Representation 10, pp39-62, 1999.
- [**Do 01**] Do M. N., and Vetterli M., ``Pyramidal directional filter banks and Curvelets``, IEEE International Conference on Image processing, Thessaloniki, Greece, Oct. 2001.
- [**Do 02**] Do M. N., and Vetterli M., ``Contourlets: a directional multiresolution image representation``, IEEE International Conference on Image processing, Rochester, USA Sept. 2002.
- [**Sta 02**] Starck J. L., Candès E. J., and Donoho D. L., ``The Curvelet transform for image denoising``, IEEE Transactions on Image Processing, Vol. 11, No. 6, pp670-684, Juin 2002.
- [**Gau 03**] Gauthier V. , `` la transformée en Contourlet pour la detection de primitives d'image``, Rapport final de projet de fin d'études, GEN1054, Département d'informatique et d'Ingénierie, Université du Québec en Outaouais, Août 2003.
- [**Can 02**] Candès E. J., and Donoho D. L., `` Recovering edges in ill-posed inverse problems: optimality of Curvelet frames``, The annals of statistics, Vol. 30, No. 3, pp784-842, 2002.

- [**Baa 92**] Baaziz N. and Labit C. “Wavelet transform and Motion compensation for image sequence coding” in International Conference on Wavelets and Applications, Toulouse, France, June 1992.
- [**Kim 95**] Kim C. W., Vaidyanathan P. P., and Ansari R. “A new class of Two-Channel biorthogonal filter banks and Wavelet bases”, IEEE Transactions on Signal Processing, Vol. 43, No. 3, pp649-665, Mars 1995.
- [**Lay 01**] Lay J. A., and Guan L. “Image retrieval in frequency domain using DCT coefficient histograms”, in “Multimedia Image and video processing”, Edited by Guan L., Kung S. Y. and Larsen J., CRC Press, 2001.
- [**Sut 01**] Suthaharan S., Kim S. W., Wu H. R., and Rao K. R., “Transform domain techniques for multimedia image and video coding”, in “Multimedia Image and video processing”, Edited by Guan L., Kung S. Y. and Larsen J., CRC Press, 2001.
- [**Man 01**] Mandal M. K., Panchanathan S., “Video segmentation in the wavelet compressed domain”, Journal of Visual Communications and Image Representation 12, pp17-28, 2001.
- [**Bow 00**] Bowyer K., Flynn P., “Introduction to content-based image retrieval at the end of the early years”, IEEE transactions PAMI, Vol. 22, No. 12, pp1348-1380, Dec. 2000.
- [**Bru 99**] Brunelli R., Mich O., and C. M. Modena “A survey on the automatic indexing of video data”, Journal of Visual Communications and Image Representation 10, pp78-112, 1999.
- [**Col 99**] Colombo C., Del Bimbo A. and Pala P., “Semantics in visual information retrieval”, IEEE Multimedia , pp38-53, July-Sept. 1999.
- [**Rui 99**] Rui Y., Huang T. S., and Chang S., “Image retrieval: current techniques, promising directions, and open issues”, Journal of Visual Communications and Image Representation 10, pp39-62, 1999.
- [**Idris 97**] Idris F., Panchanathan S., “Review of image and video indexing techniques”, Journal of Visual Communications and Image Representation Vol. 8, No. 2, pp146-166, 1997.
- [**Ahanger 96**] Ahanger G., Little T. D. C., “A survey of technologies for parsing and indexing digital video”, Journal of Visual Communications and Image Representation Vol. 7, No. 1, pp28-43, 1996.
- [**Aig 96**] Aigrain P., Zhang H. and Petrovic D., “Content-based representation and retrieval of visual media: a state of the art”, Multimedia tools and applications, 3, 179-202, 1996.