



**HAL**  
open science

# Représentation structurelle d'images par transformées locales en ondelettes orientées et codage

Guillaume Jeannic

► **To cite this version:**

Guillaume Jeannic. Représentation structurelle d'images par transformées locales en ondelettes orientées et codage. Traitement du signal et de l'image [eess.SP]. Université de Nantes, 2008. Français. NNT: . tel-00350555

**HAL Id: tel-00350555**

**<https://theses.hal.science/tel-00350555>**

Submitted on 7 Jan 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE

SCIENCES ET TECHNOLOGIES  
DE L'INFORMATION ET DES MATHÉMATIQUES

Année : 2008

Thèse de Doctorat de l'Université de Nantes

Spécialité : AUTOMATIQUE ET INFORMATIQUE APPLIQUÉE

*Présentée et soutenue publiquement par*

**Guillaume JEANNIC**

*le 5 novembre 2008*

*à Polytech'Nantes*

**TITRE**

**Représentation structurelle d'images par transformées locales en ondelettes orientées et codage**

Jury

Président	: Marc GELGON	Professeurs des Université (Université de Nantes)
Rapporteurs	: Marc ANTONINI Christine GUILLEMOT	DR CNRS DR INRIA
Examineurs	: Stéphane PATEUX Béatrice PESQUET-POPESCU James E. FOWLER Vincent RICORDEL Dominique BARBA	Ingénieur de Recherche Professeur à l'ENST Paris Professeur à l'Université du Mississippi Maître de Conférence (Université de Nantes) Professeur Émérite (Université de Nantes)

**Directeur de Thèse : Dominique BARBA**

Laboratoire : IRCCyN

Co-encadrant : Vincent RICORDEL

Laboratoire : IRCCyN

Composante de rattachement du directeur de thèse : Polytech'Nantes

N° ED 503-017

à mon père,

## Remerciements

Je tiens tout d'abord à remercier Marc Gelgon, Professeur des Universités à Polytech'Nantes, qui m'a fait l'honneur de présider ce jury.

Je remercie aussi Marc Antonini, directeur de recherche CNRS, et Christine Guillemot, directrice de recherche INRIA, d'avoir bien voulu accepter la charge de rapporteur.

Je remercie Stéphane Pateux, Ingénieur de Recherche, Béatrice Pesquet-Popescu, Professeur à l'ENST Paris, et James E. Fowler, Professeur à l'Université du Mississippi d'avoir bien voulu juger ce travail.

Je tiens à remercier chaleureusement Vincent Ricordel, Maître de Conférence, co-encadrant de cette thèse pour son soutien de tous les jours. Je remercie aussi Dominique Barba, Professeur Émérite, qui a dirigé cette thèse, pour son expérience et ses conseils avisés.

Je remercie mes voisins de bureau successifs : Nicolas Normand, Andrew Kingston, Fadi Boulos et Florent Autrusseau, qui m'ont supporté au cours de ces années.

Je n'oublie pas non plus tous les membres de l'équipe IVC qui ont permis que cette thèse ait lieu dans de bonnes conditions.

## Sommaire

# Sommaire

Sommaire	4
Introduction générale	9
<b>I Représentation multirésolution structurelle</b>	<b>13</b>
Introduction	15
<b>1 Représentations directionnelles : état de l'art</b>	<b>17</b>
1.1 Transformations fixes . . . . .	19
1.2 Transformations adaptatives . . . . .	24
<b>2 Transformation 2D en ondelettes orientées par redressement</b>	<b>29</b>
2.1 Technique de redressement pour l'implantation de la transformée en ondelettes 1D . . . . .	30
2.2 Adaptabilité du redressement dans l'implantation de la transformée en ondelettes 2D . . . . .	32
2.2.1 Redressement basé ligne et redressement basé colonne . . . . .	33
2.2.1.1 8-connexe- $(p, q)$ -redressement . . . . .	34
2.2.1.2 $(p, q)$ -connexe- $(p, q)$ -redressement . . . . .	36
2.2.1.3 $\theta$ -redressement . . . . .	36
2.2.2 Formalisme général basé sur la parité . . . . .	38
2.2.2.1 Décomposition polyphase . . . . .	38
2.2.2.2 Conclusion sur les orientations de filtrage . . . . .	39
2.3 Étude comparative des diverses implantations du redressement orienté . . . . .	40
2.3.1 Contenu rectiligne . . . . .	43
2.3.2 Contenu circulaire . . . . .	53
2.3.3 Contenu naturel . . . . .	55
<b>3 Estimation de l'orientation de filtrage et classification structurelle</b>	<b>75</b>
3.1 Gestion des frontières de blocs pour le filtrage orienté . . . . .	75
3.2 Estimation de l'orientation de filtrage d'une région de l'image . . . . .	76

3.2.1	Estimation de l'orientation des contours extraits explicitement d'une image . . . . .	78
3.2.2	Estimation implicite de l'orientation par minimisation d'une fonctionnelle . . . . .	79
3.2.2.1	Minimisation d'une mesure de régularité liée à l'erreur de prédiction du $\theta$ -redressement . . . . .	80
3.2.2.2	Estimation par modélisation stochastique de l'incertitude sur la mesure de l'orientation du gradient . . . . .	105
	<b>Conclusion</b>	<b>121</b>
	<b>II Quantification et codage</b>	<b>125</b>
	<b>Introduction</b>	<b>127</b>
<b>4</b>	<b>Codage des informations d'orientation et de classification des régions</b>	<b>129</b>
4.1	Codage arithmétique adaptatif des contours d'une image en fonction de la résolution inférieure . . . . .	130
4.1.1	Codage arithmétique et modélisation markovienne de codes de Freeman différentiels . . . . .	131
4.1.2	Entropies des probabilités conditionnelles pour un codage de contours d'images à différentes résolutions . . . . .	132
4.1.3	Codage arithmétique adaptatif de contours d'images à différentes résolutions . . . . .	135
4.1.3.1	Quantification de $C$ . . . . .	136
4.1.3.2	Comparaison avec les méthodes de l'état de l'art . . . . .	136
4.1.3.3	Initialisation du codeur arithmétique avec une loi <i>a priori</i> . . . . .	137
4.2	Codage arithmétique adaptatif des arbres quaternaires définissant les orientations de filtrage et les classes de structure . . . . .	139
4.3	Comparaison du coût de codage entre les deux méthodes . . . . .	140
<b>5</b>	<b>Quantification adaptée à la classification structurelle de l'information</b>	<b>145</b>
5.1	Quantification et codage des éléments d'une représentation en ondelettes standard 2D . . . . .	145
5.1.1	Quantification scalaire uniforme . . . . .	146
5.1.2	Prise en compte des effets de masquage . . . . .	146
5.2	Impact d'une modification de la quantification selon la classe de structure	151
5.2.1	Base de pas de quantification pour la transformée en ondelettes orientées . . . . .	152
5.2.2	Histogrammes des coefficients d'ondelettes orientées . . . . .	154
5.2.3	Pondération des pas de quantification en fonction de la classe structurelle . . . . .	159
	<b>Conclusion</b>	<b>179</b>

<i>Sommaire</i>	7
<b>Conclusion générale et perspectives</b>	<b>181</b>
<b>Annexes</b>	<b>187</b>
<b>A Filtres d'ondelettes et paramètres de redressement associés</b>	<b>187</b>
A.1 Filtres 5/3 . . . . .	187
A.2 Filtres 9/7 . . . . .	188
<b>B Paramètres de <math>\theta</math>-redressement orienté et filtres d'ondelettes orientées associés</b>	<b>191</b>
B.1 Filtres 5/3 pour $0 \leq \theta \leq 45^\circ$ . . . . .	191
B.2 Filtres 9/7 pour $0 \leq \theta \leq 45^\circ$ . . . . .	192
<b>C Base d'images</b>	<b>195</b>
<b>D Exemples de cartes de contours extraits à trois niveaux de résolution</b>	<b>203</b>
<b>E Exemples de classification structurale et d'orientations de filtrage estimées</b>	<b>207</b>
<b>Bibliographie</b>	<b>214</b>



# Introduction générale

Avec l'évolution des technologies numériques, les contenus visuels se sont plus que jamais répandus, et sur des supports de plus en plus variés : télévision sur mobile ou baladeur, vidéo sur internet, télévision haute définition, télévision 3D, etc. Malgré les améliorations des capacités de stockage (disques durs, mémoires flash, etc.) et le déploiement de réseaux de plus en plus haut débit, l'explosion de la quantité de données échangées rend l'étape de compression toujours aussi essentielle.

En supplément de la recherche du meilleur compromis débit-distorsion, des propriétés des techniques de compression qui permettent de nouveaux services sont de plus en plus souhaitées pour s'adapter à la diversité des applications, des réseaux et des terminaux existants. Parmi ces services on peut noter particulièrement la notion de scalabilité : le flux encodé d'un contenu à un certain niveau de qualité et à une certaine résolution peut être tronqué pour s'adapter à des capacités de débit plus faible, à des résolutions spatiales d'affichage plus petites, ou s'adapter à des applications ne requérant pas des niveaux maxima de qualité.

Pour le codage d'images fixes, le standard de compression scalable actuel est JPEG 2000 (ISO/CEI 15444-1). Il s'agit d'une norme commune à l'ISO, la DEI et l'UIT-T produite par le groupe de travail *Joint Photographic Experts Group*. En plus des performances en compression de JPEG 2000 supérieures à JPEG, c'est la multitude de nouvelles caractéristiques telles la scalabilité, les régions d'intérêt, la résistance aux erreurs de transmission, le codage sans perte, qui fait l'intérêt de cette norme.

Les nouveautés par rapport à JPEG du point de vue compression sont l'utilisation d'une transformée en ondelettes, qui offre une scalabilité en résolution naturelle, mais surtout d'un algorithme de codage entropique très sophistiqué basé sur l'algorithme EBCOT de David Taubman[Tau98].

Toutefois la transformée en ondelettes pour l'image n'est qu'une simple extension bidimensionnelle de la transformée en ondelettes 1D. À ce titre elle n'est pas adaptée à représenter efficacement des caractéristiques géométriques non horizontales ou non verticales. Par exemple, l'énergie des coefficients d'ondelettes relatifs à un contour qui n'est ni parallèle à l'horizontale ni à la verticale sera répartie dans les différentes sous-bandes d'un même niveau de décomposition de la transformée en ondelettes au lieu d'être concentrée dans un seul comme c'est le cas pour des contours strictement horizontaux ou verticaux. De plus, dans la transformée en ondelettes 2D, tout élément de l'image est filtré de façon symétrique suivant les deux orientations horizontale et verticale. Ce traitement isotrope ne permet pas d'exploiter la différence de dynamique

qui peut exister lorsque l'on considère des caractéristiques géométriques orientées d'une image.

Pour pallier à ces limitations, une nouvelle génération de transformées directionnelles a été développée qui s'appuie principalement sur la transformée en ondelettes. Certaines proposent une famille génératrice de fonctions orientées plus complète que les fonctions de base de la transformée en ondelettes. D'autres s'intéressent à une représentation adaptative de l'image. Soit l'image originale est déformée pour l'adapter au filtrage fixe de la transformée en ondelettes 2D, soit la base de la transformée est déformée pour s'adapter au contenu de l'image. Dans le cas des représentations adaptatives, les paramètres de déformation de l'image, ou de la base d'ondelettes doivent être transmis avec les coefficients de la transformée appliquée.

Partant des transformées adaptatives existantes, l'objectif de notre travail est de définir une représentation qui tiennent à la fois compte des orientations des caractéristiques géométriques de l'image, mais aussi du caractère anisotrope de telles caractéristiques. Une approche habituelle pour prendre en compte cette anisotropie passe par une modification des filtres utilisés pour chaque orientation de filtrage. Nous chercherons dans notre approche à prendre en compte cette différence de traitement au niveau de la quantification en fonction de la statistique des sous-bandes produites par notre représentation, et de leur impact visuel.

Ce manuscrit s'articule en deux parties. La première traitant de la construction d'une représentation satisfaisant notre objectif. La seconde de son exploitation dans un contexte de codage.

Parmi toutes ces nouvelles transformées, nous avons choisi de nous intéresser aux transformées adaptatives implantées par des techniques de « redressement » (*lifting*). La flexibilité du redressement permet la construction de transformées en ondelettes orientées. Ces transformées s'adaptent aux orientations des structures géométriques de l'image. La base de fonctions de la transformée est donc définie localement par deux orientations variables.

Après un chapitre introductif sur l'état de l'art des représentations directionnelles, un premier travail, présenté dans le deuxième chapitre, a consisté en une étude comparative de différentes variantes de transformées en ondelettes orientées implantées par des techniques de redressement. Pour le premier niveau de décomposition, nous avons considéré la première étape de ces transformées, c'est-à-dire l'opération de redressement selon la première orientation de filtrage. La capacité de chaque variante à minimiser l'énergie de la sous-bande des hautes fréquences a été analysée en fonction de l'orientation du contenu et de l'orientation de filtrage.

Pour s'adapter au contenu d'une image, les orientations locales de filtrage doivent être estimées. Le chapitre trois considère différentes approches pour analyser la géométrie de l'image. Deux informations sont alors considérées : d'une part l'orientation des structures géométriques présentes dans l'image, et d'autre part une caractérisation de ces structures.

La représentation de ces informations, qui peut prendre différentes formes en fonction la méthode d'estimation considérée, doit être transmise au décodeur pour permettre la reconstruction de l'image. Le quatrième chapitre, premier de la seconde partie, s'in-

téresse au coût de codage de ces représentations.

Dans le dernier chapitre, le codage des coefficients d'ondelettes orientées est réalisé. Les lois de quantification sont alors adaptées en fonction de la représentation structurelle de l'image.

Finalement nous concluons ce manuscrit et nous intéressons aux perspectives possibles de ce travail.



Première partie

Représentation multirésolution  
structurelle



# Introduction

Avec l'arrivée de JPEG 2000 comme successeur de JPEG pour la compression d'images fixes, la transformée en ondelettes tend à s'imposer face à la transformée en cosinus discrète. À présent les limites de cette transformée font l'objet d'attentions particulières. Une nouvelle génération de transformées se fait jour pour succéder aux ondelettes. Dans un premier chapitre introductif les limites de la transformée en ondelettes ainsi qu'une présentation non exhaustive de transformées directionnelles seront présentées.

Parmi le dédale de ces nouvelles transformées, nous avons choisi de nous intéresser aux transformées adaptatives implantées par des techniques de « redressement » (*lifting*). La flexibilité de cette technique permet la construction de différentes variantes de transformées en ondelettes orientées. Dans un deuxième chapitre le formalisme des techniques de redressement pour la transformée en ondelettes sera rappelé. À partir de la définition de transformée en ondelettes orientées implantée par redressement orienté, nous approfondirons trois techniques de redressement particulières. Ces trois transformées seront alors étudiées comparativement du point de vue de leur capacité à exploiter la régularité de la géométrie des images.

L'estimation de cette géométrie pour guider l'analyse des transformées en ondelettes orientées sera l'objet du troisième chapitre. Différentes solutions pour l'estimation de l'orientation de filtrage seront considérées.



# Chapitre 1

## Représentations directionnelles : état de l'art

### 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 efficacement les singularités appartenant à des contours horizontaux, verticaux ou diagonaux. Pour une description complète de la transformée en ondelettes, nous renvoyons le lecteur vers deux ouvrages de référence [Mal99, Ant03].

On peut toutefois noter une différence de traitement entre les contours horizontaux et verticaux d'une part, et diagonaux d'autre part. En effet la base de décomposition relative aux coefficients diagonaux résultent du produit tensoriel de deux fonctions d'ondelettes. Ces coefficients représentent donc théoriquement des singularités ponctuelles. En pratique les coefficients sont issus du passage séquentiel d'un filtre passe-haut à réponse impulsionnelle finie sur les lignes et sur les colonnes de l'image. Il en résulte que les éléments de contours pour lesquels la réponse aux filtres est maximale correspond aux orientations diagonales, car ponctuellement elles présentent une double singularité verticale et horizontale. Il est de plus important de noter que ces coefficients diagonaux vont caractériser deux orientations différentes :  $-45^\circ$  et  $45^\circ$ . Par contre la base relative aux coefficients verticaux (ou horizontaux) résulte du produit tensoriel d'une fonction d'ondelettes et d'une fonction d'échelle. La forme des fonctions de base associées fait apparaître, contrairement aux fonctions de base relatives aux coefficients diagonaux, une forme de contour. La taille de ce contour, son support, est liée à la taille du filtre passe-bas associé.

Ainsi les contours d'une image qui offrent une réponse maximale à l'un des filtres d'analyse sont orientés horizontalement, verticalement ou diagonalement. 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 aussi d'une limitation de la transformée en ondelettes. En effet si elle permet donc 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. Ce point est illustré sur la figure 1.1.

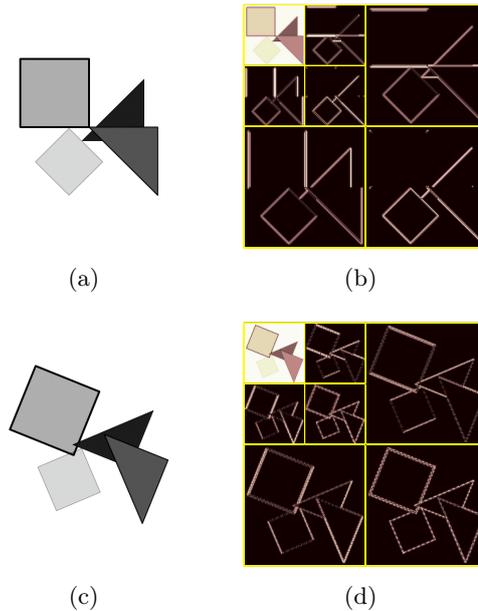


FIGURE 1.1 – (a) Une image de test avec des contours purement horizontaux, verticaux ou obliques à  $\pm 45^\circ$  et (b) sa transformée en ondelettes (analyse dyadique sur trois niveaux). (c) La même image de test mais dont les éléments ont subi une rotation de  $22^\circ$  et (d) sa transformée en ondelettes.

Pour pallier aux insuffisances de la transformée en ondelettes, une liste de souhaits a été proposée [DV05] pour une transformation idéale qui s'appuierait sur les qualités de la transformée en ondelettes, à savoir la multirésolution, la localisation en espace et en fréquence et la décimation critique, auxquels s'ajouteraient la multi-directionnalité, et la possibilité de traiter différemment une direction de filtrage plutôt qu'une autre (anisotropie du filtrage).

Les trois premières caractéristiques sont des acquis de la transformée en ondelettes. Les deux suivantes constituent les améliorations à apporter. Dans la nouvelle vague de transformées qui sera présentée on verra que la prise en compte de l'orientation et de l'anisotropie du support se fera parfois au détriment des acquis des ondelettes, et en particulier de la décimation critique.

Les contours sont au centre de toutes ces transformées. L'idée étant d'être capable de représenter les contours forts avec un faible nombre de coefficients de traiter des orientations diverses et variées.

On peut trouver dans la littérature deux voies différentes dans l'élaboration de nouvelles transformées satisfaisant au mieux les propriétés énoncées ci-dessus : les transfor-

mations fixes et les transformations adaptatives. Le premier type propose une transformée qui analyse identiquement toutes les images mais via une famille génératrice prédéfinie riche d'orientations multiples. Le second type propose une transformée qui s'adapte (sa base) au contenu de chaque image, ou bien qui adapte l'image à la transformée en ondelettes. Ceci nécessite bien entendu une étape supplémentaire de pré-traitement des images, et un codage des paramètres de la transformation, du dictionnaire de la base, etc.

## 1.1 Transformations fixes

Les transformations fixes ont ici pour objectif de décomposer un signal selon une famille génératrice de signaux plus fournie que la base de la transformée en ondelettes 2D, le but étant de permettre de traiter plus d'orientations. Parmi ces transformations nous choisissons de décrire les trois fondamentales : la transformée en ridgelettes[Can98], la transformée en curvelettes[SCD00], la transformée en contourlettes[DV05] et la transformée en ondelettes complexes[SBK05].

**Transformée en ridgelettes** Alors que la transformée en ondelettes 2D continue séparable d'un signal  $s$  peut être définie comme le produit scalaire de ce signal avec le produit tensoriel de deux fonctions d'ondelettes  $\psi_{a_1, a_2, b_1, b_2}(x_1, x_2) = \psi_{a_1, b_1}(x_1) \cdot \psi_{a_2, b_2}(x_2)$  avec :

$$\psi_{a,b}(t) = a^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), \quad (1.1)$$

la transformée en ridgelettes continue est définie comme le produit scalaire de ce signal avec la fonction de ridgelette :

$$\psi_{a,b,\theta}(x_1, x_2) = a^{-\frac{1}{2}} \psi\left(\frac{x_1 \cos \theta + x_2 \sin \theta - b}{a}\right). \quad (1.2)$$

Avec  $b, b_1, b_2$  des paramètres de translation, et  $a, a_1, a_2$  des paramètres de dilatation. Cela signifie que la fonction de base de l'espace sur lequel est projeté le signal  $s$  est une fonction d'ondelettes orientée en fonction de l'angle  $\theta$ , constante selon les lignes  $x_1 \cdot \cos \theta + x_2 \cdot \sin \theta = Cste$ . La transformée en ridgelettes continue permet donc en théorie d'être appliquée selon une orientation quelconque.

La transformée en ridgelettes est associée à la transformée de Radon, car précisément la transformée en ridgelettes correspond à une transformée en ondelettes appliquées sur la projection de la transformée de Radon.

Le problème de la discrétisation de la transformée en ridgelettes se ramène donc au problème de la discrétisation de la transformée de Radon. Pour calculer les projections d'une transformée de Radon d'une image, il existe deux stratégies. Une stratégie fréquentielle et une spatiale.

La première utilise le théorème de la tranche centrale [DM83]. La transformée de Fourier 2D de l'image est d'abord calculée. Les coefficients de Fourier de chaque ligne

$L_\theta$  d'orientation  $\theta$  passant par l'origine sont ensuite extraits, et la transformée de Fourier inverse de ces lignes correspond aux projections de la transformée de Radon (voir figure 1.2). La stratégie spatiale correspond quant à elle à une implantation discrète de la transformée elle-même.

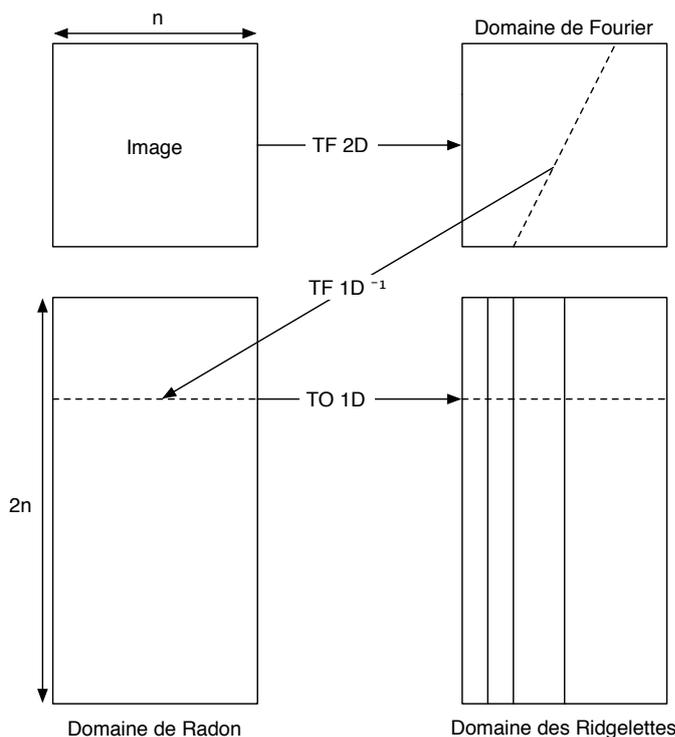


FIGURE 1.2 – Transformée en ridgelettes par la stratégie de Fourier.

Suivant la stratégie choisie, le choix dans l'implantation reste entier. Pour la méthode utilisant le théorème de la tranche centrale, Starck *et al* [SCD00] extraient  $2n$  lignes de  $n$  échantillons de la transformée de Fourier de l'image. Ces échantillons sont obtenus par interpolation des points de la grille discrète, et l'interpolateur utilisé est celui du plus proche voisin. Le choix de l'interpolateur n'est pas évident car l'interpolation des coefficients d'une transformée de Fourier n'est pas de même nature que l'interpolation des coefficients d'une image naturelle.

Carré et Andrés [CA04] définissent quant à eux des droites discrètes analytiques  $L_\theta$  de direction  $(p, q) \in \mathbb{Z}^2$  d'épaisseur  $\omega$  :

$$L_{[p,q]}^\omega = \left\{ (x_1, x_2) \in \mathbb{Z}^2 \mid |qx_1 - px_2| \leq \frac{\omega}{2} \right\}$$

L'ensemble des lignes  $L_{[p,q]}^\omega$  est défini de telle sorte qu'il doit contenir l'ensemble des coefficients de la transformée de Fourier.

Concernant la stratégie spatiale, la transformée Mojette [Gué01] formalise une transformée de Radon discrète et son inversion. Les dimensions de l'image ainsi que le choix

des projections utilisées, i.e. leurs orientations, introduisent plus ou moins de redondance permettant la reconstruction de l'image après transformée.

Sous certaines conditions, Do [DV03] propose une transformée ridgelette non redondante. Les images doivent être carrées et de dimensions premières  $p$ . L'idée est d'appliquer une transformée Mojette torique sur une image auquel on a soustrait sa moyenne (voir figure 1.3 ). De part les dimensions particulière de cette image, chaque projection contient un même nombre  $p$  de bins. Comme l'image sur laquelle se calcule la projection est de moyenne nulle, la somme des  $p$  bins est aussi nulle, par conséquent on peut représenter chaque projection par  $p - 1$  bins. La transformée s'effectue avec  $p + 1$  directions de projection soit un total de  $(p - 1)(p + 1) + 1 = p^2$  coefficients en incluant la valeur moyenne de l'image.

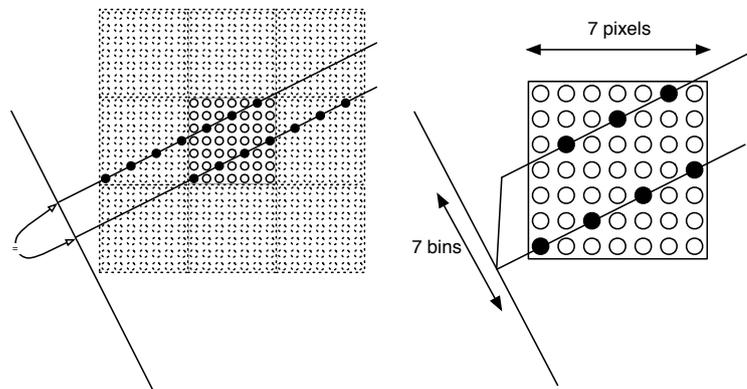


FIGURE 1.3 – Exemple de projection avec une mojette torique de direction (2,1). On remarque que deux droites traversant l'image auront la même valeur de bins en raison de la périodisation torique de l'image. Toute projection avec la mojette torique sur une image carrée de taille première renvoie donc une projection contenant autant de bins que la dimension d'un côté de l'image.

**Transformée en curvelettes** La transformée en ridgelettes est efficace surtout et seulement pour les singularités rectilignes aux dimensions de l'image. Pour pallier à ce défaut, une version par blocs a été introduite [SCD00] pour traiter plus localement les contours. L'idée étant que sur une portion de l'image suffisamment petite, les contours seront plus facilement approchés par un modèle de type rectiligne. La technique utilisée pour mettre en oeuvre cette transformée ridgelette par blocs fait appel au chevauchement de blocs (overlapping) afin d'éviter d'éventuels effets de blocs. Toutefois celle-ci tend à rajouter encore plus de redondance à la transformée.

La transformée curvelette [CD99] a été définie dans sa première version comme une transformée ridgelette par blocs appliquée sur une décomposition multi-résolution en sous-bandes de l'image originale, la taille des blocs pouvant varier selon la sous-bande considérée. Une seconde version [CD02] a été définie sans faire intervenir la transformée ridgelette.

**Transformée en contourlettes** Contrairement aux deux premières transformées présentées, les contourlettes sont définies intrinsèquement dans le domaine discret. L'objectif des contourlettes est de fournir une transformée discrète se rapprochant de l'idée de la transformation curvelette à savoir une transformée multirésolution et multidirectionnelle.

La transformée contourlette est donc une décomposition en sous-bandes sur lesquelles on applique une décomposition directionnelle. Ces deux étapes se font séquentiellement à l'aide de bancs de filtres. La décomposition obtenue n'est pas radiale mais elle est directionnelle.

**Pyramide laplacienne** La première étape de la décomposition fait appel à une pyramide laplacienne. À partir d'une image en entrée  $x$  on obtient la composante basse-fréquence décimée  $c$ , et l'erreur de prédiction  $d$  entre l'image en entrée et la sortie  $c$  interpolée et filtrée (voir figure 1.4).

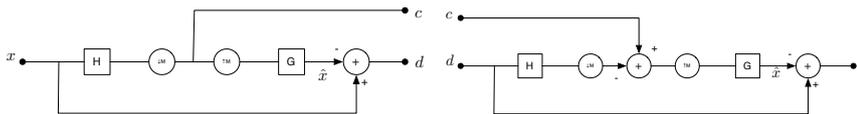


FIGURE 1.4 – Schéma de décomposition et de reconstruction d'une pyramide laplacienne.

Le désavantage de cette décomposition est qu'elle n'est pas à décimation critique. Seule la composante basse fréquence est décimée. Le taux de sur-échantillonnage est d'au maximum  $\frac{4}{3}$ , taux obtenu si on ré-applique à l'infini la décomposition laplacienne sur la composante basse-fréquence  $c$ .

**Banc de filtres directionnels** La décomposition par banc de filtres directionnels reprend le principe de la décomposition de Bamberger et Smith[BS92] qui était à décimation critique et reconstruction parfaite. La version de Bamberger et Smith impliquait l'utilisation de modulation, de banc de filtres quinconces avec des filtres diamant. La nouvelle version du banc de filtres directionnels fait intervenir uniquement des filtres diamants et des opérateurs de sous-échantillonnage. Le principe de ces opérateurs est de transformer l'allure du spectre afin de l'adapter au filtre diamant utilisé (voir figure 1.6).

Les deux premiers étages permettent de décomposer le signal en quatre sous-bandes (voir figure 1.5) à l'aide de filtre diamant et de deux opérateurs de sous-échantillonnage. (voir figure 1.6). L'opérateur de décimation par la matrice  $Q_0$  induit une rotation de l'image originale de  $-45^\circ$ , par conséquent le spectre en sortie du premier filtre sablier se retrouve coupé en deux après passage à travers le deuxième filtre sablier. Le même principe est appliqué ensuite pour redécouper les différents spectres obtenus.

**Banc pyramidal de filtres directionnels** La décomposition directionnelle est appliquée sur les images de différence de notre pyramide laplacienne. Le découplage entre la décomposition radiale et la décomposition angulaire permet à chaque niveau

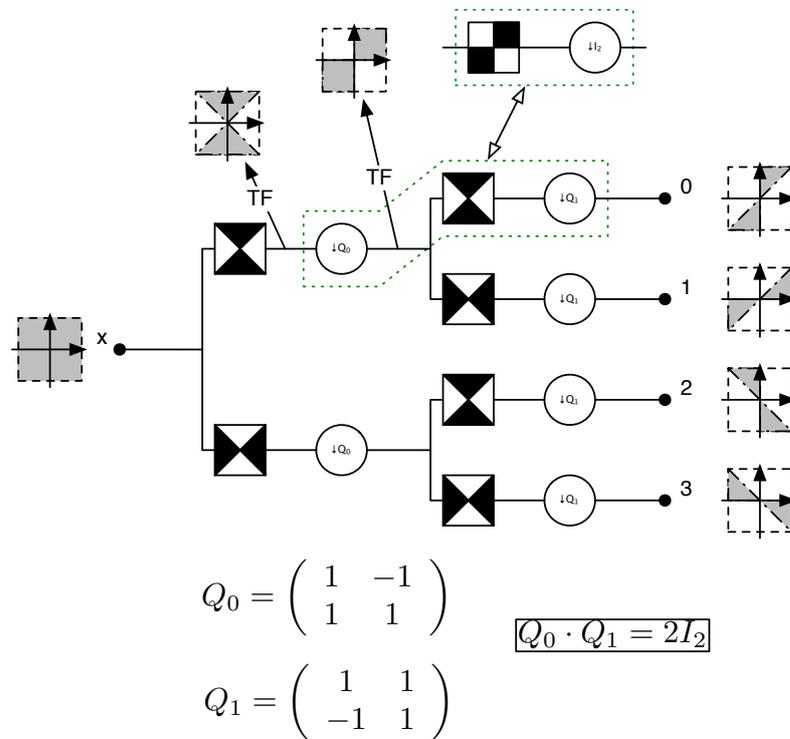
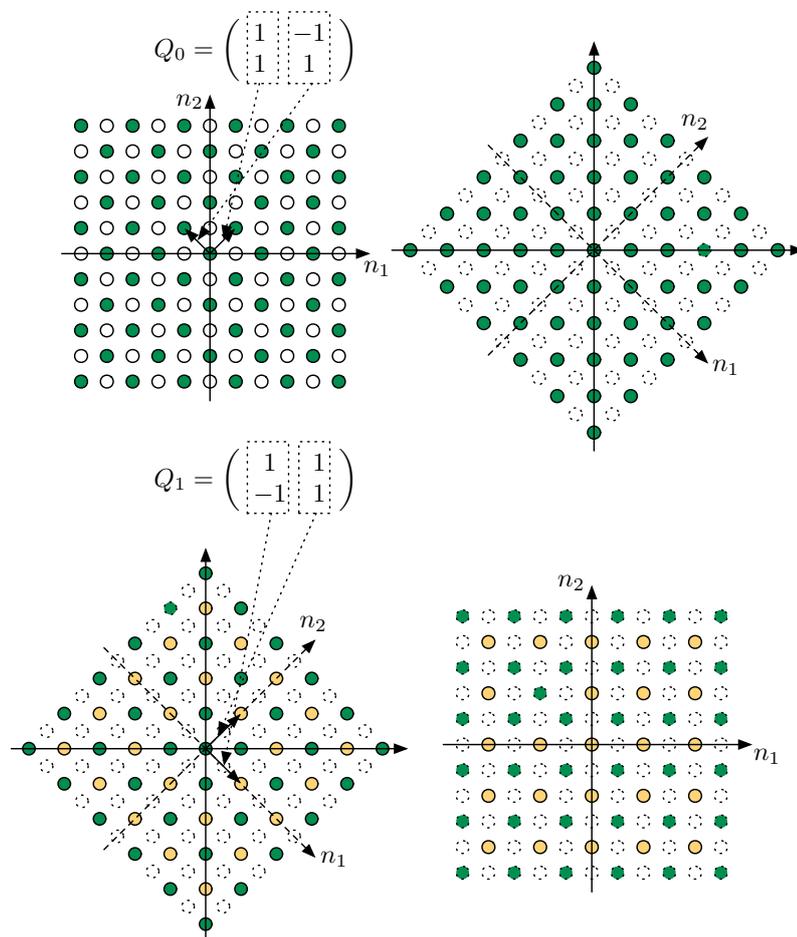


FIGURE 1.5 – Décomposition d’un signal en quatre sous-bandes représentant quatre orientations dans l’image.

de la décomposition radiale de fixer indépendamment le nombre de niveaux de la décomposition angulaire (voir figure 1.7). De plus, étant donnée que la décomposition est dichotomique, on peut décomposer indépendamment chaque sous-bande pour obtenir une décomposition similaire aux paquets d’ondelettes.

L’implantation de la transformée en contourlettes fait appel à des filtres diamant qui ont un grand support, entraînant un étalement des coefficients d’ondelettes, effet de vagues, etc. L’aspect redondant induit par les contourlettes n’est pas un facteur de rejet aussi important que pour la transformée en ridgelettes ou curvelettes, d’ailleurs une nouvelle version de la transformée contourlette [LD03] a été développée à décimation critique, mais conservant les autres problèmes précédemment énoncés.

**Transformée en ondelettes complexes** La « 2D dual-tree complex wavelet transform » [SBK05] permet de décomposer pour un niveau donné une image en 6 sous-bandes hautes-fréquences correspondant aux orientations  $-75^\circ$ ,  $-45^\circ$ ,  $-15^\circ$ ,  $15^\circ$ ,  $45^\circ$  et  $75^\circ$  avec une redondance de 4.

FIGURE 1.6 – Sous-échantillonnage spatial par les matrices  $Q_0$  et  $Q_1$ .

## 1.2 Transformations adaptatives

Comme leur nom l'indique, les transformations adaptatives s'adaptent au contenu de l'image sur laquelle elles sont appliquées. Les transformations que nous considérerons s'appuient sur la transformation en ondelettes. Deux philosophies existent alors concernant la manière de procéder pour construire une telle transformation en ondelettes adaptatives. D'un côté adapter l'image pour la transformation en ondelettes, de l'autre adapter la transformation en ondelettes pour l'image. Pour l'une comme pour l'autre, une description réduite des caractéristiques de l'image est nécessaire pour guider l'application de la transformation, et permettre la reconstruction. Si l'intérêt de ces transformations est d'hériter de la propriété de décimation critique, leur inconvénient est l'estimation, la représentation et le codage de la description réduite nécessaires à leur bon fonctionnement.

Dans [LG08], l'idée est de déformer le contenu d'une image pour l'adapter à la

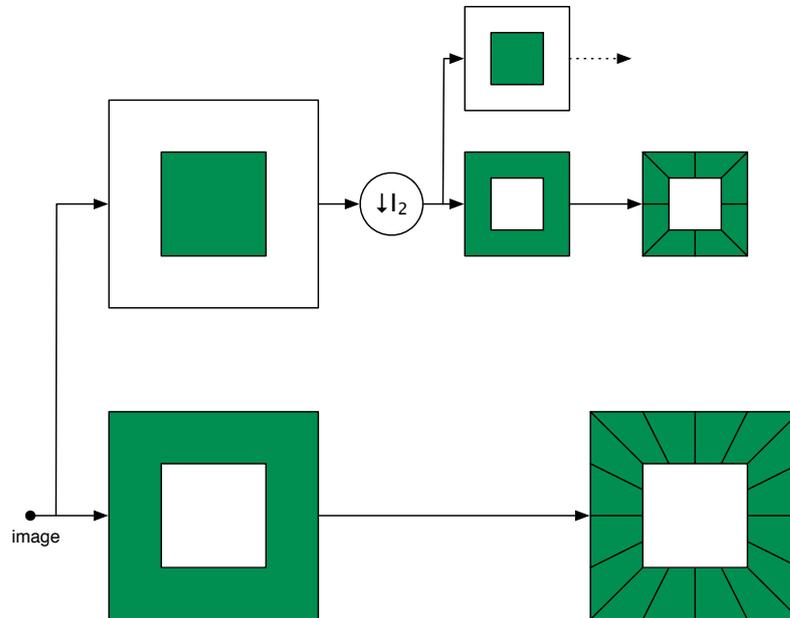


FIGURE 1.7 – Banc de filtre pyramidal directionnel.

transformation en ondelettes standard. La déformation de l'image est modélisée par un maillage déformable. Le critère d'adaptation utilisé est le coût de description de l'image déformée. Les paramètres du maillage sont estimés par la minimisation d'une fonction de coût. L'image après déformation est transmise à un codeur JPEG 2000. À la reconstruction, l'image déformée est décodée. L'opération de déformation n'étant toutefois pas parfaitement réversible, l'image originale n'est pas reconstruite à l'identique. Ceci n'est toutefois pas gênant dans un contexte de compression avec pertes. L'avantage de cette méthode est de pouvoir bénéficier de l'implantation avancée du codage JPEG 2000.

La transformation peut au contraire être adaptée au contenu de l'image. Les filtres utilisés pour l'implantation de la transformée en ondelettes séparable peuvent être adaptés localement [PPPH02, PPPH05, HPPP06]. Nous considérerons toutefois les transformations qui adaptent géométriquement la transformée en ondelettes.

La transformée en bandelettes a pour but d'exploiter localement la régularité géométrique de l'image afin d'optimiser la transformée en ondelettes séparables. Si l'idée est d'effectuer une déformation d'une base d'ondelettes séparable afin que celle-ci s'adapte mieux aux contours de l'image, dans la pratique pour la première version de la transformée en bandelettes [LP02], un redressement local des contours importants de l'image est opéré de telle sorte qu'ils deviennent soit verticaux, soit horizontaux.

Ces contours sont extraits explicitement par un algorithme *ad hoc*. Plusieurs versions ont été présentées par Le Pennec pour obtenir une segmentation à partir des chaînages obtenus. La première [LP02] utilise la méthode du feu de prairie. La région autour du contour est propagée dans la direction orthogonale au contour jusqu'à une certaine

distance maximale. Les zones de l'image non atteinte constitue le fond. Cette distance maximale doit être déterminée de telle sorte que l'approximation qui consiste à dire que les iso-intensités dans la région considérée sont parallèles à la direction de régularité maximale reste valable. Ce type de segmentation étant très coûteuse dans une optique de compression, deux autres techniques [LPM03, LPM05] ont été développées dans le but d'économiser sur la segmentation. Ces deux méthodes utilisent une décomposition de l'image en arbre quaternaire. La première part d'un arbre complètement développé, et opère une fusion si le coût de débit-distorsion estimé d'un père est moins élevé que la somme de celui de ses quatre fils. La deuxième débute elle avec la racine de l'arbre et scinde chaque élément récursivement s'il ne vérifie pas des conditions géométriques du type taille minimal, proximité d'un unique contour inférieur à une certaine distance, etc.

La géométrie est modélisée pour chaque région par une fonction spline qui permet de paramétrer la direction de régularité maximale. Dans le cas d'un contour plutôt horizontal (voir figure 1.8) (respectivement vertical), l'opération de redressement consiste alors en une translation dans la direction verticale (respectivement horizontale). Cette translation peut être non entière, Le Penec [LP02] propose l'utilisation d'interpolateur basé sur des fonctions splines cubiques. L'opérateur inverse de redressement est approché par la translation opposée.

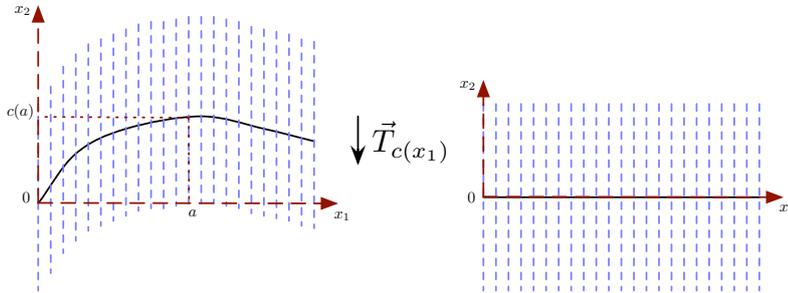


FIGURE 1.8 – Redressement d'un contour plutôt horizontal par translation suivant le vecteur  $\vec{T}_{c(x_1)}$ .

L'image est décomposée en différents domaines, ceux où la direction de régularité est plutôt horizontale, ceux où elle est plutôt verticale, et le fond, où aucune direction n'est privilégiée. Dans chaque région où une direction privilégiée a été détectée, un redressement local s'opère pour qu'enfin on puisse y appliquer un processus de bandelettisation : une transformée en ondelettes 2D est appliquée, suivi d'une transformée en ondelettes 1D selon les lignes de la sous-bande des détails horizontaux (respectivement les colonnes de la sous-bande des détails verticaux) si la direction de régularité est plutôt horizontale (respectivement verticale). Le processus de bandelettisation est illustré sur la figure 1.9.

Dans la seconde version de la transformée en ondelettes [PM05], l'idée d'exploiter la direction de régularité est conservée mais pas la première étape consistant à redresser localement l'image. Au contraire, ici c'est la transformée qui s'adapte à l'image. Une

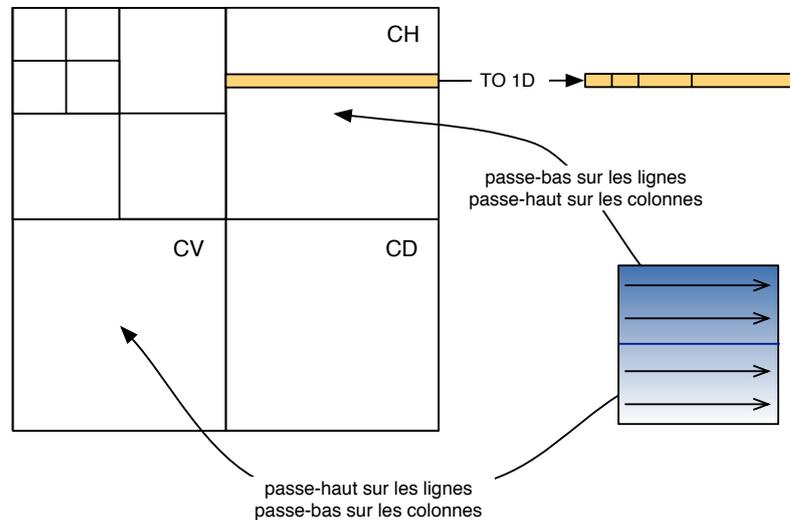


FIGURE 1.9 – Processus de bandelettisation pour un bloc d'une image redressé dont la direction de régularité est maintenant horizontale. Une transformée en ondelettes 2D est appliquée sur le bloc, puis une transformée en ondelettes 1D est appliquée sur chaque ligne de la sous-bande des détails horizontaux afin d'exploiter la régularité le long de ces lignes.

transformée en ondelettes 2D est appliquée sur une image, puis chaque sous-bande est divisée en blocs. Les éléments de chaque bloc sont réorganisés en un signal 1D. En pratique ce signal 1D correspond à la transformée de Radon discrète selon une direction  $(p, q)$  telle que le nombre d'éléments de la projection (nombre de bins) est égal au nombre d'éléments dans le bloc. La direction de projection est choisie pour que la représentation en ondelettes 1D de la projection soit la plus intéressante possible.

Pour la majorité des autres transformations adaptatives, c'est aussi la transformée en ondelettes qui est adaptée au contenu de l'image. Ces transformations s'appuient sur la flexibilité de l'implantation par *lifting* de la transformée en ondelettes [DS98] pour définir différentes transformées en ondelettes directionnelles. Le chapitre qui suit est consacré à l'étude de telles transformées.

## Conclusion

La transformée en ondelettes pour la représentation d'images mérite d'être améliorée pour représenter efficacement les structures non horizontales et non verticales, et de façon anisotrope. De nouvelles transformées ont été développées pour permettre une telle analyse directionnelle et anisotrope des images. Les transformations dites fixes permettent une décomposition sur un ensemble de sous-bandes plus importants que la transformée en ondelettes au prix généralement d'une certaine redondance. Les transformées dites adaptatives adaptent soit l'image à la transformée en ondelettes 2D, soit

la transformée en ondelettes au contenu de l'image. Parmi ces transformées, nous avons choisi de nous intéresser aux transformées directionnelles implantées par techniques de *lifting*. L'objet du chapitre suivant porte sur l'implantation de la transformée en ondelettes 1D par les techniques de *lifting*, la définition de transformées orientées par ces techniques, et leur étude comparative.

## Chapitre 2

# Transformation 2D en ondelettes orientées par redressement

### Introduction

Avant de commencer ce chapitre, un choix linguistique se pose à l'auteur. La technique du « lifting » est au cœur des travaux ici présentés, mais ne connaît pas de traduction française, en tout cas pas une suffisamment répandue pour être arrivée jusqu'à ses oreilles. La question se pose donc soit d'utiliser le terme anglais tel quel comme il est devenu courant de le trouver ou alors de le substituer par une traduction la plus adéquate possible. Pour des raisons aussi diverses que variées, dans ce manuscrit, nous ne parlerons pas de « lifting », mais de « redressement ». Redressement en raison d'une traduction littérale possible du terme anglais, et aussi car il traduit le « redressement » du signal sous-échantillonné lors de l'étape de mise à jour.

Considérant de prime abord le cas unidimensionnel, toute transformée en ondelettes peut être implantée par une technique de redressement [DS98], l'extension au cas bidimensionnel se faisant classiquement par l'application de deux transformées successives : l'une selon les lignes, puis selon les colonnes sur le résultat de la première transformation. L'adaptabilité du redressement permet toutefois de définir des extensions bidimensionnelles correspondant à ce que nous nommerons les « transformées en ondelettes orientées ».

Différentes implantations pour de telles transformées orientées sont toutefois possibles, toutes essayant de s'adapter au mieux pour permettre un filtrage selon une direction donnée. Diverses versions ont été présentées dans la littérature que nous résumerons ici en proposant une technique la plus globale possible. Ces différentes implantations de transformées en ondelettes orientées seront alors comparées d'un point de vue statistique, et à partir d'exemples précis, sans tenir compte d'un contexte de codage d'image afin de présenter au lecteur les caractéristiques inhérentes à chacune.

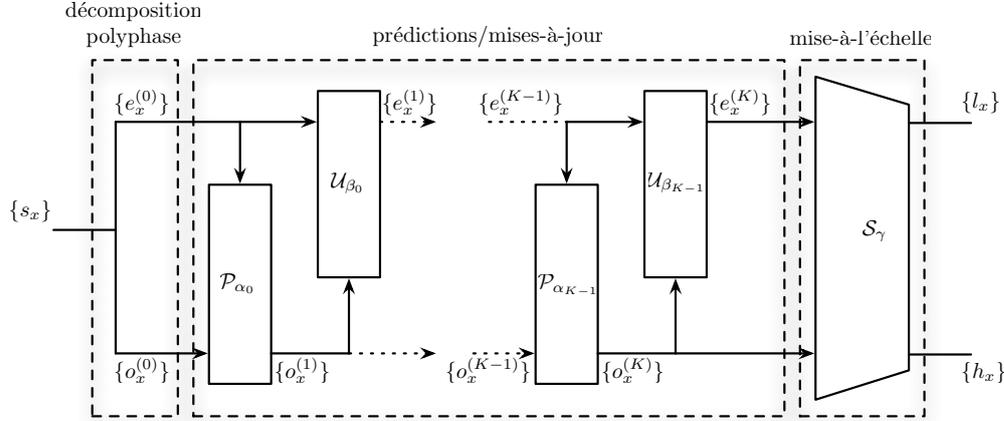


FIGURE 2.1 – Les étapes de la technique de redressement d'un signal unidimensionnel.

## 2.1 Technique de redressement pour l'implantation de la transformée en ondelettes 1D

La technique de redressement d'une signal  $s$  peut être décomposée en trois étapes successives comme illustré sur la figure 2.1. La première, consiste simplement en une séparation des échantillons de numéros pairs  $e$  (*even*) et impairs  $o$  (*odd*). Cette opération, notée *Split* est définie ainsi

$$\textit{Split} : \{s_x\}_{x \in \mathbb{Z}} \mapsto \{e_x^{(0)}\}_{x \in \mathbb{Z}}, \{o_x^{(0)}\}_{x \in \mathbb{Z}} \text{ avec } \begin{cases} e_x^{(0)} = s_{2x} \\ o_x^{(0)} = s_{2x+1} \end{cases} \quad (2.1)$$

Cette étape est appelée ironiquement en anglais « lazy transform ». L'opération inverse, *Merge*, est elle définie par :

$$\textit{Merge} : \{e'_x\}_{x \in \mathbb{Z}}, \{o'_x\}_{x \in \mathbb{Z}} \mapsto \{s'_x\}_{x \in \mathbb{Z}} \text{ avec } \begin{cases} s'_{2x} = e'_x \\ s'_{2x+1} = o'_x \end{cases} \quad (2.2)$$

Cette première étape est suivie d'une succession de couples d'étapes de prédiction et de mise à jour. À un étage  $k$  donné, l'étape de prédiction modifie les échantillons impairs à l'aide d'un prédicteur utilisant les échantillons pairs originaux. Ces derniers sont quant à eux corrigés lors de l'étape de mise à jour à partir des échantillons impairs précédemment modifiés. Le nombre d'étapes nécessaires pour réaliser une transformée en ondelettes dépend directement des filtres passe-haut et passe-bas définissant une telle transformée dans l'hypothèse où les filtres de prédiction et de mise à jour sont de longueur 2. Notons  $\mathcal{P}_\alpha$  l'opération de prédiction paramétrée par le coefficient  $\alpha$ ,  $\alpha \in \mathbb{R}^*$ , et  $\mathcal{U}_\beta$  le prédicteur de l'échantillon  $o_x$  associé,  $\mathcal{P}_\alpha$  est définie par :

$$\mathcal{P}_\alpha : \{e_x^{(k)}\}_{x \in \mathbb{Z}}, \{o_x^{(k)}\}_{x \in \mathbb{Z}} \mapsto \{e_x^{(k)}\}_{x \in \mathbb{Z}}, \{o_x^{(k+1)}\}_{x \in \mathbb{Z}} \quad (2.3)$$

avec :

$$o_x^{(k+1)} = o_x^{(k)} + \alpha \cdot P_x(\{e_{x'}^{(k)}\}) \quad (2.4)$$

L'opération de mise à jour est elle définie par :

$$\mathcal{U}_\beta : \{e_x^{(k)}\}_{x \in \mathbb{Z}}, \{o_x^{(k+1)}\}_{x \in \mathbb{Z}} \mapsto \{e_x^{(k+1)}\}_{x \in \mathbb{Z}}, \{o_x^{(k+1)}\}_{x \in \mathbb{Z}} \quad (2.5)$$

avec :

$$e_x^{(k+1)} = e_x^{(k)} + \beta \cdot U_x(\{o_{x'}^{(k+1)}\}) \quad (2.6)$$

où  $\beta \in \mathbb{R}^*$ .

Les étapes de prédiction et de mise à jour sont trivialement inversibles, et dans le cas où  $P$  et  $U$  sont des opérateurs linéaires :

$$\mathcal{P}_\alpha^{-1} = \mathcal{P}_{-\alpha} \quad (2.7)$$

$$\mathcal{U}_\beta^{-1} = \mathcal{U}_{-\beta} \quad (2.8)$$

La troisième et dernière étape est constituée de l'opération de mise-à-l'échelle  $\mathcal{S}_\gamma$  :

$$\mathcal{S}_\gamma : \{e_x^K\}_{x \in \mathbb{Z}}, \{o_x^K\}_{x \in \mathbb{Z}} \mapsto \{l_x\}_{x \in \mathbb{Z}}, \{h_x\}_{x \in \mathbb{Z}} \quad (2.9)$$

avec :

$$\begin{cases} l_x = \gamma \cdot e_x^K \\ h_x = \frac{1}{\gamma} \cdot o_x^K \end{cases} \quad (2.10)$$

Son inverse est elle aussi triviale :

$$\mathcal{S}_\gamma^{-1} = \mathcal{S}_{\frac{1}{\gamma}} \quad (2.11)$$

où  $\gamma \in \mathbb{R}^*$ .

La transformée en ondelettes unidimensionnelle est finalement implantée par redressement selon la formule suivante :

$$\mathcal{L}_{A,B,\gamma} = \mathcal{S}_\gamma \circ \bigcirc_{k=0}^{K-1} (\mathcal{U}_{\beta_k} \circ \mathcal{P}_{\alpha_k}) \circ Split \quad (2.12)$$

et son inverse par :

$$\mathcal{L}_{A,B,\gamma}^{-1} = Merge \circ \mathcal{S}_\gamma^{-1} \circ \bigcirc_{k=K-1}^0 (\mathcal{P}_{\alpha_i}^{-1} \circ \mathcal{U}_{\beta_i}^{-1}) \quad (2.13)$$

avec :

$$A = (\alpha_k)_{0 \leq k < K} \quad (2.14)$$

$$B = (\beta_k)_{0 \leq k < K} \quad (2.15)$$

et :

$$\bigcirc_{k=0}^{K-1} \mathcal{F}_k = \mathcal{F}_{K-1} \circ \mathcal{F}_{K-2} \circ \cdots \circ \mathcal{F}_1 \circ \mathcal{F}_0 \quad (2.16)$$

L'article de Daubechies et Swelden de 1998 [DS98] décrit à la façon d'un tutoriel comment décomposer une transformation en ondelettes définie par ses deux filtres d'analyse en étapes de redressement dans le cas particulier où les filtres de prédiction et de mise à jour sont simplement définis comme une somme des deux échantillons connexes, c'est-à-dire :  $P_x : \{e_{x'}\}_{x \in \mathbb{Z}} \mapsto e_x + e_{x+1}$  et  $U_x : \{o_{x'}\}_{x \in \mathbb{Z}} \mapsto o_{x-1} + o_x$ . Dans ce cas particulier, mais aussi dans le cas plus général où  $P$  et  $U$  sont des opérateurs linéaires, les relations (2.4) et (2.6) deviennent :

$$o_x^{(k+1)} = o_x^{(k)} + \alpha \cdot (e_x^{(k)} + e_{x+1}^{(k)}) \quad (2.17)$$

$$e_x^{(k+1)} = e_x^{(k)} + \beta \cdot (o_{x-1}^{(k+1)} + o_x^{(k+1)}) \quad (2.18)$$

La technique de redressement pour l'implantation de la transformée en ondelettes a ainsi trois avantages par rapport à l'implantation par bancs de filtres :

**complexité mémoire :** le redressement permet des calculs en place, c'est-à-dire que contrairement à l'implantation par bancs de filtres, un autre emplacement mémoire n'est pas nécessaire pour effectuer la transformation ;

**complexité calculatoire :** l'application sur un signal de longueur  $N$  de deux filtres, l'un passe-bas, l'autre passe-haut, tous les deux suivis d'une décimation d'ordre deux a un coût global de  $\frac{N}{2}(L_l + L_h)$  multiplications –  $L_l$  et  $L_h$  représentant respectivement les longueurs des filtres passe-bas et passe-haut – et  $\frac{N}{2}(L_l - 1 + L_h - 1)$  additions tandis qu'une technique de redressement constituée de  $K$  étapes de prédiction/mise à jour et une étape de mise-à-l'échelle a, elle, un coût global de  $N(2K + 1)$  multiplications et  $2NK$  additions. À titre d'exemple, le tableau 2.1 compare les coûts calculatoires de l'implantation des filtres 5/3 et 9/7 ;

**adaptabilité :** les opérateurs  $P$  et  $U$  de prédiction et de mise à jour autorisent une certaine flexibilité dans le choix des échantillons utilisé tout en maintenant une reconstruction exacte. La technique de redressement peut être étendu en deux dimensions (voire plus). C'est de cette flexibilité que peut être tirée une définition d'une technique de redressement orienté.

À titre d'exemple les techniques de redressement associées aux filtres d'ondelettes 5/3 et 9/7 sont donnés à l'annexe A.

## 2.2 Adaptabilité du redressement dans l'implantation de la transformée en ondelettes 2D

La transformée en ondelettes bidimensionnelle est classiquement implantée par deux transformées unidimensionnelles appliquées successivement selon les lignes et les colonnes d'une image, ou inversement. La flexibilité du choix des échantillons dans les

type d'implantation	5/3 (1 étage)		9/7 (2 étages)	
	#×	#+	#×	#+
banc de filtres	$4N$	$3N$	$8N$	$7N$
redressement	$3N$	$2N$	$5N$	$4N$

TABLE 2.1 – Comparaison de la complexité entre une implantation par banc de filtres et une implantation avec le schéma de redressement associé.

étapes de prédiction et de mise à jour d'un schéma de redressement permet d'envisager toutes sortes d'extensions bidimensionnelles exotiques. Parmi celles-ci, ce sont les définitions de transformées en ondelettes orientées implantées par redressement qui attirent notre attention. De tels schémas de redressement s'appuient sur une certaine décomposition polyphase et sur des choix techniques d'implantations des orientations de filtrage considérées.

Des extensions bidimensionnelles s'appuyant sur une décomposition polyphase basée sur un échantillonnage régulier seront tout d'abord présentées à travers le concept de redressement basé ligne et redressement basé colonne afin de fixer les esprits sur la flexibilité permise par les méthodes par redressement. Le concept de transformée orientée sera ensuite élargie à tous types d'échantillonnages. Pour cela des règles concernant le choix des échantillons seront alors énoncées. Les diverses transformées orientées connues à l'heure actuelle seront finalement détaillées et comparées.

### 2.2.1 Redressement basé ligne et redressement basé colonne

La transformée en ondelettes bidimensionnelle étant classiquement implantée par deux transformées unidimensionnelles appliquées successivement selon les lignes et les colonnes d'une image, ou inversement, considérons le cas où le premier filtrage se fait selon les lignes de l'image. Un échantillon impair est prédit à partir de son voisin de gauche, d'indice pair, et de son voisin de droite, d'indice pair lui aussi. De façon générale, comme cela a été défini plus haut via le schéma de redressement d'un signal unidimensionnel, cet échantillon peut même être prédit à partir de n'importe quels échantillons pairs issus de la même ligne auquel il appartient. L'ensemble des échantillons pouvant servir à la prédiction d'un échantillon d'indice impair peut être étendu dans le cas bidimensionnel à tous les échantillons d'indice pair quelles que soient les lignes auxquelles ils appartiennent. En particulier, au lieu de prédire un échantillon impair uniquement à partir de son voisin de gauche et de son voisin de droite, la prédiction peut utiliser n'importe quels éléments de la colonne de gauche, et n'importe quels éléments de la colonne de droite. Pour généraliser, les éléments d'une colonne impaire peuvent être prédits à partir des éléments des colonnes paires. Le schéma de redressement associé est alors dit « basé colonne », il est illustré sur la figure 2.2.

Cette définition d'un redressement basé colonne ouvre la porte à toutes sortes d'implantations parmi lesquelles nous relèverons trois variantes :

**8-connexe-( $p, q$ )-redressement** : filtrage le long d'une droite discrète 8-connexe définie par le vecteur  $(p, q)$  (figure 2.3-a) ;

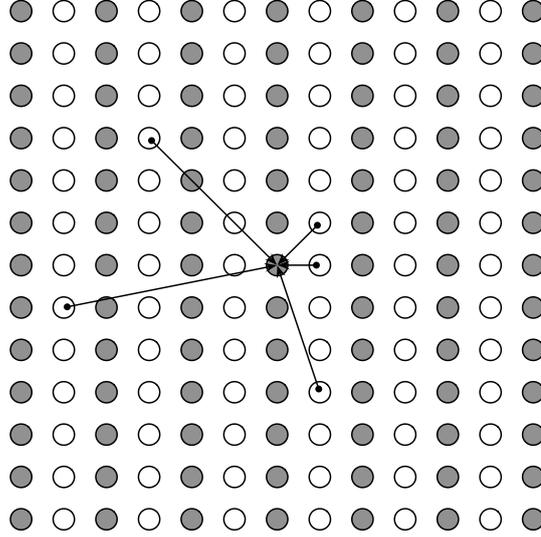


FIGURE 2.2 – Exemple, dans le cas d’un schéma de redressement basé colonne, de la prédiction d’un échantillon d’une colonne impaire (en gris) à partir de divers échantillons de colonnes paires (en blanc).

**$(, p, q)$ -connexe- $(p, q)$ -redressement** : prédiction et mise à jour utilisent des échantillons situés à  $(p, q)$  échantillons du point à prédire ou à mettre à jour (figure 2.3-b) ;

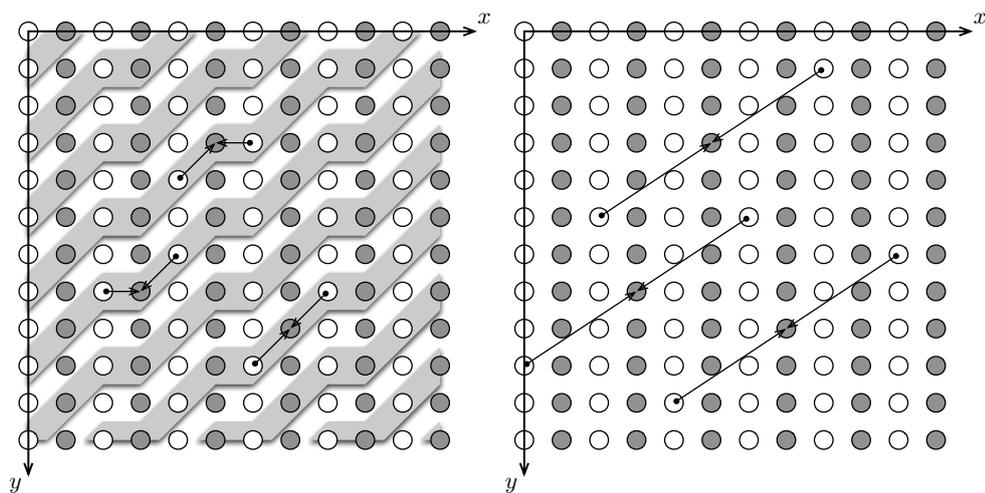
**$\theta$ -redressement** : les valeurs utilisées sont issues d’une interpolation (figure 2.3-c).

### 2.2.1.1 8-connexe- $(p, q)$ -redressement, ou Réveillès- $(p, q)$ -redressement

Dans cette première variante [JRB07a, VBLVD06], illustrée figure 2.3-a, l’image est partitionnée en lignes discrètes 8-connexes [Rev91]. L’orientation de ces lignes est définie par un vecteur  $(p, q)$  avec  $p$  et  $q$  deux entiers premiers entre eux, et les lignes elles-mêmes par :

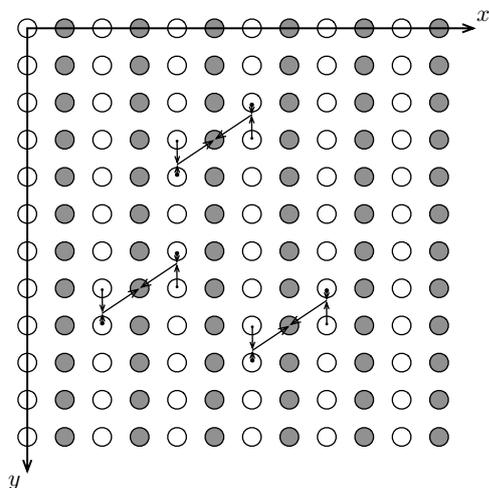
$$0 \leq -q \cdot x + p \cdot y + C < \max(|p|, |q|) \quad (2.19)$$

où  $(x, y)$  appartient à la ligne correspondant à la constante  $C$ . Un échantillon d’une colonne impaire est alors prédit à partir d’échantillons de colonnes paires appartenant à la même ligne discrète. En pratique, la valeur  $\left\lfloor \frac{-q \cdot x + p \cdot y}{\max(|p|, |q|)} \right\rfloor$  correspondant à un indice de ligne est calculée pour chaque point de l’image. Le pixel  $(x_c, y_c)$  d’indice de ligne  $\left\lfloor \frac{-q \cdot x_c + p \cdot y_c}{\max(|p|, |q|)} \right\rfloor$  doit donc être prédit en utilisant des échantillons ayant le même indice de ligne. La position de ces échantillons n’étant pas définie analytiquement, un parcours du tableau des indices de ligne est nécessaire pour la connaître. Cette recherche est limitée au test de six valeurs pour un filtre de prédiction n’utilisant que les deux voisins



(a) Réveillès- $(p, q)$ -redressement

(b) Mojette- $(p, q)$ -redressement



(c)  $\theta$ -redressement

FIGURE 2.3 – Exemples d'implantations pour l'étape de prédiction d'un schéma de redressement basé colonne selon une orientation  $(p, q)$  de filtrage d'environ  $33^\circ$  par rapport à l'horizontale.

connexes. Le schéma de redressement est alors rigoureusement identique à une transformée en ondelettes implantée par banc de filtres appliquée selon les droites discrètes 8-connexes définies par la relation (2.19). Un filtre de prédiction linéaire utilisé dans la relation (2.4) pour le Réveillès- $(p, q)$ -redressement aura la forme :

$$P_{x,y} : \{e_{x',y'}\}_{(x',y') \in \mathbb{Z}^2} \mapsto \sum_{n=-\frac{N_p}{2}+1}^{\frac{N_p}{2}} P[n] \cdot e_{x+n,y_n} \quad (2.20)$$

où  $y$  et  $y_n$  vérifient la relation :

$$\left\lfloor \frac{-q \cdot x + p \cdot y}{\max(|p|, |q|)} \right\rfloor = \left\lfloor \frac{-q \cdot (x+n) + p \cdot y_n}{\max(|p|, |q|)} \right\rfloor \quad (2.21)$$

### 2.2.1.2 $(p, q)$ -connexe- $(p, q)$ -redressement, ou Mojette- $(p, q)$ -redressement

L'orientation de filtrage dans cette deuxième version [VBLVD06, CG07] est, elle aussi, définie par un vecteur  $(p, q)$ , sauf que cette fois-ci les droites discrètes définies par ce vecteur ne sont pas 8-connexes, mais  $(p, q)$  connexes [Gué01] avec la contrainte supplémentaire que  $p$  soit impair de telle sorte qu'un échantillon sur deux de la droite discrète ainsi définie soit pair. Contrairement au précédent schéma de redressement, la position des échantillons utilisés pour la prédiction est directement connue à partir du vecteur  $(p, q)$ . Un filtre de prédiction linéaire aura donc ici la forme suivante :

$$P_{x,y} : \{e_{x',y'}\}_{(x',y') \in \mathbb{Z}^2} \mapsto \sum_{n=-\frac{N_p}{2}+1}^{\frac{N_p}{2}} P[n] \cdot e_{\frac{2x+1+(2n-1) \cdot p+1}{2}, y+(2n-1) \cdot q} \quad (2.22)$$

De la même façon que précédemment, ce schéma de redressement est rigoureusement identique à une transformée en ondelettes implantée par banc de filtres et appliquée sur réseau régulier de points défini par le vecteur  $(p, q)$ .

### 2.2.1.3 $\theta$ -redressement

Contrairement aux deux versions précédentes, pour le  $\theta$ -redressement [WZS06, JRB07a, DWW<sup>+</sup>07], les valeurs utilisées pour la prédiction peuvent être interpolées. Dans l'article de Ding *et al.* [DWW<sup>+</sup>07], un interpolateur de type sinus cardinal est utilisé. Sur la figure 2.3-c, le  $\theta$ -redressement est illustré avec un simple interpolateur linéaire du premier ordre. De la même façon que pour la version Mojette, la position des échantillons utilisés pour la prédiction est entièrement connue lorsque l'interpolateur utilisé est connu. À titre d'exemple, pour le cas de l'interpolateur linéaire du premier ordre, un filtre de prédiction linéaire aura la forme :

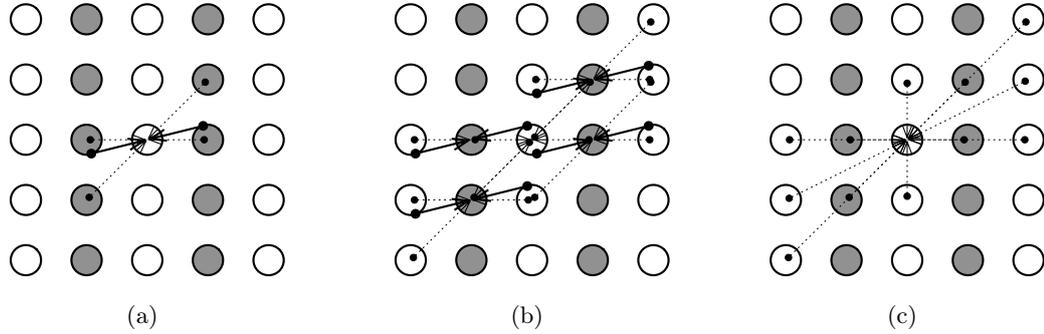


FIGURE 2.4 – Illustration de l'effet « ricochet » dans l'implantation d'une transformation en ondelettes orientées par un schéma de  $\theta$ -redressement basé sur une transformation en ondelettes 5/3. Sur la figure (a), l'échantillon central (pair) est mis-à-jour d'après des valeurs interpolées à partir de quatre échantillons impairs. Ces échantillons impairs eux-mêmes correspondant à des erreurs de prédiction calculées d'après des valeurs interpolées à partir d'échantillons pairs (b). La figure (c) rend finalement compte à partir de quels échantillons un élément pair est réellement filtré.

$$P_{x,y} : \{e_{x',y'}\}_{(x',y') \in \mathbb{Z}^2} \mapsto \sum_{n=-\frac{N_p}{2}+1}^{\frac{N_p}{2}} P[n] \cdot (e_{x+n,y+(2n-1)\cdot \tan \theta}) \quad (2.23)$$

soit :

$$P_{x,y} : \{e_{x',y'}\}_{(x',y') \in \mathbb{Z}^2} \mapsto \sum_{n=-\frac{N_p}{2}+1}^{\frac{N_p}{2}} P[n] \cdot (([t] - t) \cdot e_{x+n,[y+t]} + (t - [t]) \cdot e_{x+n,[t]}) \quad (2.24)$$

avec :

$$t = (2n - 1) \cdot \tan \theta$$

Contrairement aux deux précédentes variantes, le schéma de redressement présenté ici ne correspond pas à la transformée en ondelettes à laquelle on pourrait penser intuitivement : une transformée où pour filtrer un échantillon  $(x, y)$  les échantillons utilisés seraient prélevés sur une droite de coefficient directeur déterminé par l'angle  $\theta$  passant par  $(x, y)$ . Ici les interpolations successives pour les étapes de prédiction et de mise à jour entraînent ce qu'on pourrait appeler un effet « ricochet » qui conduit à un filtrage différent de celui attendu intuitivement. Cet effet « ricochet » est illustré sur la figure 2.4.

Les filtres d'ondelettes orientées équivalents à la technique de  $\theta$ -redressement,  $0 \leq \theta \leq 45^\circ$ , pour les filtres 5/3 et 9/7 sont donnés à l'annexe B.

Au regard de ces exemples de filtres, on constate que contrairement aux deux autres variantes de redressement orienté, l'énergie de la réponse impulsionnelle de chaque filtre varie en fonction de l'orientation du redressement. Cette constatation amène donc à une première conclusion concernant cette variante du redressement orienté : les énergies en sortie des filtres des sous-bandes produites pour différentes orientations ne sont pas comparables entre elles.

## 2.2.2 Formalisme général basé sur la parité

Dans le paragraphe précédent a été formalisé ce que nous avons nommé le redressement basé colonne et le redressement basé ligne. Ces deux schémas représentent une sous-partie de l'ensemble des schémas de redressement réalisant une transformée en ondelettes orientée. Trois variantes d'implantation ont été présentées pour un schéma basé colonne en précisant l'opération de prédiction utilisée. Des schémas basés ligne correspondants peuvent être construits par analogie. Dans ce paragraphe, le formalisme précédent sera généralisé pour une décomposition polyphase quelconque. Une seule et unique direction n'est pas utilisée pour le filtrage d'une image donnée, mais un certain nombre. L'opération de mise à jour sera précisée au regard de cette possibilité. Une transformée en ondelettes bidimensionnelle fait bien évidemment intervenir deux filtrages successifs qui seront précisés l'un par rapport à l'autre.

### 2.2.2.1 Décomposition polyphase

Dans les implantations qui ont été jusqu'ici présentées, la décomposition polyphase employée est basée sur un échantillonnage régulier. Selon une première direction, par exemple, un redressement basé colonne est appliqué, ce qui sous-entend un échantillonnage préalable des colonnes. Ensuite, selon la seconde direction, un redressement basé ligne est appliqué sur l'ensemble des colonnes paires d'une part et l'ensemble des colonnes impaires d'autre part. Ceci implique donc un échantillonnage préalable des lignes. Au final les échantillons d'une image sont donc, après décomposition, séparés en quatre sous-ensembles correspondant chacun à une sous-bande en fonction de la parité des échantillons au regard des deux filtrages successifs opérés. La sous-bande contenant l'information des basses fréquences selon les deux directions de filtrage contiendra les échantillons « pairs-pairs », c'est-à-dire provenant d'une colonne paire par rapport un premier filtrage basé colonne, et d'une ligne paire par rapport au second filtrage basé ligne. Celle contenant l'information des basses fréquences selon la première direction, et hautes fréquences selon la seconde contiendra les échantillons « pairs-impairs ». L'information des basses fréquences selon la première direction, et hautes fréquences selon la seconde contiendra les échantillons « impairs-pairs ». Enfin, les échantillons « impairs-impairs » représenteront l'information des hautes fréquences dans les deux directions de filtrage.

Cet « étiquetage » de la parité des échantillons dépend de l'ordre des deux types de redressement appliqué : les échantillons n'auront pas la même parité suivant qu'un redressement basé colonne ou basé ligne aura été appliqué en premier. Pour des raisons

techniques qui n'apparaîtront clairement que plus tard, mais aussi pour avoir bien en tête les notions géométriques mises en œuvre dans le filtrage bidimensionnel opéré par redressement, les échantillons ne seront pas réellement séparés préalablement à chacun des deux schémas de redressement, mais uniquement une fois ces deux schémas opérés. La décomposition polyphase préalable à une technique de redressement est alors vue comme un simple étiquetage.

Ce point de vue permet d'englober dans ce formalisme les ondelettes orientées, basées sur un échantillonnage quinconce[Cha05, Gou02].

Ce qu'il convient de garder en tête lors de l'implantation d'un quelconque schéma de redressement est la contrainte de parité suivante :

- pour le filtrage selon la première direction :
  - un élément « impair » ne peut être prédit qu'à partir d'éléments « pairs » ;
  - un élément « pair » ne peut être mis-à-jour qu'à partir d'éléments « impairs » ;
- pour le filtrage selon la seconde direction :
  - un élément « pair-impair » ne peut être prédit qu'à partir d'éléments « pairs-pairs » ;
  - un élément « pair-pair » ne peut être prédit qu'à partir d'éléments « pairs-impairs » ;
  - un élément « impair-impair » ne peut être prédit qu'à partir d'éléments « impairs-pairs » ;
  - un élément « impair-pair » ne peut être prédit qu'à partir d'éléments « impairs-impairs ».

### **2.2.2.2 Conclusion sur les orientations de filtrage**

Il ne suffit pas de parler de filtrage selon une orientation donnée pour caractériser un tel filtrage. Comme cela a été présenté plus haut, un filtrage orienté peut-être implanté de diverses manières pour une même orientation. Exemple évident : le Réveillès- $(p, q)$ -redressement et le Mojette- $(p, q)$ -redressement. Exemple plus complexe : deux schémas de  $\theta$ -redressement, pour la même valeur de  $\theta$ , mais l'un basé ligne, l'autre basé colonne (voir figure 2.5). Une orientation de filtrage n'est donc pas seulement définie par un angle  $\theta$  ou un vecteur  $(p, q)$ , mais aussi par la façon dont, en pratique, le filtrage s'opère, soit par choix technique comme dans le premier exemple, soit par ce que permet ou ne permet pas la décomposition polyphase utilisée, comme dans le deuxième exemple. La valeur même des orientations de filtrage possibles est tributaire de l'implantation. Par exemple, sans utiliser d'interpolation, on ne peut pas mettre en place de Mojette- $(0, 1)$ -redressement basé colonne ou d'ondelettes orientées de Chappelier à  $45^\circ$ [CG05] pour le premier niveau. Le filtrage orienté dépend donc et de l'orientation choisie, et (compte tenu de ce qui est possible de faire étant données les contraintes de parité) des choix d'implantation.

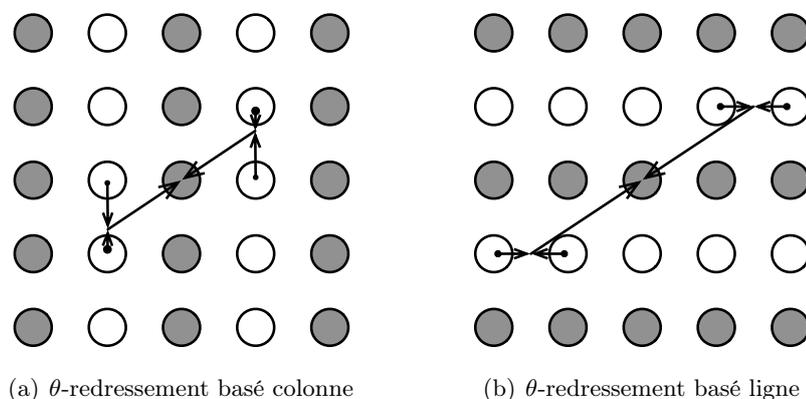


FIGURE 2.5 – Deux implantations du  $\theta$ -redressement pour un même  $\theta$  mais chacune basée sur une décomposition polyphase différente.

Qui dit filtrage orienté bidimensionnel, dit deux directions de filtrage, respectant chacune les conditions de parité établies plus haut. Ces deux directions peuvent être indépendantes, par exemple chacune cherchant à minimiser l'énergie des sous-bandes résultantes, ou elles peuvent être liées : en général, la seconde étant soit la direction orthogonale à la première (dans la mesure du possible), soit une direction horizontale ou verticale en fonction de la première.

### 2.3 Étude comparative des diverses implantations du redressement orienté

Les directions de filtrage doivent être idéalement choisies selon un compromis débit-distorsion dans un contexte de codage. En pratique on s'intéresse souvent aux directions qui vont minimiser l'énergie dans les sous-bandes des hautes fréquences. Dans cette section, les trois variantes de techniques de redressement présentées plus haut (voir sous-section 2.2.1) sont donc comparées entre elles du point de vue de leur capacité à minimiser une telle énergie. Cette étude s'intéresse uniquement au premier filtrage orienté de la transformée complète qui compte deux filtrages. En pratique une technique de redressement basé ligne ou basé colonne est donc appliquée sur un certain contenu pour produire une sous-bande des basses fréquences et une sous-bande des hautes fréquences.

Pour qu'une comparaison en termes d'énergie soit possible et concerne différentes sous-bandes produites après différentes techniques de redressement, la dynamique des signaux doit être la même.

Or, comme indiqué dans la sous-section 2.2.1.3, la dynamique du signal de cette sous-bande varie en fonction de l'orientation de filtrage pour cette technique de redressement, et est amenée à varier selon la technique de filtrage employée. 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 de minimiser l'énergie de cette image reconstruite pour la comparaison des différentes techniques de redressement.

Les trois variantes des techniques de redressement testées ici utilisent les filtres d'ondelettes 5/3 ou 9/7 avec les filtres de prédiction/mise à jour de longueur 1 ou 3. Pour chacune de ces variantes un ensemble assez restreint d'orientations de filtrage est considéré. Pour le  $\theta$ -redressement cet ensemble correspond à un échantillonnage des orientations. 16 orientations ont ici été choisies. L'échantillonnage en question peut être uniforme auquel cas l'ensemble des orientations pour un redressement basé colonne (BC) sera :

$$\{-45^\circ; -33,75^\circ; -22,5^\circ; -11,25^\circ; 0^\circ; 11,25^\circ; 22,5^\circ; 33,75^\circ; 45^\circ\},$$

et l'ensemble des orientations pour un redressement basé ligne (BL) sera :

$$\{45^\circ; 56,25^\circ; 67,5^\circ; 78,75^\circ; 90^\circ; 101,25^\circ; 112,5^\circ; 123,75^\circ; 135^\circ\}.$$

L'échantillonnage est non uniforme s'il est basé sur la notion d'interpolation au demi-pixel ou au quart-de-pixel, et donc dans ce cas l'ensemble des orientations pour un redressement basé colonne (BC) sera approximativement :

$$\{-45^\circ; -36,87^\circ; -26,57^\circ; -14,04^\circ; 0^\circ; 14,04^\circ; 26,57^\circ; 36,87^\circ; 45^\circ\}$$

et l'ensemble des orientations pour un redressement basé ligne (BL) sera approximativement :

$$\{45^\circ; 53,13^\circ; 63,43^\circ; 75,96^\circ; 90^\circ; 104,04^\circ; 116,57^\circ; 126,87^\circ; 135^\circ\}.$$

Ces deux échantillonnages particuliers sont résumés sur la figure 2.6. Pour les deux autres variantes où les orientations de filtrage sont basées sur des vecteurs  $(p, q)$ , l'ensemble des orientations est défini par la suite de Farey d'ordre 3 (voir figure 2.7). L'ensemble des orientations pour un redressement basé colonne (BC) sera donc approximativement :

$$\{-71,57^\circ; -63,43^\circ; -45^\circ; -33,69^\circ; -18,43^\circ; 0^\circ; 18,43^\circ; 33,69^\circ; 45^\circ; 63,43^\circ; 71,57^\circ\},$$

et celui pour un redressement basé ligne (BL) :

$$\{18,43^\circ; 26,57^\circ; 45^\circ; 56,31^\circ; 71,57^\circ; 90^\circ; 108,43^\circ; 123,69^\circ; 135^\circ; 153,43^\circ; 161,57^\circ\}.$$

La première étude portera d'abord sur l'analyse de signaux 2D correspondant à des structures de type contours rectilignes puis à celui de contours courbes. Enfin des comparaisons seront faites sur des contenus naturels plus complexes.

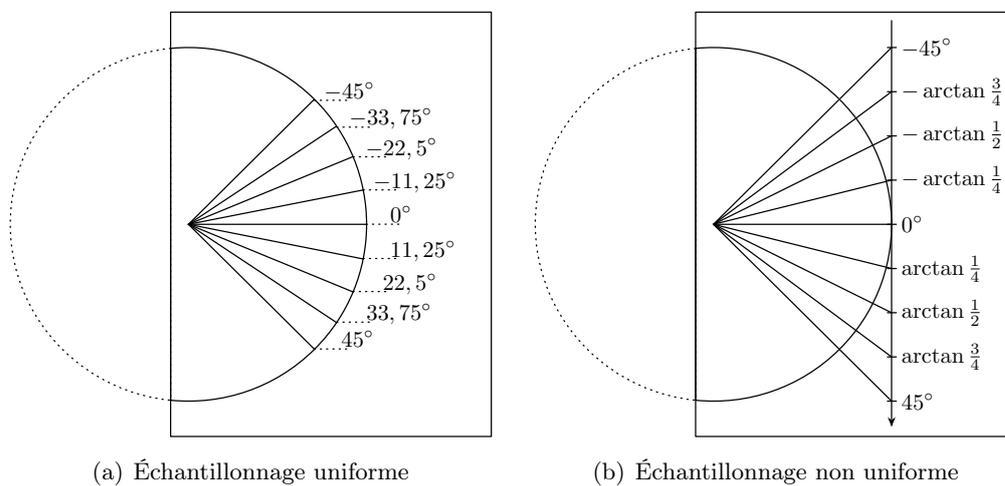


FIGURE 2.6 – Échantillonnages possibles des orientations utilisées pour le  $\theta$ -redressement.

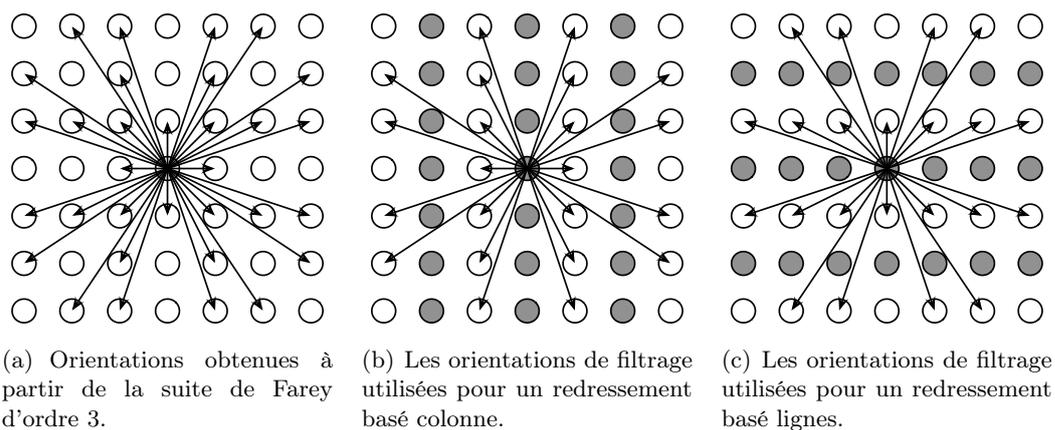


FIGURE 2.7 – Les orientations utilisées pour les techniques de  $(p, q)$ -redressement basé ligne et basé colonne à partir de la série de Farey d'ordre 3.

### 2.3.1 Contenu rectiligne

Dans ce paragraphe, la capacité des techniques de redressement présentées précédemment, à minimiser l'énergie de la sous-bande des hautes fréquences dans la direction de filtrage est étudiée d'abord dans le cas assez favorable où le contenu testé est rectiligne. Pour ce faire, des images de tests de dimension  $N \times N$ , avec  $N = 64$ , sont générées à partir de quatre types de contenus rectilignes différents :

1. la droite discrète d'orientation  $(p_{ech}, q_{ech})$  définie par la relation (2.19), où  $C$  est telle que la droite passe par le pixel  $(x_c, y_c) = (\frac{N-1}{2}, \frac{N-1}{2})$  et se situe à la frontière de deux régions uniformes. Les pixels  $(x, y)$  vérifiant  $\frac{-q_{ech} \cdot (x-x_c) + p_{ech} \cdot (y-y_c)}{\max(|p_{ech}|, |q_{ech}|)} \geq 0$  appartiennent à la région de valeur 64, les autres à la région de valeur 192. La droite frontière appartient donc à la première région. Ce choix d'appartenance à une région plutôt que, par exemple, lui donner la valeur 128 a été fait pour la raison suivante : dans le cas d'un filtrage par ondelettes 5/3, pour une orientation de  $0^\circ$ , la prédiction d'un pixel de valeur 128 serait correctement prédit, accidentellement en quelque sorte, par la moyenne des valeurs du pixel situé à sa gauche et de celui situé à sa droite ( $128 = \frac{192+64}{2}$ ) ce qui avantagerait cette orientation plus qu'il n'est nécessaire ;
2. un échelon idéal, d'amplitude 128 centré sur la valeur 128, orienté selon  $\theta_{ech}$  est idéalement échantillonné ;
3. ce même échelon orienté selon  $\theta_{ech}$  est filtré par un passe-bas de réponse impulsionnelle constante sur le support de la taille d'un pixel carré et idéalement échantillonné ;
4. une sinusoïde orientée selon  $\theta_{ech}$  idéalement échantillonnée. Cette sinusoïde est définie par la relation :

$$\sin 2D(x, y) = \frac{M - m}{2} \cos \left( \frac{2\pi}{T} (\sin \theta_{ech} \cdot x' + \cos \theta_{ech} \cdot y') \right) + \frac{M + m}{2} \quad (2.25)$$

avec  $x' = x - \frac{N-1}{2}$ ,  $y' = y - \frac{N-1}{2}$ ,  $m = 64$ ,  $M = 192$  et  $T = 5$ .

Ces quatre contenus sont illustrés sur la figure 2.8. Pour le premier contenu, un ensemble d'images de test est généré avec  $(p_{ech}, q_{ech})$  décrivant l'ensemble des éléments de la suite de Farey d'ordre 6, *id est* : (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 trois autres contenus,  $\theta_{ech}$  varie de  $0^\circ$  à  $45^\circ$  avec un pas de  $5^\circ$ . La répartition de ces orientations est volontairement plus fine que celle des orientations de filtrage de façon à pouvoir étudier le comportement d'une transformée associée à une certaine orientation de filtrage pour un contenu d'orientation proche mais non identique.

Considérant de prime abord les filtres d'ondelettes 5/3 associés aux filtres de prédiction et de mise à jour nommés plus haut L1, les énergies de l'image après reconstruction à partir uniquement de la sous-bande des hautes fréquences en fonction de l'orientation de filtrage et de l'orientation du contenu sont données dans les figures 2.10 pour le premier contenu rectiligne, 2.11 pour le deuxième, 2.12 pour le troisième et 2.13 pour

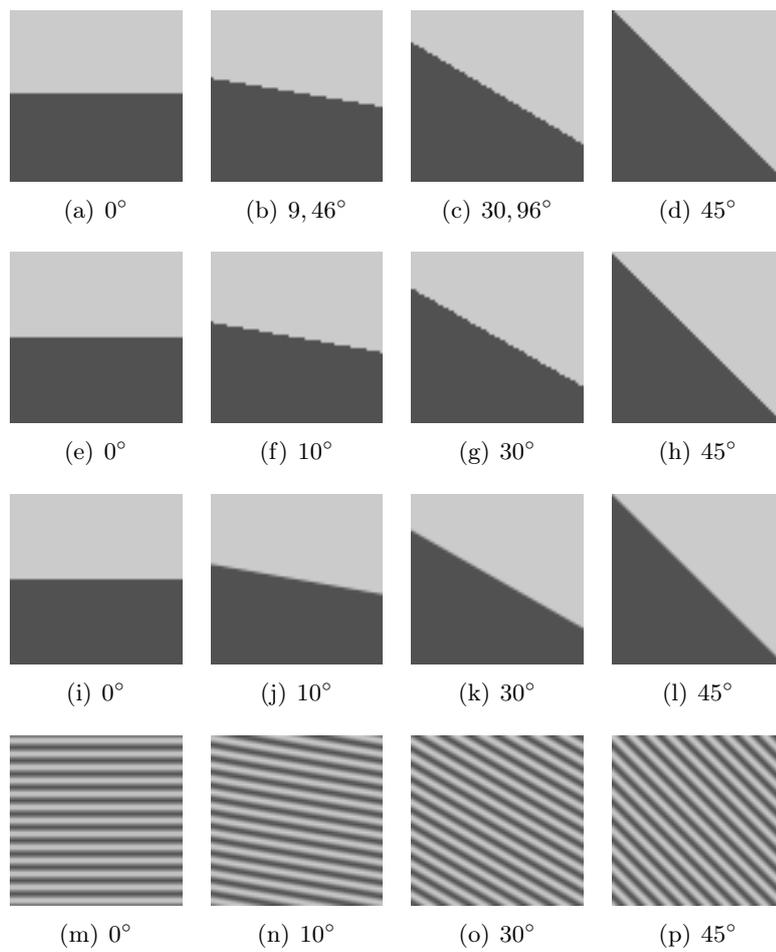


FIGURE 2.8 – Sous-ensembles des images de test associées au quatre contenus (ici, un par ligne) représentant des structures rectilignes.

le dernier. À noter que l'énergie moyenne représentée dans les graphiques suivants est calculée à partir de tous les pixels des images de test  $64 \times 64$  n'appartenant pas à une bordure large de 16 pixels pour s'affranchir complètement des effets de bord.

**Contenu 1** À première vue, le premier contenu, basé sur des droites discrètes 8-connexes, semble taillé sur mesure pour la technique de Réveillès- $(p, q)$ -redressement. En pratique, une différence de phase peut exister entre la définition des droites discrètes du contenu et les droites discrètes de filtrage.

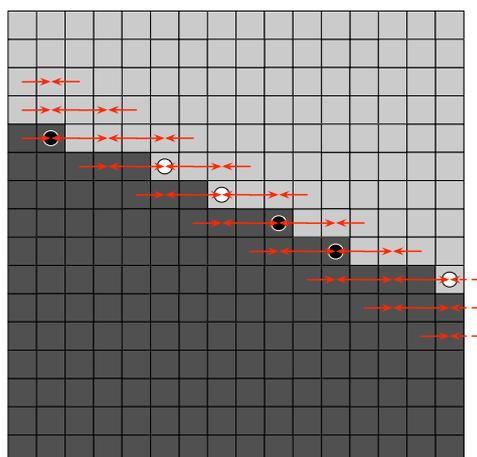
Ici, volontairement, la droite d'indice 0 définissant le contenu passe par le centre de l'image de test, alors que celle définissant le filtrage passe par le point de coordonnées  $(0, 0)$ . Ainsi même pour une orientation de filtrage parfaitement adaptée à l'orientation du contenu, la capacité à minimiser 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 2.10-a, pour une même orientation du contenu et du filtrage (i.e même  $(p, q)$ ), une énergie nulle pour une orientation de  $18, 43$  ( $(p, q) = (3, 1)$ ), une énergie non nulle pour une orientation de  $26, 57$  ( $(p, q) = (2, 1)$ ) mais inférieure aux énergies obtenues avec d'autres orientations de filtrage, et une énergie non nulle et non minimale, comparée aux autres filtrages, pour  $33, 69$  ( $(p, q) = (3, 2)$ ).

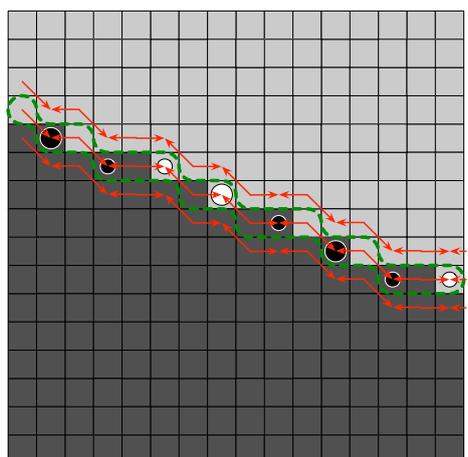
Ce dernier cas est illustré sur la figure 2.9 par une image de test de dimension  $16 \times 16$  afin d'expliquer le fait qu'un filtrage selon une orientation qui semble être la « bonne » peut être moins performant qu'un filtrage selon  $0^\circ$ . Pour un filtrage horizontal de l'image de test considérée, 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 2.9-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 2.9-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  $0^\circ$ , mais comme on l'observe sur les figures 2.9-a et 2.9-b, le nombre de pixels « mal prédits » est plus élevé pour l'orientation de filtrage  $(5, 2)$  que pour  $(1, 0)$ .

La deuxième technique, le Mojette- $(p, q)$ -redressement, profite pour ce contenu des avantages de la précédente technique sans ses inconvénients (exemple sur la figure 2.9-c). On observe donc bien sur la figure 2.10-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.

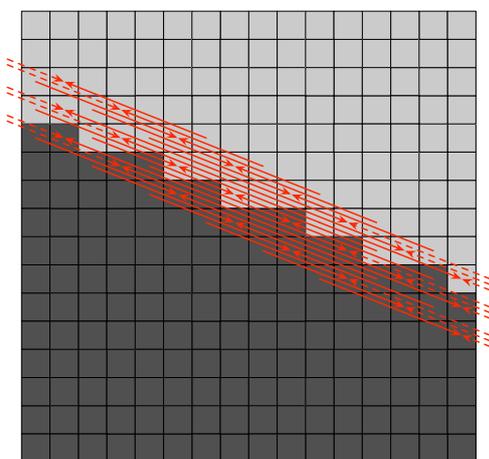
Les résultats pour la dernière technique, le  $\theta$ -redressement, sont donnés figure 2.10-c. On observe que quelle que soit l'orientation du contenu, n'importe quelle technique de redressement nécessitant une interpolation ne va pas réduire l'énergie plus qu'un filtrage à  $0^\circ, 45^\circ, 90^\circ$  ou  $135^\circ$ . 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



(a) Réveillés-(1,0)-redressement



(b) Réveillés-(5,2)-redressement



(c) Mojette-(5,2)-redressement

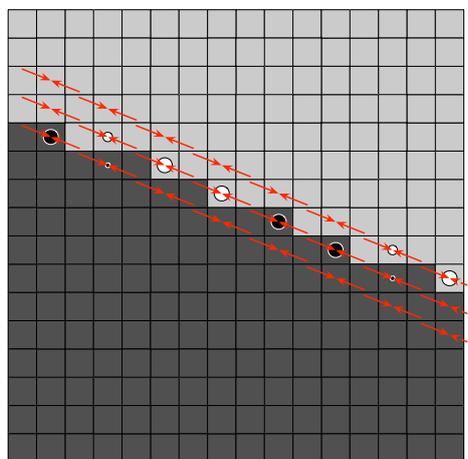
(d)  $\arctan \frac{2}{5}$ -redressement ( $\approx 21,80^\circ$ )

FIGURE 2.9 – 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.

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 2.9-a et 2.9-d, toutes les erreurs de prédiction non nulles pour un filtrage à  $0^\circ$  se retrouvent aussi pour un filtrage à environ  $21,80^\circ$ . Par contre d'autres erreurs apparaissent pour cette orientation qui n'existaient pas pour un filtrage strictement horizontal.

**2.3.1.0.1 Contenu 2** Contrairement au contenu précédent qui correspondait à un contenu spécialement généré pour correspondre avec les techniques de  $(p, q)$ -redressements, le deuxième contenu représente un échelon analogique orienté échantillonné idéalement. Les figures 2.11-a et c montrent, pour les mêmes raisons que pour le premier contenu, leur incapacité à diminuer le terme d'énergie étudié pour des orientations autres que  $0^\circ$  et  $45^\circ$ . Par contre on remarque que la technique de Mojette- $(p, q)$ -redressement reste efficace pour un contenu non spécifiquement généré pour elle.

**2.3.1.0.2 Contenu 3** Le troisième et dernier 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 2.12 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  $\theta$ -redressement.

**2.3.1.0.3 Contenu 4** Le quatrième contenu rectiligne semble à première vue être le plus favorable à la technique de  $\theta$ -redressement. Considérons l'erreur de prédiction d'un pixel  $(x, y)$  pour un filtrage selon la direction de la sinusoïde  $\theta_{ech}$  :

$$err_{\theta_{ech}}^{\theta_{ech}}(x, y) = \frac{1}{\sqrt{2}} \left[ \sin 2D(x, y) - \frac{\sin 2D(x-1, y-t) + \sin 2D(x+1, y+t)}{2} \right], \quad (2.26)$$

avec  $t = \tan \theta_{ech}$ .

On peut constater aisément que si l'interpolation était idéale, l'erreur de prédiction serait parfaitement nulle. Toutefois lorsqu'on analyse les résultats des tests menés au regard de la figure 2.13, on ne peut que constater comme pour les autres contenus que la technique de redressement qui fait intervenir une interpolation linéaire n'est pas efficace.

Comparons donc l'erreur de prédiction d'un pixel  $(x, y)$  d'un contenu orienté selon  $\theta_{ech}$  pour un filtrage selon cette orientation ( $err_{\theta_{ech}}^{\theta_{ech}}$ ),  $0^\circ$  ( $err_{0^\circ}^{\theta_{ech}}$ ) et  $45^\circ$  ( $err_{45^\circ}^{\theta_{ech}}$ ). En posant  $t = -\sin \theta_{ech} \cdot (x - x_c) + \cos \theta_{ech} \cdot (y - y_c)$ , ces erreurs de prédiction valent :

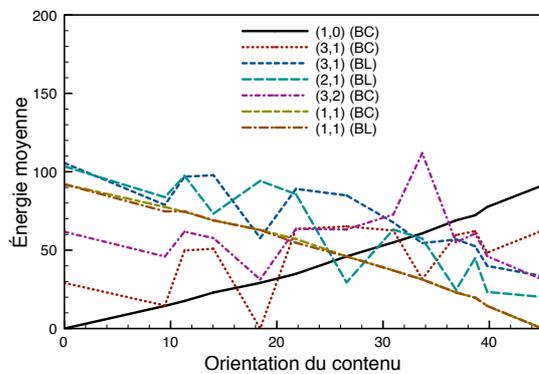
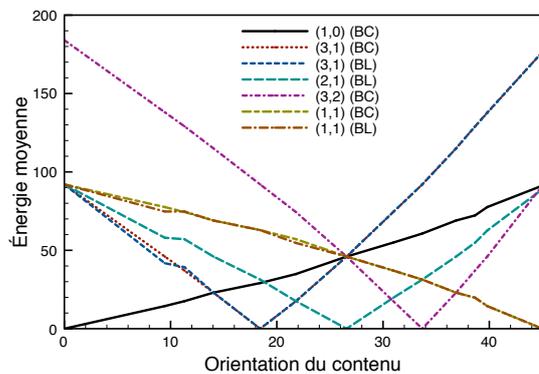
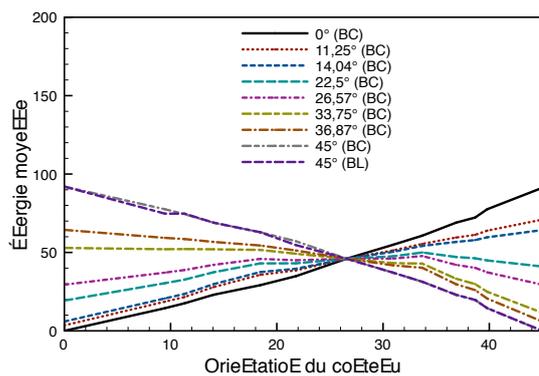
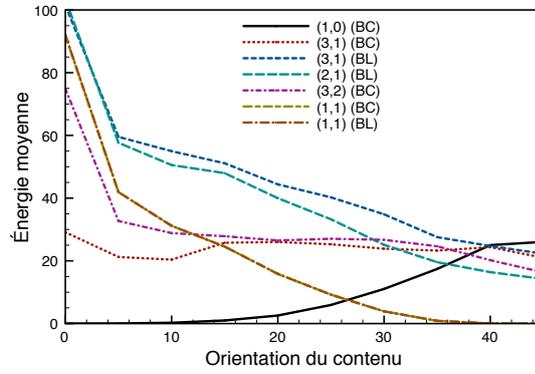
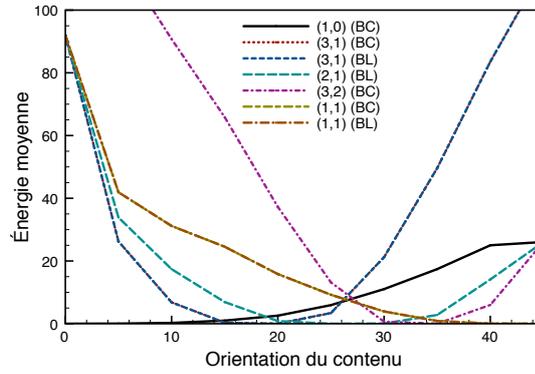
(a) Réveillès- $(p, q)$ -redressement(b) Mojette- $(p, q)$ -redressement(c)  $\theta$ -redressement

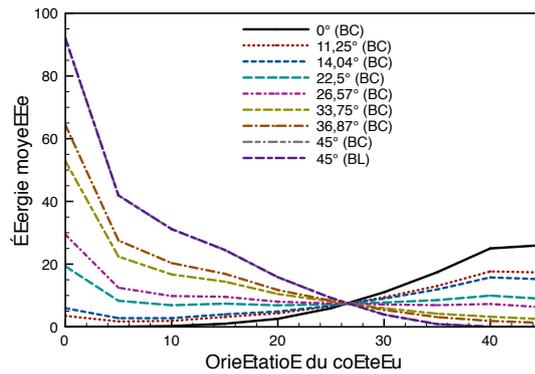
FIGURE 2.10 – Énergie des images représentant le contenu 1, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.



(a) Réveillès- $(p, q)$ -redressement



(b) Mojette- $(p, q)$ -redressement



(c)  $\theta$ -redressement

FIGURE 2.11 – Énergie des images représentant le contenu 2, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.

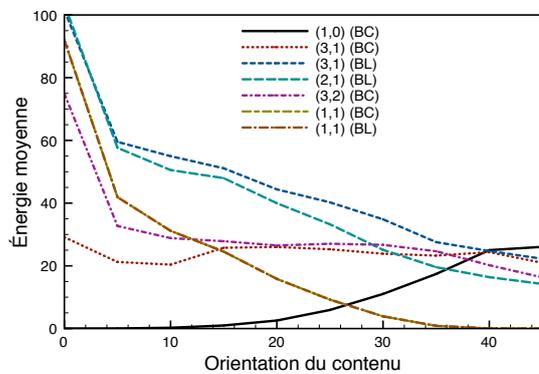
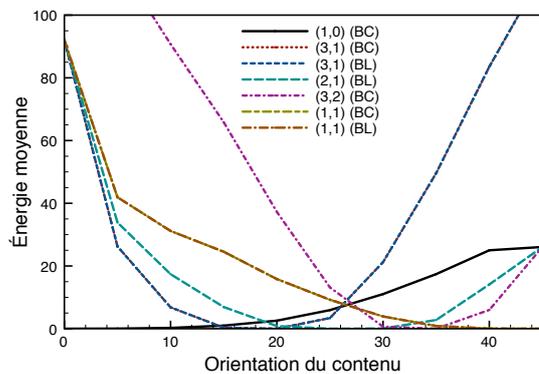
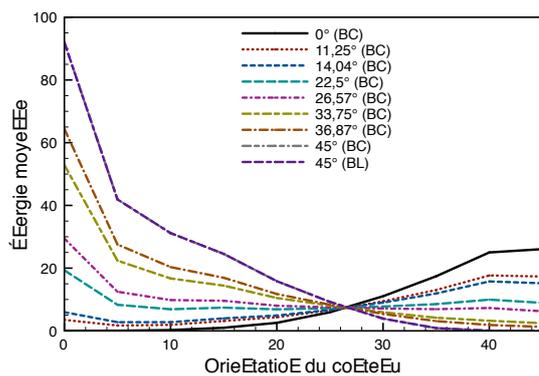
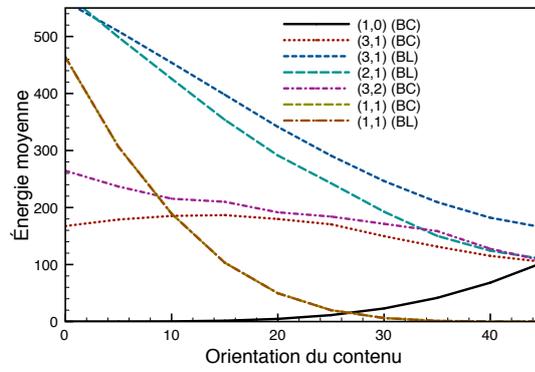
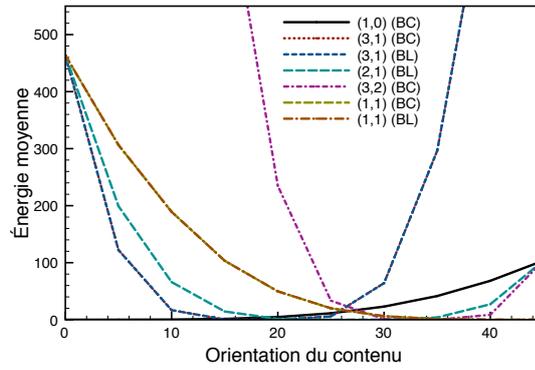
(a) Réveillès- $(p, q)$ -redressement(b) Mojette- $(p, q)$ -redressement(c)  $\theta$ -redressement

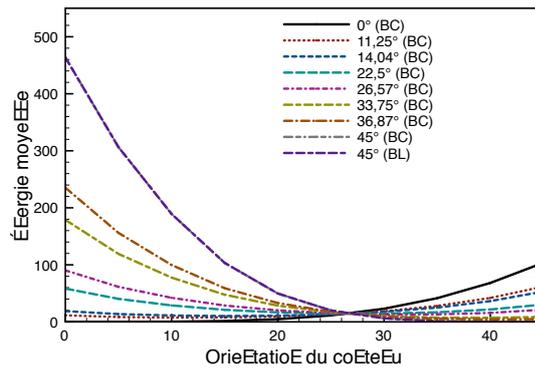
FIGURE 2.12 – Énergie des images représentant le contenu 3, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.



(a) Réveillès- $(p, q)$ -redressement



(b) Mojette- $(p, q)$ -redressement



(c)  $\theta$ -redressement

FIGURE 2.13 – Énergie des images représentant le contenu 4, pour différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.

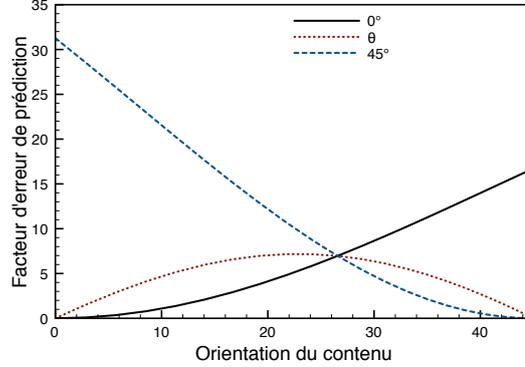


FIGURE 2.14 – Facteurs d’erreur de prédiction pour des filtres selon  $0^\circ$  et  $45^\circ$  de sinusoides, de période  $T = 5$ , orientées selon  $\theta$ ,  $\theta$  variant de  $0^\circ$  à  $45^\circ$ , mis en concurrence avec le facteur d’erreurs de prédiction pour un filtrage adapté de ces mêmes sinusoides selon leur propre orientation.

$$err_{0^\circ}^{\theta_{ech}} = \lambda \cdot \left[ 1 - \cos \left( \frac{2\pi}{T} \sin \theta_{ech} \right) \right] \quad (2.27)$$

$$err_{\theta_{ech}}^{\theta_{ech}} = \lambda \cdot \left[ 1 - \left( (1-t) \cdot \cos \left( \frac{2\pi}{T} \sin \theta_{ech} \right) + t \cdot \cos \left( \frac{2\pi}{T} (\cos \theta_{ech} - \sin \theta_{ech}) \right) \right) \right] \quad (2.28)$$

$$err_{45^\circ}^{\theta_{ech}} = \lambda \cdot \left[ 1 - \cos \left( \frac{2\pi}{T} (\cos \theta_{ech} - \sin \theta_{ech}) \right) \right] \quad (2.29)$$

avec  $\lambda = \frac{1}{\sqrt{2}} \cdot \frac{M-m}{2} \cdot \cos \left( \frac{2\pi t}{T} \right)$ .

On remarque que ces erreurs, qui concernent un pixel à la position  $(x, y)$ , sont toutes proportionnelles à  $\cos \left( \frac{2\pi t}{T} \right)$  qui est fonction de ses coordonnées. La moyenne des erreurs pour différentes coordonnées est donc aussi proportionnelle à la moyenne des terme  $\cos \left( \frac{2\pi t}{T} \right)$ . Le graphique de la figure 2.14 représente pour chaque contenu orienté selon  $\theta_{ech} - \tan \theta_{ech}$  variant de 0 à 1 par pas de 0,1 – le facteur d’erreur de prédiction correspondant à un filtrage selon  $0^\circ$ , selon  $45^\circ$  et selon l’orientation du contenu. Si les deux premières courbes correspondent à des erreurs de prédiction relatives à de mauvaises prédictions dues à une orientation de filtrage plus ou moins différente de l’orientation du contenu, pour la troisième courbe, ces deux orientations coïncident. L’erreur de prédiction correspond alors strictement à l’erreur d’interpolation.

On remarque que le facteur d’erreur d’interpolation est toujours plus élevé que le minimum des facteurs d’erreur de prédiction correspondant à  $0^\circ$  et  $45^\circ$ .

**Conclusion** Les trois techniques de redressement présentées dans la section 2.2 ont été appliquées sur différents contenus rectilignes dont l’orientation varie entre  $0^\circ$  et

45°, avec un certain nombre d'orientations de filtrage, dans le but de comparer leur capacité à minimiser l'énergie de la sous-bande reconstruite des hautes fréquences dans la direction de filtrage.

Seule la technique de Mojette- $(p, q)$ -redressement permet, pour tous les contenus considérés, de minimiser le terme d'énergie étudié pour des orientations de filtrage autres que l'horizontale ou la diagonale (45°). Les techniques de Réveillès- $(p, q)$ -redressement et de  $\theta$ -redressement sont quant à elles inefficaces pour ces types de contenus car d'une part localement des éléments situés d'un côté d'un contour se retrouvent prédits ou mis à jour à partir uniquement d'éléments situés de l'autre côté de ce contour, et d'autre part à cause de l'erreur d'interpolation introduite pour la technique avec interpolation linéaire.

La capacité de la technique de Mojette- $(p, q)$ -redressement à produire un terme d'énergie plus faible que celui obtenu par un filtrage selon l'horizontale ou la verticale varie d'un contenu à l'autre. Nous avons observé que cette diminution d'énergie est plus faible pour les contenus dont la structure linéique représentant un contour présente une transition douce.

### 2.3.2 Contenu circulaire

Dans cette sous-section, la capacité des techniques de redressement présentées précédemment, à minimiser l'énergie de la sous-bande des hautes fréquences dans la direction de filtrage est étudiée dans le cas moins favorable où le contenu testé est courbe. Pour ce faire, des images de tests de dimension  $N \times N$ , avec  $N = 128$ , sont générées de façon à contenir une portion angulaire d'un disque de valeur 192 défini par l'inéquation  $(x - x_c)^2 + (y - y_c)^2 < R^2$ , où  $(x_c, y_c)$  représente le centre d'un tel disque, et  $R$  son rayon, le tout sur un fond uniforme de valeur 64.

Différentes portions d'un disque sont testées de façon à évaluer les capacités de minimisation du terme d'énergie en fonction de l'orientation du contour courbe et du rayon de courbure. Ces différents contenus sont organisés en fonction de la longueur de l'arc du disque qu'ils contiennent. Les longueurs d'arc testées sont  $\frac{\pi}{4}$ ,  $\frac{\pi}{8}$  et  $\frac{\pi}{16}$ .

Pour chacune des longueurs d'arc  $l$  testées, différentes portions d'un disque de rayon  $R = \frac{N-1}{2 \sin \frac{l}{2}}$  sont étudiées de façon à considérer des contours courbes d'orientation proche de  $k \frac{l}{2}$ ,  $k$  variant de 0 à  $\frac{\pi}{4} \cdot \frac{2}{l}$ .

Chaque orientation donne lieu à une image de test  $I_{k \frac{l}{2}}^l$  représentant une portion d'un disque de rayon  $R$ . Les coordonnées du centre de ce disque sont :

$$\left( N - 1 + R \sin \frac{(k-1) \cdot l}{2}, -R \cos \frac{(k+1) \cdot l}{2} \right)$$

sauf pour  $k = 0$  où l'ordonnée est définie par  $\frac{N-1}{2} - R$  afin de profiter du recentrage possible de la portion du disque.

Les figures 2.15, 2.18 et 2.21 illustrent comment les images de test comprenant une portion de disque de longueur d'arc  $\frac{\pi}{16}$  ont été générées. Ces images sont quant à elles présentées sur les figures 2.16, 2.19 et 2.22.

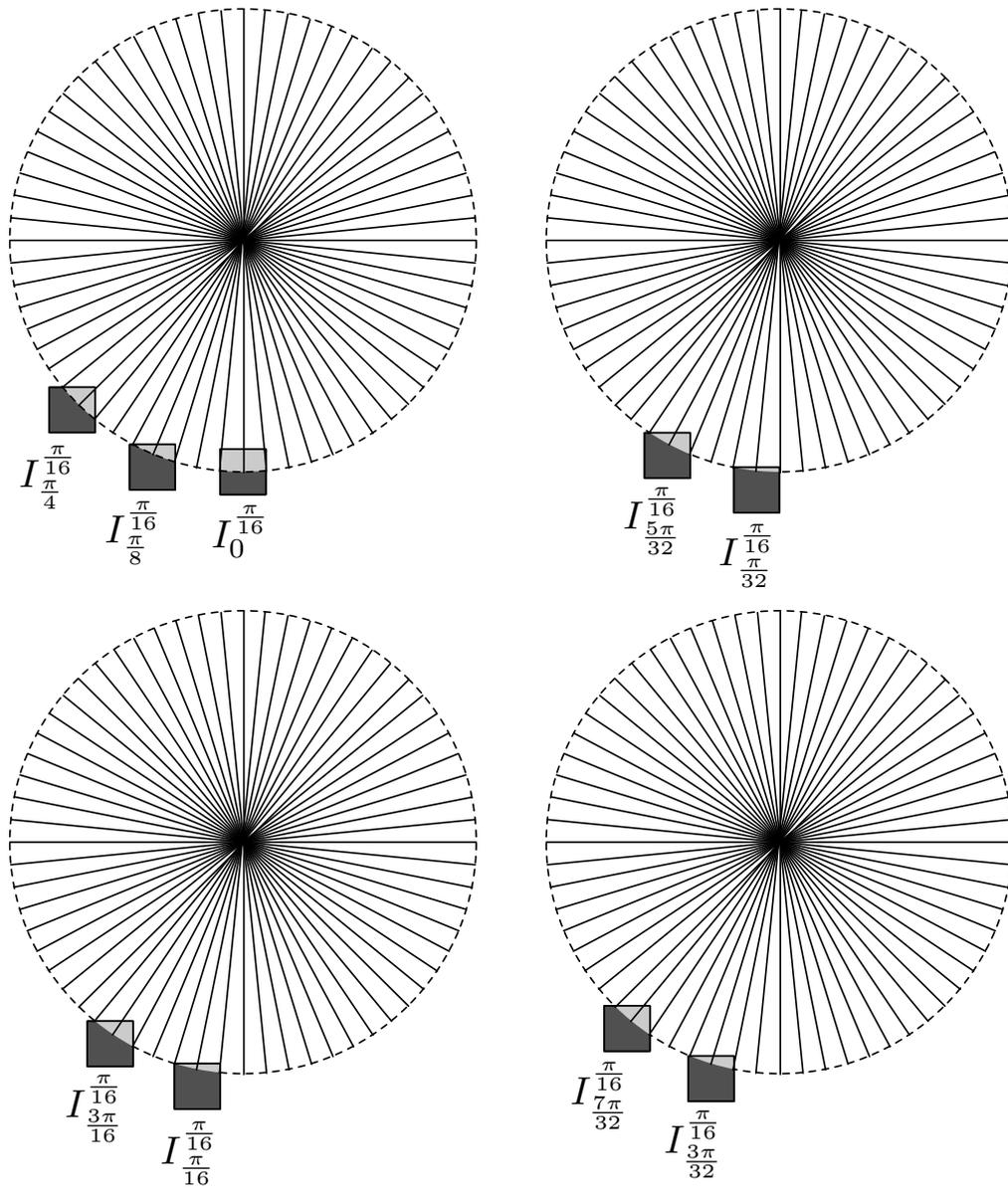
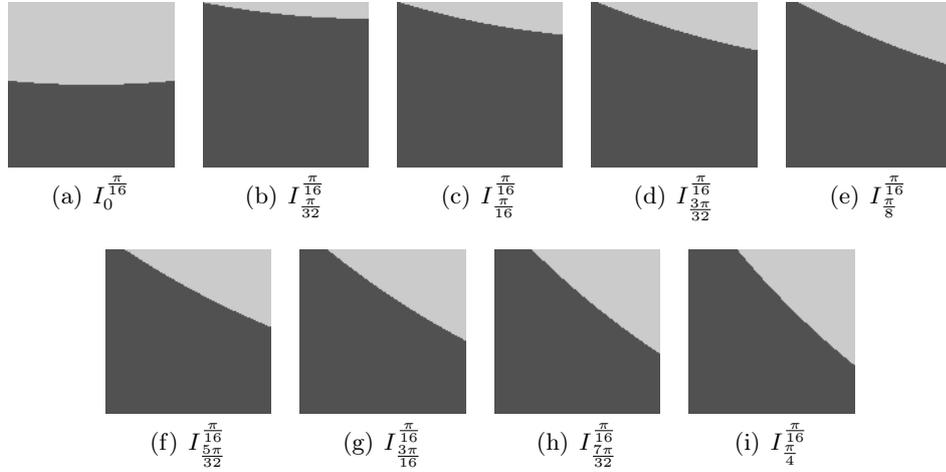


FIGURE 2.15 – Illustration de la génération des images de tests contenant une portion de disque de longueur d'arc  $\frac{\pi}{16}$ .

FIGURE 2.16 – Images de test contenant une portion de disque de longueur d’arc  $\frac{\pi}{16}$ .

Concernant les deux techniques de redressement qui n’ont pas donné satisfaction pour des orientations de filtrage autres que l’horizontale, la verticale et la diagonale, les résultats présentés sur les figures 2.17, 2.20 et 2.23 vont dans le même sens. Si l’on se focalise sur la technique de Mojette- $(p, q)$ -redressement, on observe que plus le rayon de courbure est faible (et donc la courbure forte), moins les orientations de filtrage  $(p, q)$  de norme supérieure à 1 sont efficaces. En effet, plus la courbure est forte, et plus la tentative de prédire un pixel de la portion du disque par des pixels éloignés ( $\|(p, q)\| > 1$ ) se révèle infructueuse. De plus, toutes les portions de disques ne sont pas filtrées pareillement. Par exemple, l’énergie d’un contour dont l’orientation est proche de  $0^\circ$ , filtré selon la direction  $(1, 0)$ , sera plus faible que l’énergie d’un contour dont l’orientation est proche de  $26,57^\circ$ , filtré selon  $(2, 1)$ .

### 2.3.3 Contenu naturel

Si les différents types de contenus synthétiques précédemment expérimentés permettent une compréhension fine des mécanismes mis en œuvre dans les différentes techniques de redressement et de leur impact sur leurs performances, il est difficile d’établir lequel de ces contenus correspond le mieux à un contour “naturel”. À défaut, les trois techniques de filtrage orienté sont appliquées sur des blocs extraits d’images naturelles de test bien connues. Ces parties d’images ont été choisies de façon à contenir un contenu ne présentant qu’une seule orientation principale. Ce type de contenu sera par la suite appelé “mono-orienté”. Parmi les différents blocs analysés, certains correspondent à des textures orientées plus ou moins parasitées par du bruit et des orientations secondaires (figure 2.24-a et b), les autres correspondent à des contours forts selon différentes orientations plus ou moins rectilignes (figure 2.24-c, d, e, f et g).

Tout d’abord quelques considérations générales sur les différents résultats présentés sur les figures 2.25 à 2.33. Des considérations qui peuvent apparaître contradictoires

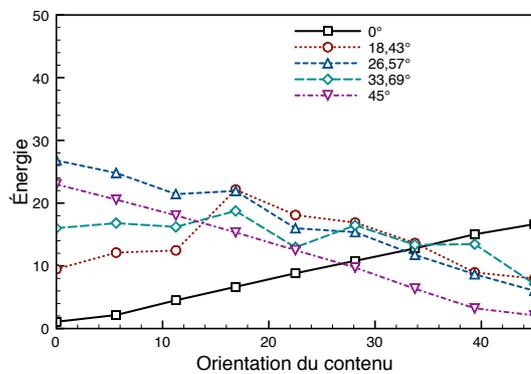
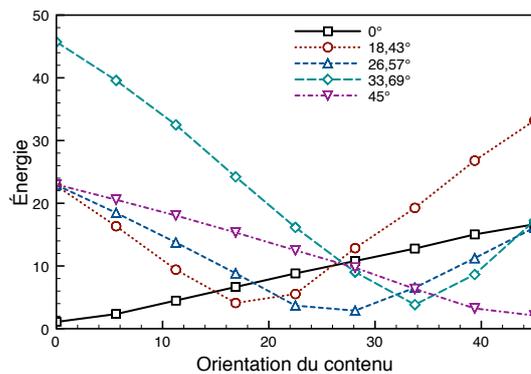
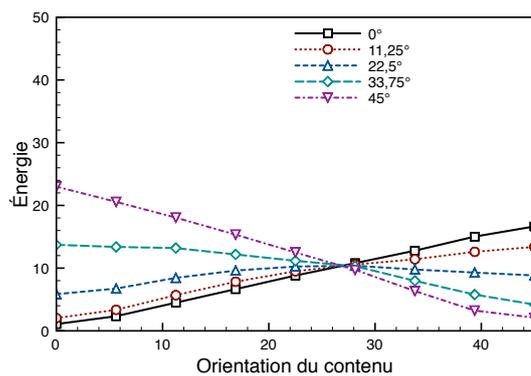
(a) Réveillès-( $p, q$ )-redressement(b) Mojette-( $p, q$ )-redressement(c)  $\theta$ -redressement

FIGURE 2.17 – Énergie des images représentant des portions de disque de longueur d'arc  $\frac{\pi}{16}$ , de différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.

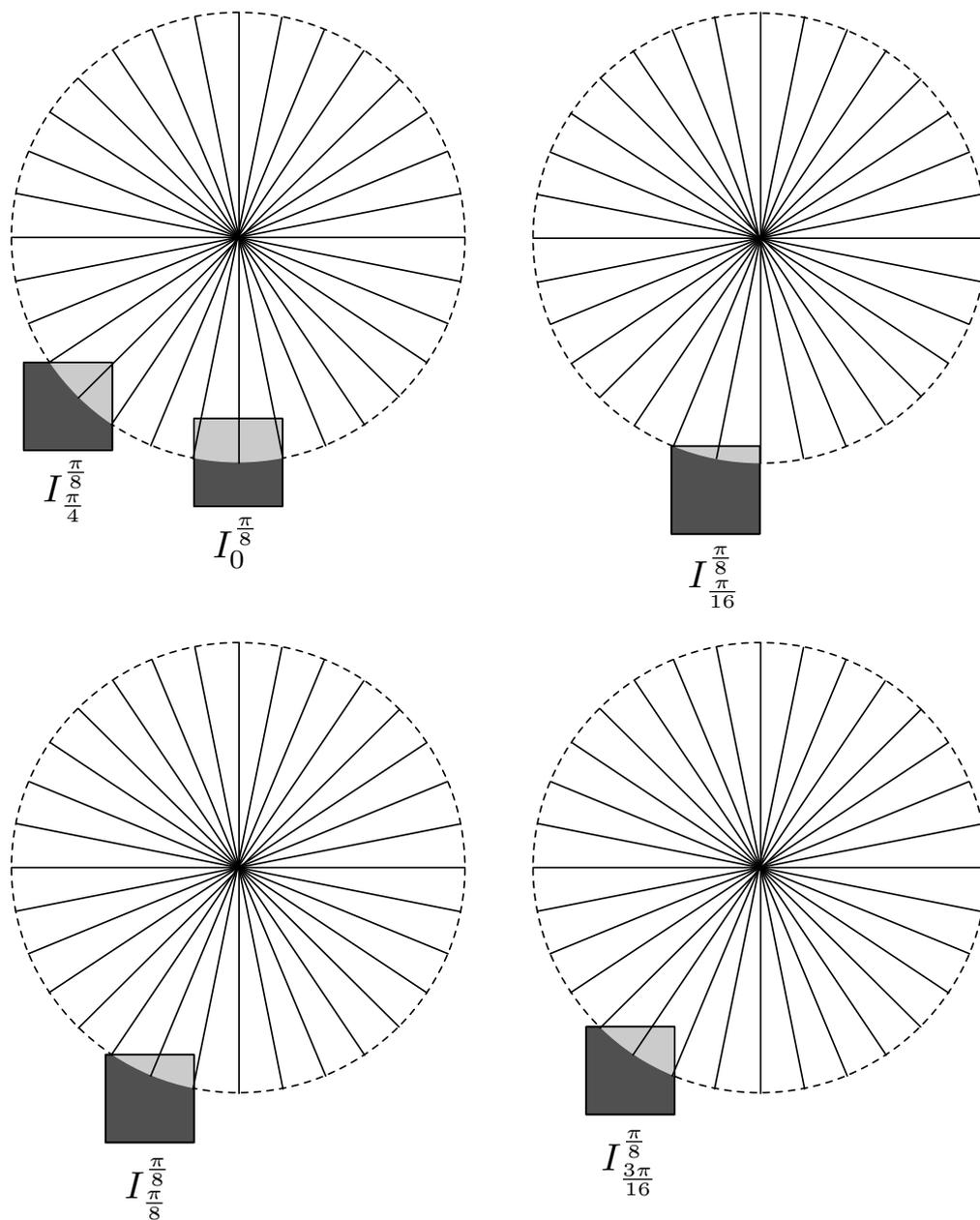


FIGURE 2.18 – Illustration de la génération des images de tests contenant une portion de disque de longueur d'arc  $\frac{\pi}{8}$ .

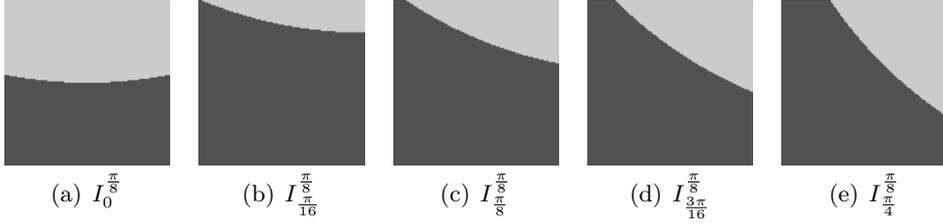


FIGURE 2.19 – Images de test comprenant une portion de disque de longueur d’arc  $\frac{\pi}{8}$ .

aux conclusions tirées des expérimentations sur les différents contenus synthétiques.

Contrairement aux précédentes expérimentations pour le  $\theta$ -redressement, certaines orientations de filtrage différentes de l’horizontale, la verticale et les diagonales arrivent à minimiser l’énergie des hautes fréquences. De plus l’orientation de filtrage en question peut s’avérer plus proche de l’orientation que nous percevons du contours ou de la texture (voir figure 2.25, 2.26, 2.27 et 2.29). En ce qui concerne le Mojette- $(p, q)$ -redressement, on observe des résultats instables d’une orientation de filtrage à l’autre, là où la stabilité était plutôt de mise pour les expérimentations synthétiques. À l’opposé, le Réveillès- $(p, q)$ -redressement apparaît plus stable que précédemment.

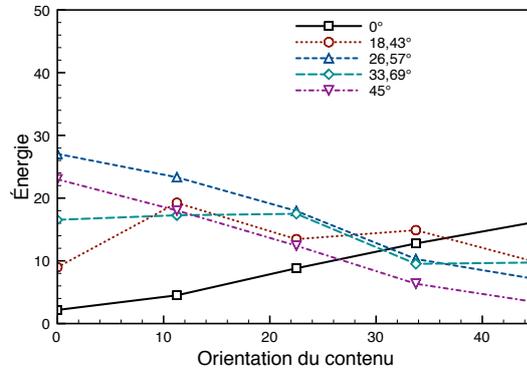
D’autres observations peuvent être faites concernant maintenant l’efficacité comparée des différentes techniques de redressement. Au regard des neuf morceaux choisis, on peut regrouper les différents résultats en trois catégories. Les figures 2.26 et 2.32, où la technique Mojette- $(p, q)$ -redressement est la plus performante. Les figures 2.25, 2.27 et 2.29, où le  $\theta$ -redressement produit un minimum d’énergie plus faible. Enfin sur les figures 2.28, 2.30, 2.31 et 2.33 le minimum est obtenu pour une direction de filtrage horizontale, verticale ou diagonale, et donc toutes les trois techniques produisent la même valeur.

Afin d’expliquer les différences de comportement des différentes techniques de redressement sur les contenus synthétiques et les contenus naturels, deux phénomènes doivent d’être analysés. D’une part pourquoi les résultats avec le Mojette- $(p, q)$ -redressement sont si variables. Et d’autre part comment expliquer que le  $\theta$ -redressement puisse apporter des résultats plus intéressants que précédemment pour des orientations de filtrage impliquant une interpolation.

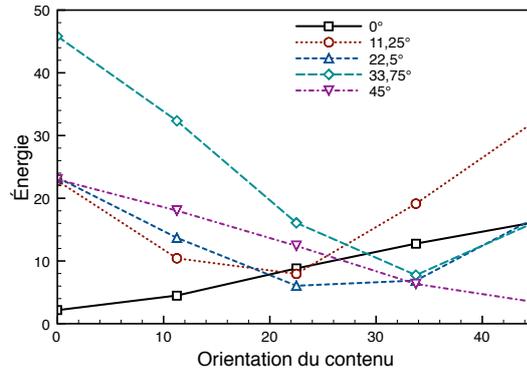
Dans les images synthétiques des sous-sections 2.3.1 et 2.3.2, deux éléments ne sont pas intervenus : la cohérence spatiale et le bruit.

En ce qui concerne le Mojette- $(p, q)$ -redressement, pour peu que l’on aille chercher la valeur d’un pixel du bon côté du contour, peu importe l’éloignement de ce pixel, l’uniformité des images synthétisées ou leur périodicité permettrait une très bonne prédiction dans certains cas. En pratique, plus le pixel sera loin, plus la probabilité que sa valeur soit proche sera faible. Ceci explique que pour des orientations telles que  $\|(p, q)\| > 1$ , la technique de redressement ne soit pas efficace. On observe qu’elle surpasse néanmoins les autres méthodes pour des contours très forts et rectilignes 2.26 et 2.32.

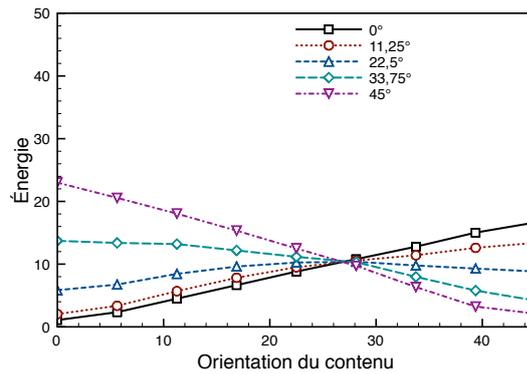
Le bruit peut quant à lui expliquer en partie le fait que le redressement avec interpolation soit plus intéressant que sans interpolation. Comme nous l’avons vu pré-



(a) Réveillès-( $p, q$ )-redressement



(b) Mojette-( $p, q$ )-redressement



(c)  $\theta$ -redressement

FIGURE 2.20 – Énergie des images représentant des portions de disque de longueur d'arc  $\frac{\pi}{8}$ , de différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.

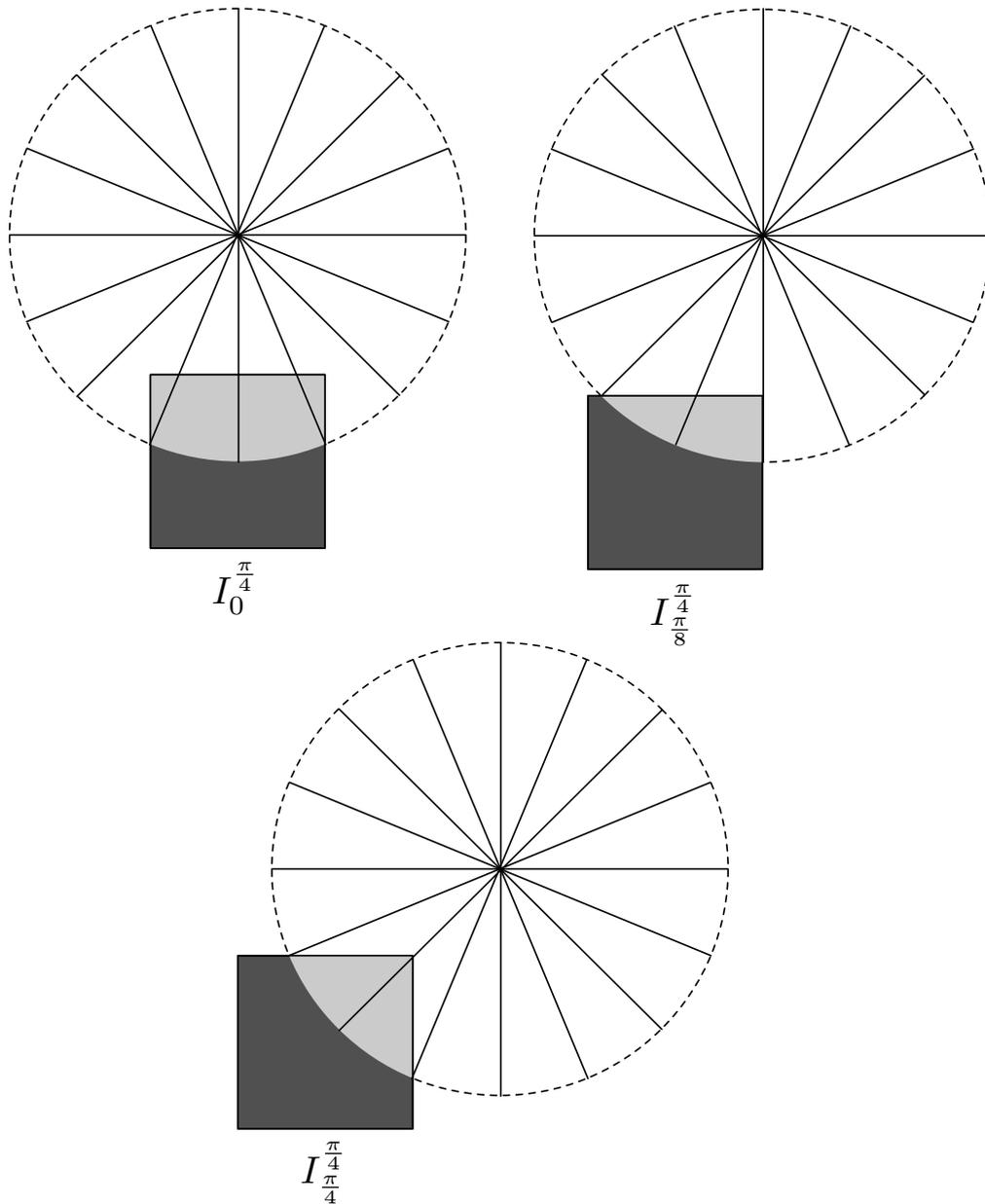


FIGURE 2.21 – Illustration de la génération des images de tests contenant une portion de disque de longueur d'arc  $\frac{\pi}{4}$ .

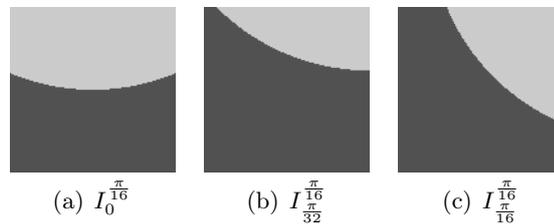


FIGURE 2.22 – Images de test comprenant une portion de disque de longueur  $\frac{\pi}{4}$ .

cédemment pour le cas de sinusôides, l'erreur d'interpolation introduite a tendance à augmenter l'erreur de prédiction. Or l'interpolation, qui introduit un filtrage passe-bas supplémentaire, a tendance à réduire le bruit, et donc diminuer l'erreur de prédiction. La différence d'énergie très faible entre l'orientation minimale qui nécessite une interpolation et l'orientation horizontale, verticale ou diagonale la plus proche étant très faible, ce simple effet lisseur pourrait donc bien suffire à faire pencher la balance.

Finalement on observe tout de même que certains contours, bien que présentant des orientations particulières, ont leur énergie minimisée par un simple filtrage horizontal, vertical ou diagonal (figures 2.28, 2.30 et 2.31).

## Conclusion

Ce chapitre a mis en évidence la flexibilité de la technique de redressement pour définir des transformées en ondelettes orientées. Une telle souplesse permet donc diverses implantations, dont trois ont fait l'objet d'une étude particulière. L'une basée sur la définition de droites discrètes 8-connexes, la deuxième sur le choix d'échantillons sur un réseau de points, la dernière sur une interpolation linéaire des voisins 8-connexes.

Des expérimentations sur des images synthétiques non bruitées représentant des contours ou textures, rectilignes ou légèrement courbes, ont permis de mettre en évidence une meilleure minimisation de l'énergie des hautes fréquences selon l'orientation  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  ou  $135^\circ$  la plus proche de l'orientation contenue dans l'image. Pour la technique basée « lattice », le Mojette- $(p, q)$ -redressement, il apparaît qu'une subdivision plus fine des orientations permet d'améliorer cette minimisation, contrairement à la méthode avec interpolation à cause justement de l'introduction d'erreurs d'interpolation. Ceci reste toutefois contraint à une faible courbure des contours ou textures.

Appliquées sur des portions d'images naturelles de contenu mono-orienté, la notion de cohérence spatiale entre alors en ligne de compte en défaveur de la technique de redressement basée lattice. À l'opposé l'effet lisseur de l'interpolation permet d'améliorer très légèrement les performances du  $\theta$ -redressement, mais suffisamment pour qu'une orientation nécessitant de l'interpolation soit plus intéressante que les simples directions horizontale, verticale et diagonales.

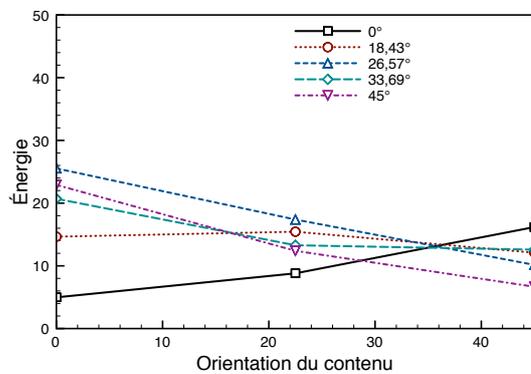
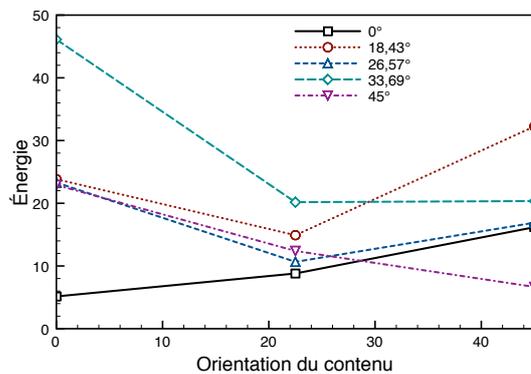
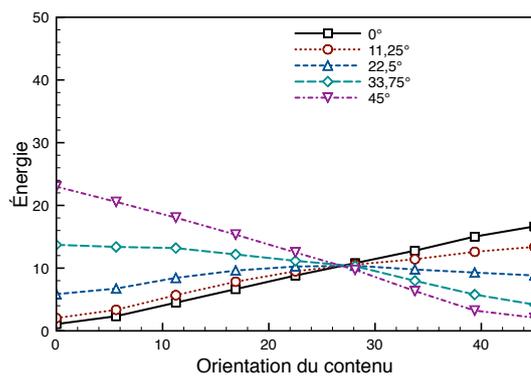
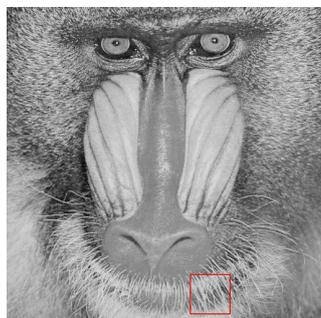
(a) Réveillès- $(p, q)$ -redressement(b) Mojette- $(p, q)$ -redressement(c)  $\theta$ -redressement

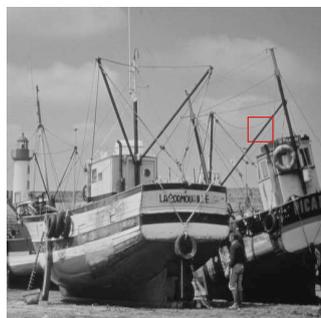
FIGURE 2.23 – Énergie des images représentant des portions de disque de longueur d'arc  $\frac{\pi}{4}$ , de différentes orientations, reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement.



(a) baboon



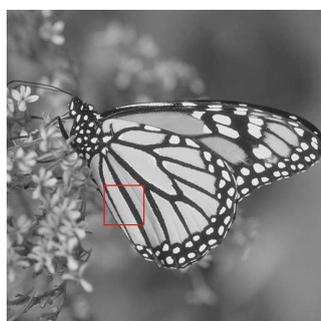
(b) barbara



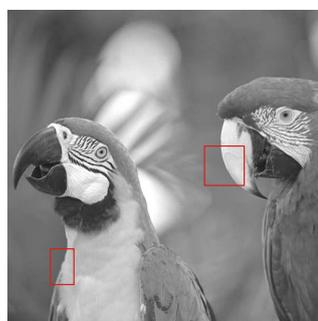
(c) boats



(d) fruits



(e) monarch



(f) parrots



(g) plane

FIGURE 2.24 – Les morceaux choisis de différentes images de test naturelles.

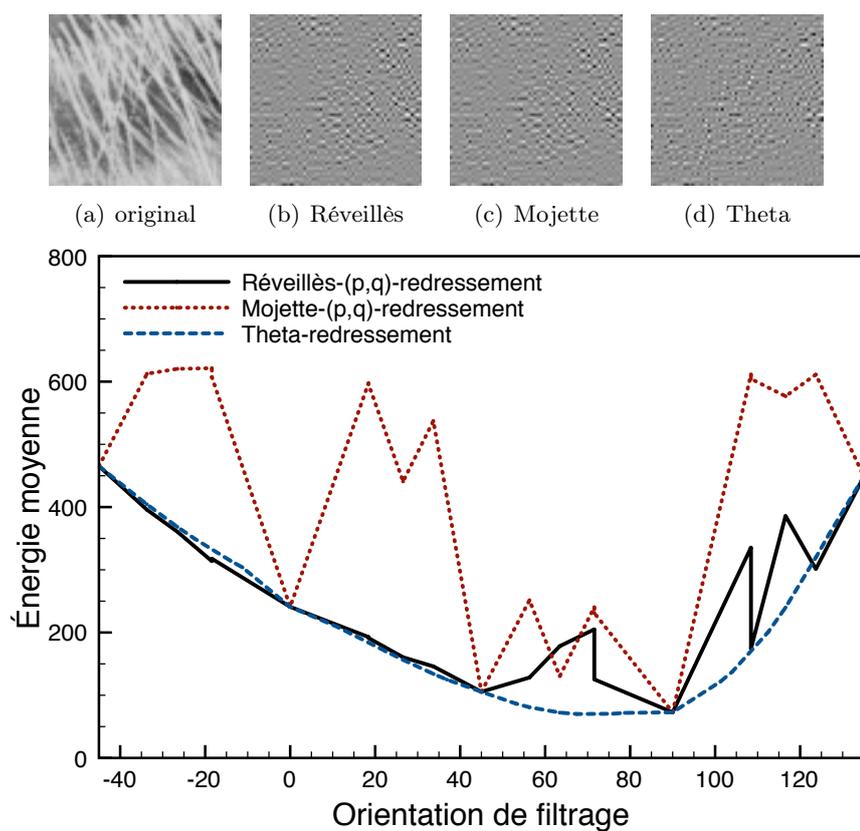


FIGURE 2.25 – Énergie de l'image baboon\_mouth\_hair (a) de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $90^\circ$ ,  $90^\circ$  et  $67,5^\circ$ .

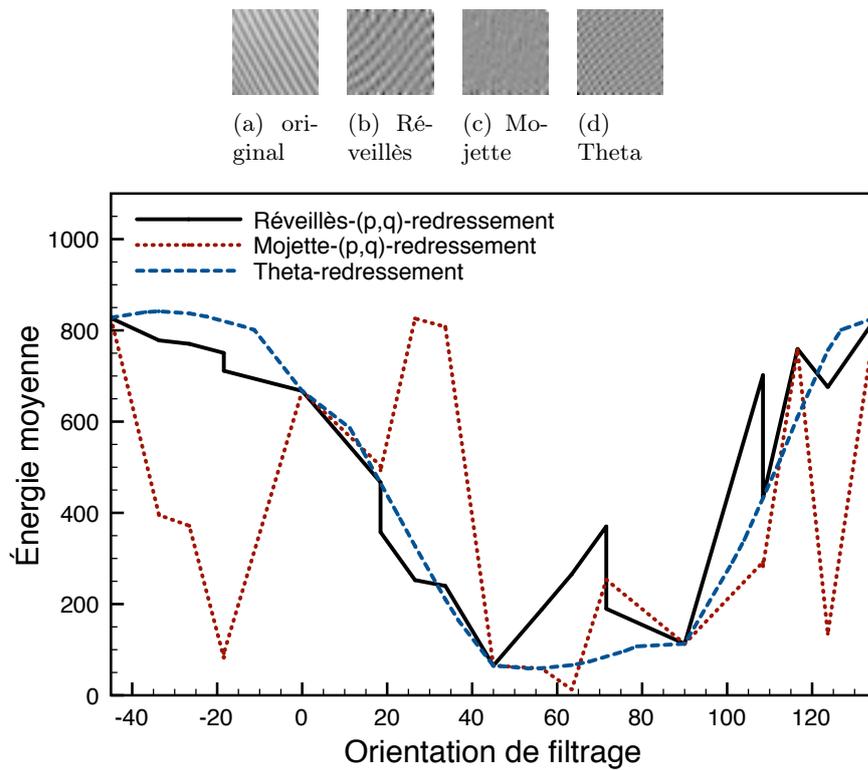


FIGURE 2.26 – Énergie de l'image `barbara_pants` (a) de dimension  $32 \times 32$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $45^\circ$ ,  $63,43^\circ$  et  $53,13^\circ$ .

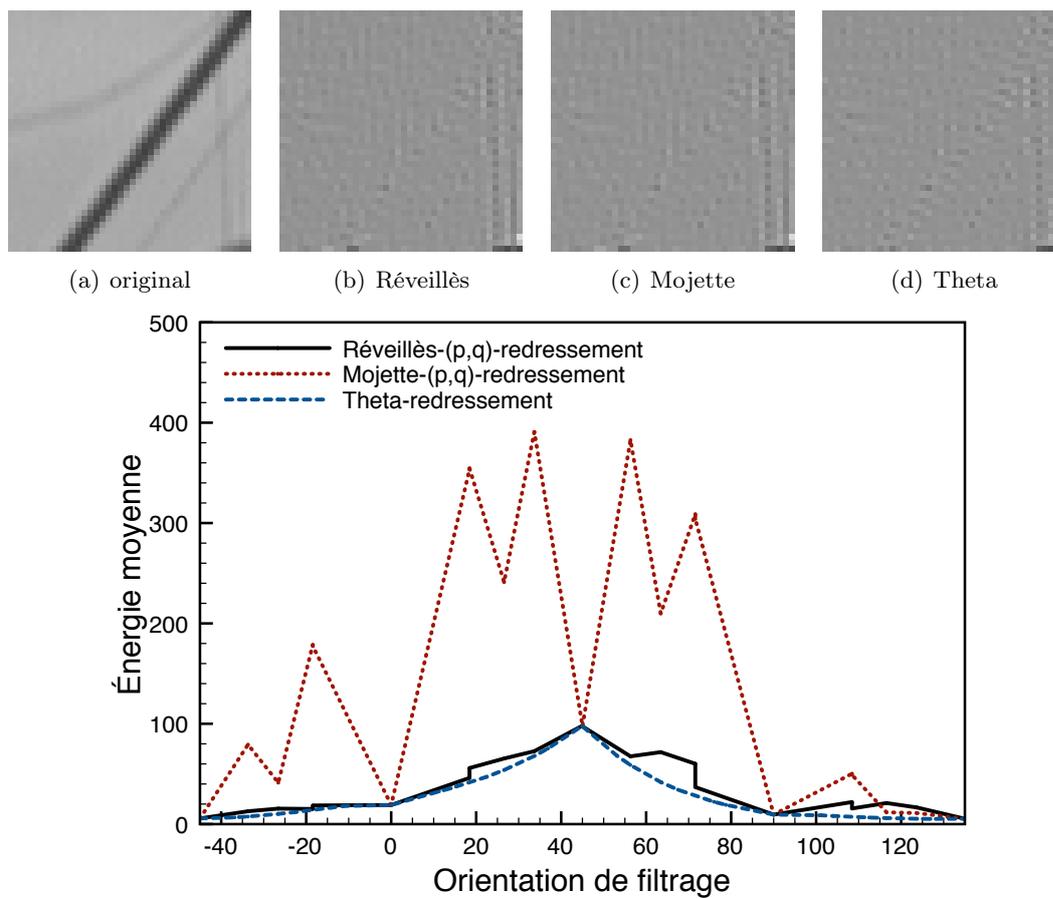


FIGURE 2.27 – Énergie de l'image `boats_mast` (a) de dimension  $40 \times 40$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $135^\circ$ ,  $135^\circ$  et  $126,87^\circ$ .

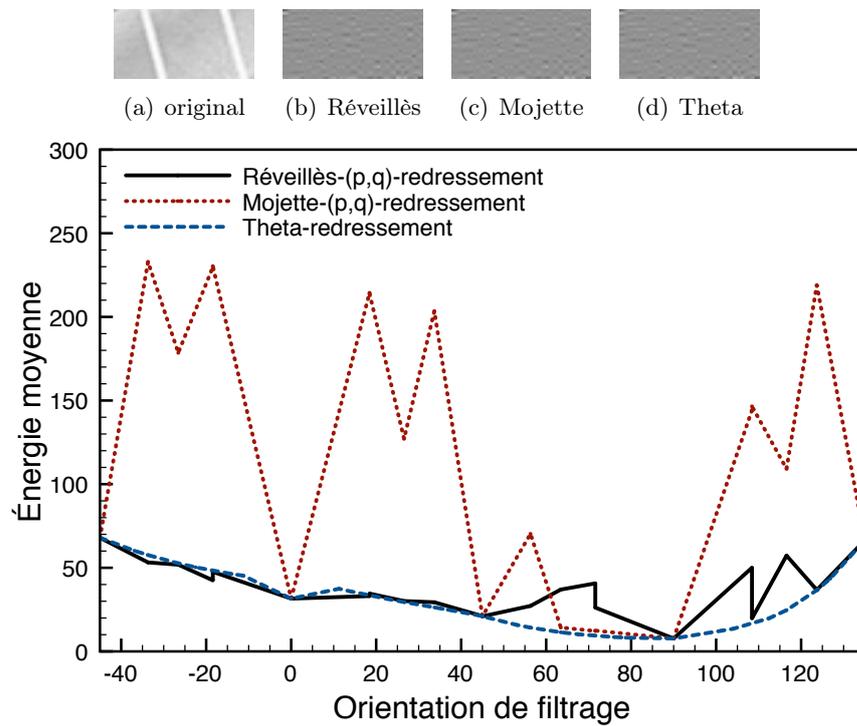


FIGURE 2.28 – Énergie de l'image `fruits_racket` (a) de dimension  $52 \times 26$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $90^\circ$ ,  $90^\circ$  et  $90^\circ$ .

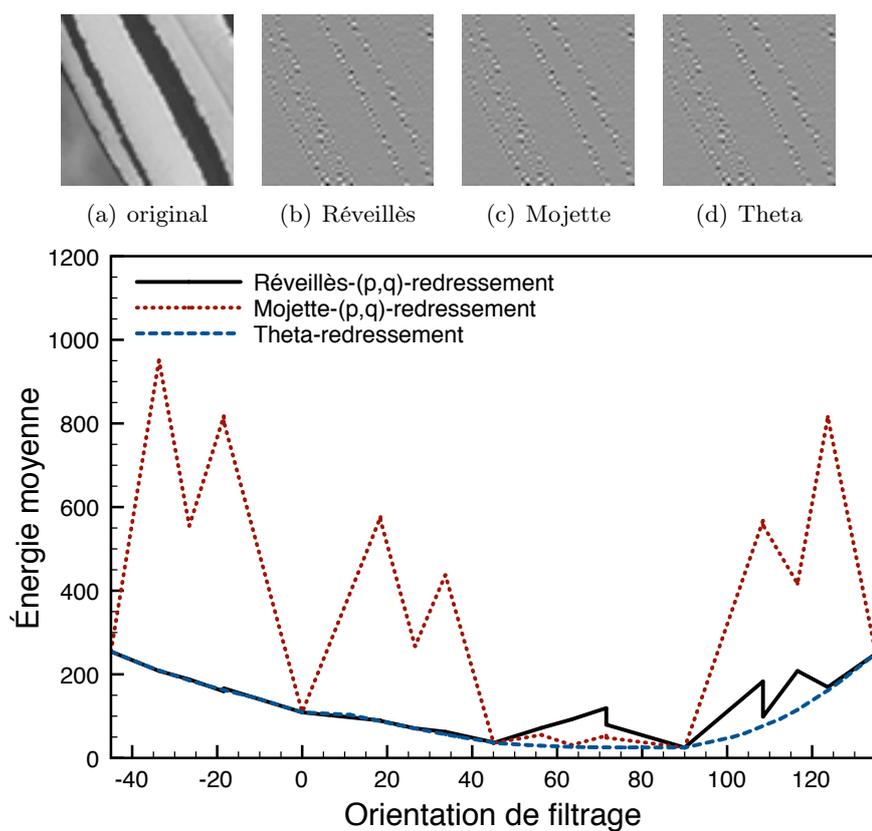


FIGURE 2.29 – Énergie de l'image `monarch_wing` (a) de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $90^\circ$ ,  $90^\circ$  et  $90^\circ$ .

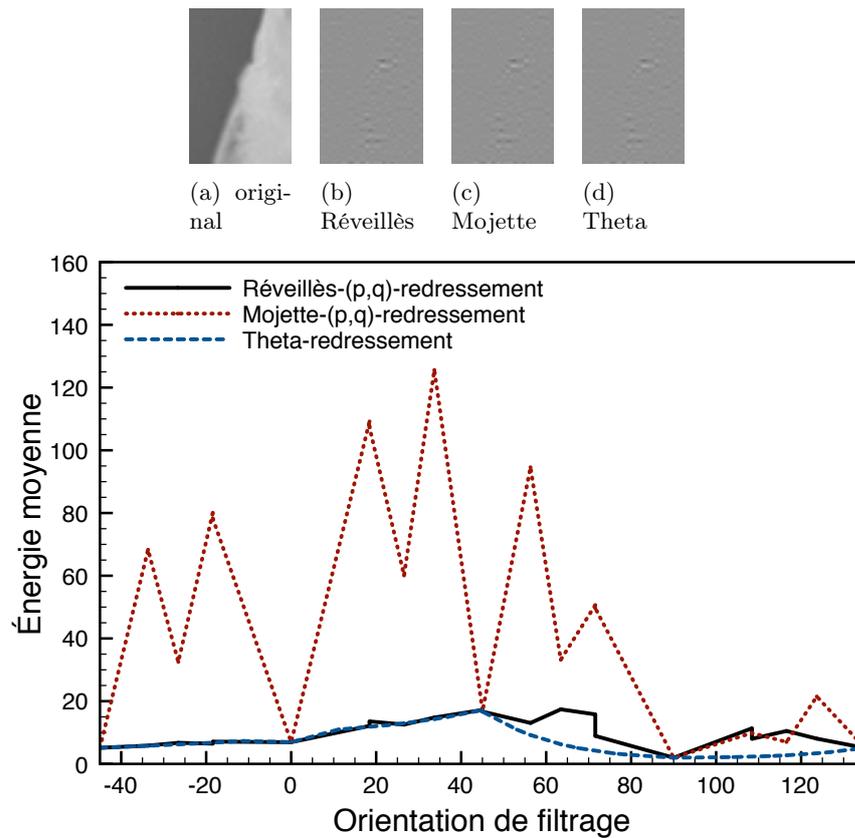


FIGURE 2.30 – Énergie de l'image `parrots_fether` (a) de dimension  $38 \times 58$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $90^\circ$ ,  $90^\circ$  et  $90^\circ$ .

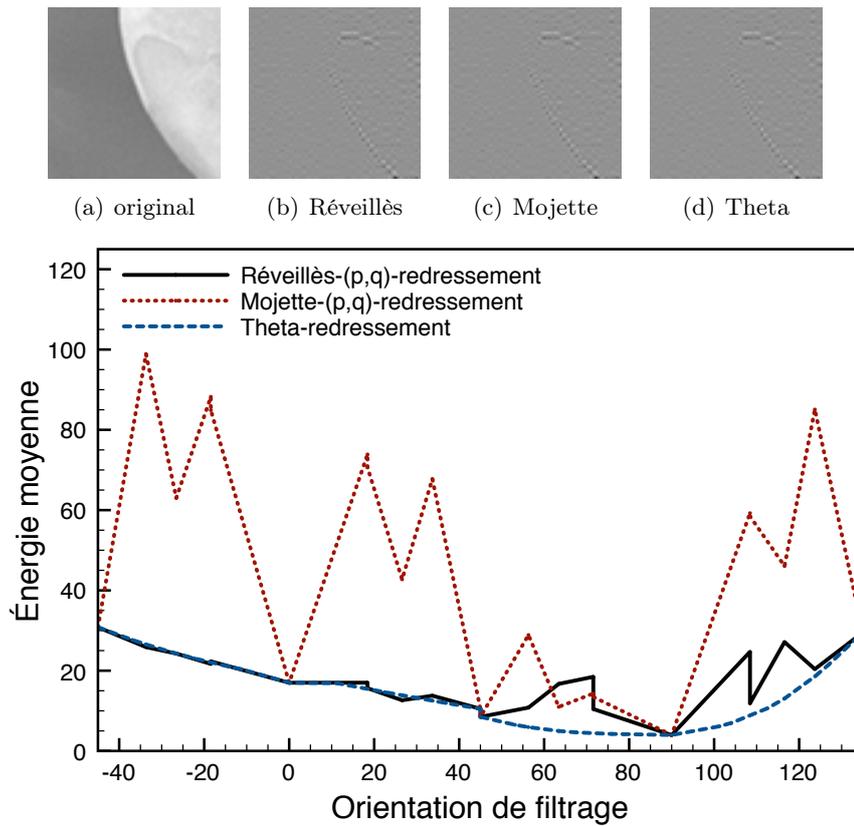


FIGURE 2.31 – Énergie de l'image `parrots_nose` (a) de dimension  $64 \times 64$ , reconstruites à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $90^\circ$ ,  $90^\circ$  et  $90^\circ$ .

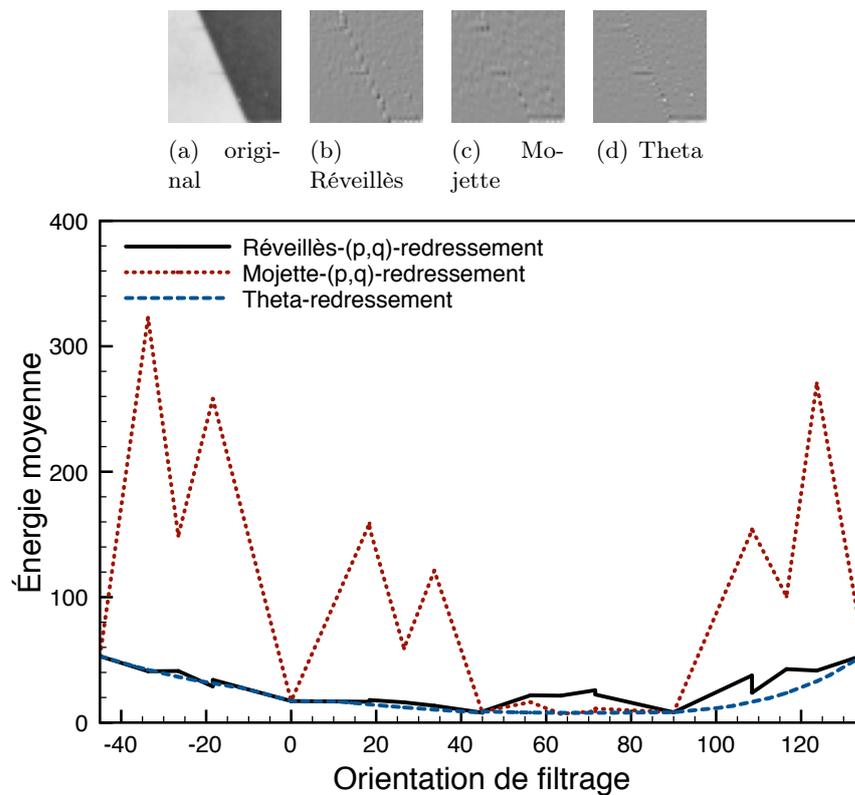


FIGURE 2.32 – Énergie de l'image `plane_wing_back` (a) de dimension  $42 \times 42$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $45^\circ$ ,  $63,43^\circ$  et  $67,5^\circ$ .

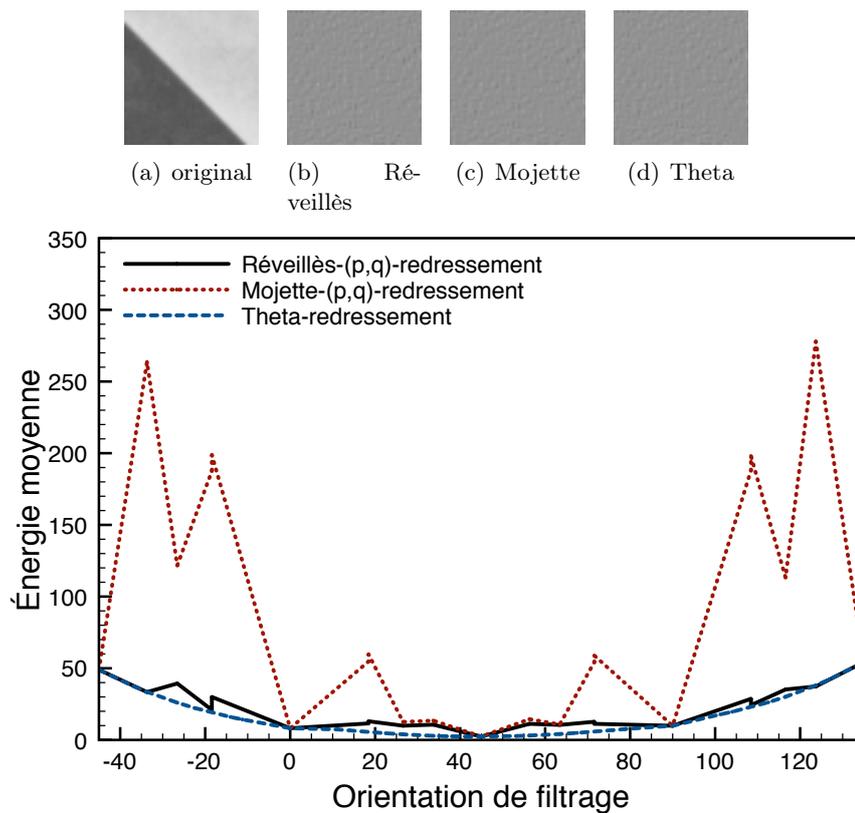


FIGURE 2.33 – Énergie de l'image `plane_wing_front` (a) de dimension  $50 \times 50$ , reconstruite à partir des coefficients des hautes fréquences uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage. (b), (c) et (d) représentent cette image reconstruite pour chacune des trois techniques lorsque l'énergie de l'image est minimale : respectivement pour  $45^\circ$ ,  $45^\circ$  et  $45^\circ$ .

Si la méthode basée sur des droites discrètes 8-connexes n'a quant à elle jamais donné satisfaction, les deux autres techniques ont montré qu'elles pouvaient permettre la minimisation de l'énergie des hautes fréquences dans la direction de filtrage pour un ensemble de directions assez restreint. Au final, suivant certaines caractéristiques du contour filtré (niveau du gradient, rectitude, bruit, etc.) l'une de ces deux techniques de redressement sera plus efficace que l'autre pour minimiser l'énergie des hautes fréquences dans la direction de filtrage. Des réserves peuvent toutefois être émises quant au gain apporté par une subdivision supérieure aux quatre directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ .



## Chapitre 3

# Estimation de l'orientation de filtrage et classification structurelle

### Introduction

Les transformations en ondelettes orientées n'ont pas vocation à s'appliquer selon une seule et unique orientation de filtrage pour toute une image. Leur intérêt réside au contraire dans leur adaptation au contenu de l'image. Par conséquent l'orientation de filtrage d'une technique de redressement est définie localement. Cela induit au moins trois problématiques :

- la définition de cette localité (dimension et forme de la région sur laquelle une même orientation de filtrage est appliquée) ;
- le choix de l'orientation de filtrage locale ;
- la gestion du filtrage à la frontière des régions.

Dans ce chapitre, la segmentation de l'image en régions de même orientation de filtrage est considérée par blocs. Le problème de la gestion du filtrage à la frontière des différents blocs est étudié dans une première section. Ensuite, différentes solutions pour l'estimation de l'orientation de filtrage adéquate sont proposées.

### 3.1 Gestion des frontières de blocs pour le filtrage orienté

Bien que les orientations de filtrage soient définies par bloc, le filtrage mis en œuvre par les techniques de redressement ne se limite pas, dans un bloc, à n'utiliser que des échantillons de ce bloc. Trois cas sont à considérer. Dans les deux premiers cas, la décomposition polyphase est la même pour tous les blocs de l'image. Le premier cas est celui dégénéré où à chaque bloc est associée la même orientation de filtrage. Dans le deuxième, l'orientation de filtrage peut varier d'un bloc à l'autre. Enfin dans le dernier cas, la décomposition polyphase peut varier d'un bloc à l'autre.

Dans le premier cas, les effets de bord aux frontières des blocs n'existent pas de fait. L'opération de mise à jour est simplement définie par les mêmes règles que l'opération de prédiction.

Dans le deuxième cas, deux façons d'implanter l'opération de mise à jour peuvent être rencontrées dans la littérature. Soit la mise à jour s'effectue avec les mêmes règles que l'opération de prédiction, comme dans le cas précédent. On nommera cette politique de mise à jour « symétrique ». Comme illustré figure 3.1-b, un élément peut se retrouver mis à jour par des erreurs de prédiction qui n'ont pas été calculées à partir de cet élément mis à jour. Soit la mise à jour s'effectue en utilisant toutes les erreurs de prédiction qui ont utilisé l'élément mis à jour pour être calculées [CG07, CG05]. Nommons cette façon d'implanter « non symétrique ». Comme illustré figure 3.1-c, un élément peut être mis à jour par un nombre variable d'erreurs de prédiction, contrairement au cas précédent où un élément est toujours mis à jour à partir de deux erreurs de prédiction. Comme indiqué dans [CG05], ces différentes erreurs de prédiction peuvent être pondérées de différentes façons en fonction des filtres d'ondelettes utilisées et des propriétés statistiques souhaitées pour les sous-bandes produites.

Finalement, dans le dernier cas, lorsque les décompositions polyphases peuvent varier d'un bloc à l'autre, à un bloc peut être assignée une orientation de filtrage requérant une technique de redressement basé colonne, et à un autre bloc voisin peut être assignée une autre orientation de filtrage requérant une technique de redressement basé ligne (illustré figure 3.2). Comme on peut l'observer, certains éléments ne peuvent pas être prédits par des éléments qui pourraient être utilisés si la décomposition polyphase était partout la même. Ce problème a conduit dans un premier temps dans [JRB07a] à considérer en deux composantes indépendantes les blocs où était appliqué un redressement basé ligne, et ceux des blocs où était appliqué un redressement basé colonne, introduisant par là même des effets de blocs aux frontières de ces deux composantes. Toutefois si l'information nécessaire au filtrage n'est pas toujours disponible, elle l'est parfois, si bien que le filtrage peut s'effectuer comme précédemment pour certains éléments, et pour les cas où un élément n'est pas disponible, il peut être remplacé selon des règles communes définies au codeur et au décodeur sans ajouter ainsi d'informations codées supplémentaires (*overhead*).

## 3.2 Estimation de l'orientation de filtrage d'une région de l'image

Pour procéder à un filtrage orienté adapté à la géométrie d'une image, il faut segmenter cette image et associer à chaque région une orientation de filtrage particulière. Il convient donc d'estimer la géométrie de cette image afin de déterminer quelles orientations choisir. Cette géométrie peut être estimée de façon explicite ou implicite. Explicite, des contours peuvent être extraits de l'image par exemple. Implicite, le modèle qui correspond le mieux à la géométrie de l'image est choisi.

Dans une première sous-section, les contours forts de l'image sont extraits et l'orientation des éléments de ces contours sert de mesure pour segmenter l'image. Dans les deux sous-sections suivantes, l'estimation de la géométrie est implicite. Elle est d'abord directement en relation avec la meilleure orientation de filtrage pour une technique de redressement donnée, puis liée à une modélisation stochastique de l'orientation du

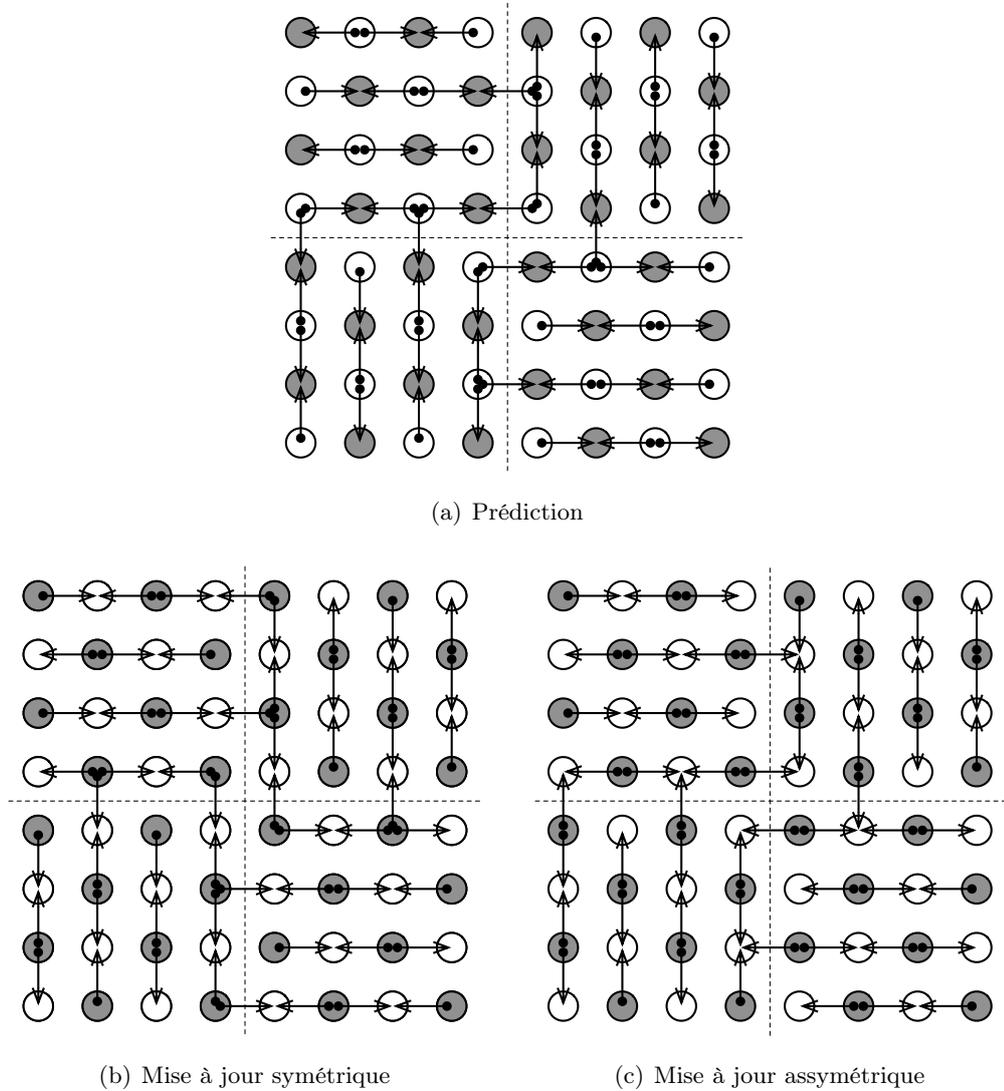


FIGURE 3.1 – Différentes politiques de mise à jour aux frontières des blocs.

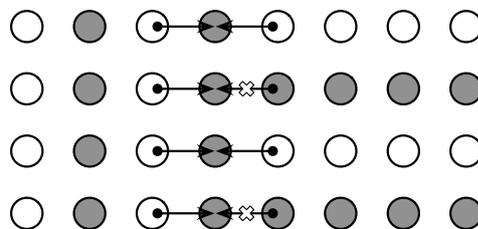


FIGURE 3.2 – Problème de la gestion des bords lorsque des blocs voisins ont une décomposition polyphase différente.

gradient en chaque pixel de l'image.

### 3.2.1 Estimation de l'orientation des contours extraits explicitement d'une image

L'estimation de la géométrie d'une image, à savoir les directions de régularité dans cette image, peuvent être obtenues à travers une extraction des contours. Pour cela l'algorithme d'extraction de contours décrit dans la thèse de Jean-François Bertrand [Ber87] est utilisé. Pour résumer cet algorithme, après seuillage de la réponse maximale à quatre filtres de gradient orientés définis comme suit :

$$G_{0^\circ} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & \underset{\circ}{0} & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (3.1)$$

$$G_{45^\circ} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & \underset{\circ}{0} & 1 \\ -2 & -1 & 0 \end{bmatrix} \quad (3.2)$$

$$G_{90^\circ} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & \underset{\circ}{0} & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (3.3)$$

$$G_{135^\circ} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & \underset{\circ}{0} & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad (3.4)$$

où le symbole  $\underset{\circ}{\phantom{x}}$  indique la position de l'élément du filtre correspondant à l'élément de l'image sur lequel est appliqué la convolution.

Les structures composées des éléments seuillés sont amincies afin d'extraire les éléments extrema locaux du gradient. Une série de traitements est alors appliquée pour étiqueter les contours en tenant compte, pour les éléments de jonction, de la continuité de l'orientation du gradient. Les contours de longueur inférieure à un seuil sont supprimés.

Les cartes de contours extraites à différents niveaux de résolution pour quelques images de tests sont présentées à l'annexe D.

Les contours extraits, l'image est segmentée en arbre quaternaire de façon descendante selon le critère suivant :

- si une feuille ne contient aucun élément de contour, alors cette feuille correspond à une zone uniforme, et aucune orientation particulière n'est associée à cette feuille ;
- sinon l'orientation du contour est estimée en chaque élément de contour. Si un nombre significatif d'éléments de contours, par exemple 75%, appartiennent à la même classe d'orientation, alors cette orientation est choisie comme la direction de régularité et la feuille correspond à une zone uniforme. Si le nombre d'éléments de

contours représente un pourcentage inférieur à un seuil haut (75%) mais supérieur à un seuil bas, par exemple 50%, alors cette orientation est aussi choisie comme la direction de régularité, mais la feuille est classée multi-orientée ;

- sinon la feuille est subdivisée si ses dimensions n'ont pas atteint une limite prédéterminée, auquel cas la feuille est considérée comme une zone texturée, et là aussi aucune orientation n'est associée à cette feuille.

L'objectif souhaité de cette méthode était initialement de profiter de la connaissance de la position des éléments de contours pour améliorer localement le filtrage. En effet, nous avons constaté qu'un élément situé d'un côté d'un contour pouvait être prédit ou mis à jour à partir d'éléments situés de l'autre côté de ce contour. La connaissance a priori de la position du contour devrait donc permettre d'éviter ce type de comportement. Toutefois en pratique ici, un contour est défini par une suite de pixels 8-connexes (chemin), et non pas comme une frontière entre pixels. Un contour, tel qu'il est extrait par la méthode utilisée, ne correspond pas nécessairement à une ligne de niveau, c'est-à-dire que les éléments du chemin 8-connexe n'ont pas tous la même valeur. Il n'est donc pas particulièrement intéressant de filtrer strictement le long de ce contour extrait.

Cette méthode d'extraction, de contours a deux inconvénients majeurs. Outre les difficultés de choix des seuils de la méthode d'extraction des contours, le premier concerne l'estimation de l'orientation d'une chaîne d'éléments de contour. Le deuxième concerne le codage des contours extraits. Sans évoquer de techniques particulières de codage de chaînes qui seront elles évoquées dans la seconde partie de ce document, on peut d'ores et déjà observer que si la position des éléments de contours n'est plus utile, la représentation des chaînes d'éléments de contour codée sera redondante. Par exemple, plusieurs contours de même orientation dans une même région seront codés pour une même information d'orientation. Autre exemple, les éléments de contours des zones texturées sont elles codées pour rien. Pour de telles raisons, l'arbre quaternaire des orientations calculé à partir des éléments de contours pourrait alors être plus intéressant à coder.

### 3.2.2 Estimation implicite de l'orientation par minimisation d'une fonctionnelle

La minimisation de l'énergie de la sous-bande des hautes fréquences est le critère le plus souvent utilisé pour déterminer l'orientation de filtrage à appliquer.

Par exemple, pour les « Curved Wavelet » [WZS06], l'image est décomposée en blocs, et pour chaque bloc une technique de  $\theta$ -redressement basé ligne (ou colonne) est appliquée selon cinq orientations :  $\frac{3\pi}{4}$ ,  $\frac{5\pi}{8}$ ,  $\frac{\pi}{2}$ ,  $\frac{3\pi}{8}$  et  $\frac{\pi}{4}$  (ou respectivement  $-\frac{\pi}{4}$ ,  $-\frac{\pi}{8}$ ,  $0$ ,  $\frac{\pi}{8}$ ,  $\frac{\pi}{4}$ ). L'énergie du signal de la sous-bande des hautes fréquences spatiales produite pour chacune de ses orientations est calculée. L'orientation qui produit l'énergie minimale est choisie comme première orientation de filtrage si cette énergie est inférieure à l'énergie produite par un filtrage selon la direction verticale (respectivement horizontale) pondérée par un facteur compris entre zéro et un (dans l'article considéré ; il vaut 0,9). Cette dernière comparaison s'explique par le codage de l'information d'orientation. Il prend en compte qu'un choix différent de l'orientation par défaut engendre un surcoût qui doit être justifié par une diminution significative de l'énergie des hautes fréquences.

La deuxième orientation de filtrage est choisie en répétant le procédé sur la sous-bande des basses fréquences spatiales produite après un filtrage complet de l'image selon la première dimension.

Autre exemple, pour la « Direction-Adaptive Discrete Wavelet Transform » [CG07], la première orientation de filtrage choisie est celle minimisant une fonction de coût modélisant un compromis débit-distorsion faisant intervenir l'énergie du signal de la sous-bande des hautes fréquences spatiales produite après un Mojette- $(p, q)$ -redressement et le coût de codage de l'information d'orientation et de segmentation. La direction  $(p, q)$  est alors choisie parmi l'ensemble :  $\{(-3, 1), (-2, 1), (-1, 1), (-1, 3), (0, 1), (1, 3), (1, 1), (2, 1), (3, 1)\}$ . La seconde direction de filtrage est alors toujours  $(1, 0)$ .

Cette approche permet donc de déterminer l'orientation de filtrage, qui globalement, minimise l'énergie des hautes fréquences spatiales dans cette direction. Cela soulève cependant un problème et une interrogation.

Un problème : appliquer un certain nombre de fois une transformée en ondelettes orientée a un coût calculatoire non négligeable. Nous étudierons donc tout d'abord d'autres métriques moins coûteuses basées sur un simple calcul de l'erreur de prédiction dans un voisinage réduit  $3 \times 3$  avec interpolation linéaire.

Une interrogation : si cette méthode a de fortes chances de choisir la « meilleure » orientation de filtrage dans le cas d'un contenu mono-orienté, qu'en est-il d'un contenu multi-orienté, c'est-à-dire représentant plusieurs contours ? Dans le cas où une image présente plusieurs orientations privilégiées, on observe rarement plusieurs minimums d'énergie. Pire que cela, à deux orientations privilégiées assez proches mais correspondant à des contours vraiment différents, le minimum d'énergie fera correspondre une orientation ne correspondant à aucune de ces deux orientations.

Si l'on reformule la question : faut-il se satisfaire d'un tel résultat ? D'une certaine façon oui, car l'orientation choisie est bien celle qui minimise globalement l'énergie des hautes fréquences spatiales dans la direction de filtrage choisie. Toutefois, si l'énergie du signal de la sous-bande des hautes fréquences spatiales reconstruite est globalement plus basse, tous les contours sont présents. Dès lors il peut être jugé plus souhaitable de minimiser l'énergie de façon à faire disparaître le plus possible les termes d'énergie relatifs à un contour particulier d'une région, par exemple celui considéré comme le plus « fort », quitte à être moins performant pour les autres. Une mesure trop globale n'est dans ce cas pas efficace, et doit donc être modifiée pour prendre en compte la pertinence locale de l'information. La régularité en un point est d'autant plus pertinente par exemple que le gradient est fort, ou qu'elle est faible par rapport aux autres orientations. Une telle mesure sera proposée de façon à extraire plusieurs orientations privilégiées d'une région d'une image.

### 3.2.2.1 Minimisation d'une mesure de régularité liée à l'erreur de prédiction du $\theta$ -redressement

Plutôt que de calculer successivement pour chaque orientation de filtrage la transformée orientée, puis son inverse en ne considérant que les coefficients relatifs aux hautes fréquences, et l'énergie de l'image ainsi reconstruite, on se propose de ne calculer que les

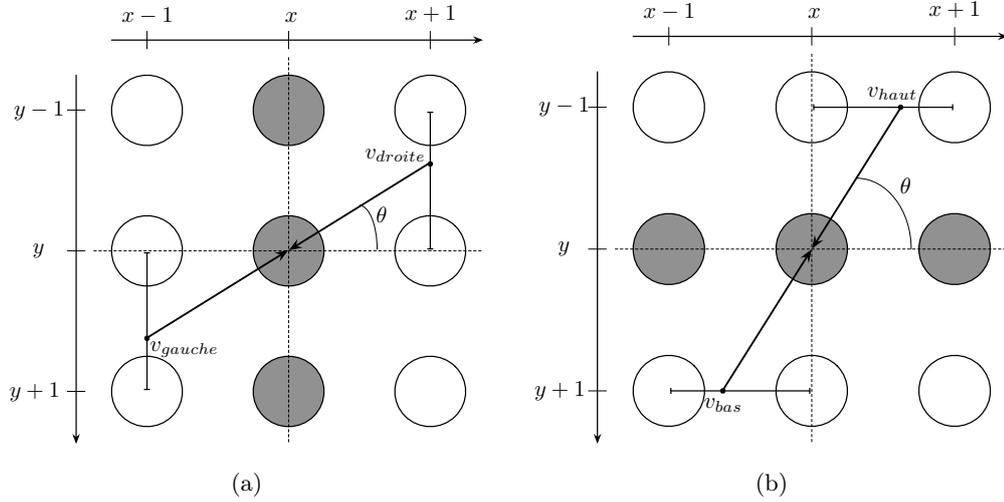


FIGURE 3.3 –  $ep_{gauche/haut}^\theta$  et  $ep_{droite/bas}^\theta$  représentent respectivement la différence entre l'élément central de coordonnées  $(x, y)$  et l'élément  $v_{gauche}$  ou  $v_{haut}$ , respectivement  $v_{droite}$  ou  $v_{bas}$ , dont la valeur est interpolée.

erreurs de prédiction correspondant à l'unique étape de prédiction (d'un  $\theta$ -redressement) utilisant les filtres d'ondelettes 5/3 et les filtres de prédiction et de mise à jour  $L1$ . On définit pour cela tout d'abord les erreurs de prédiction suivantes :

$$ep_{gauche/haut}^\theta(x, y) = v(x, y) - \begin{cases} v(x-1, y - \tan \theta) & \text{si } -\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4} \\ v(x - \tan(\frac{\pi}{2} - \theta), y-1) & \text{si } \frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4} \end{cases} \quad (3.5)$$

$$ep_{droite/bas}^\theta(x, y) = v(x, y) - \begin{cases} v(x+1, y + \tan \theta) & \text{si } -\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4} \\ v(x + \tan(\frac{\pi}{2} - \theta), y+1) & \text{si } \frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4} \end{cases} \quad (3.6)$$

avec  $v(x, y)$  la valeur de l'image au point de coordonnées  $(x, y)$  (voir figure 3.3).

L'erreur de prédiction qui nous intéresse correspond alors à :

$$ep^\theta = \frac{ep_{gauche/haut}^\theta + ep_{droite/bas}^\theta}{2} \quad (3.7)$$

Toutefois il a déjà été fait état dans le chapitre précédent qu'un filtrage horizontal pouvait produire une énergie plus faible qu'un filtrage strictement dans la direction d'un contour pour la raison que les termes  $ep_{gauche/haut}^\theta$  et  $ep_{droite/bas}^\theta$  peuvent se compenser mutuellement. Dans le but de déterminer l'orientation qui minimise plutôt la différence de valeur de l'image pixel à pixel, il convient d'introduire l'erreur de prédiction modifiée suivante :

$$epm^\theta = \frac{|ep_{gauche/haut}^\theta| + |ep_{droite/bas}^\theta|}{2} \quad (3.8)$$

Notons que cette erreur de prédiction risque de nous éloigner du choix de l'orientation de filtrage qui minimise l'énergie globale après reconstruction, en espérant qu'elle nous permette de déterminer l'orientation du ou des contours de l'image.

Pour chacune de ces deux erreurs, deux métriques sont alors considérées : l'erreur absolue moyenne (EAM) et l'erreur quadratique moyenne (EQM) :

$$EAM_{ep}^{\theta} = \frac{1}{\text{card}(\text{image})} \sum_{(x,y) \in \text{image}} |ep^{\theta}(x,y)| \quad (3.9)$$

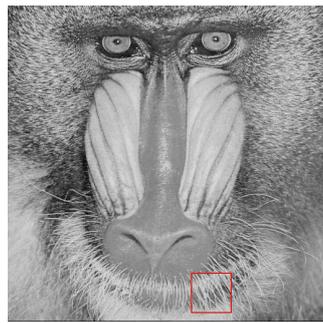
$$EQM_{ep}^{\theta} = \frac{1}{\text{card}(\text{image})} \sum_{(x,y) \in \text{image}} [ep^{\theta}(x,y)]^2 \quad (3.10)$$

**3.2.2.1.1 Contours mono-orientés :** Les résultats de la technique du  $\theta$ -redressement sur des contours mono-orientés présentés dans le chapitre précédent sont ici repris pour être comparés avec nos quatre métriques d'erreurs de prédiction. Pour chacune des figures 3.5 à 3.13, le graphique de gauche représente l'erreur absolue moyenne des erreurs de prédiction et celle des erreurs de prédiction modifiées comparées à l'énergie moyenne des hautes fréquences spatiales dans la direction de filtrage, le graphique de droite représente quant à lui l'erreur quadratique moyenne des erreurs de prédiction et celle des erreurs de prédiction modifiées comparées elles aussi à l'énergie moyenne des hautes fréquences spatiales.

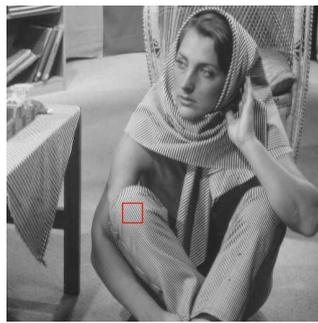
L'observation des différentes figures permettent deux observations préliminaires. Tout d'abord l'EAM et l'EQM des erreurs de prédiction  $pe$  semblent fortement corrélées à l'énergie moyenne des hautes fréquences spatiales dans la direction de filtrage pour des orientations de filtrage proches de l'orientation de filtrage qui minimise cette énergie moyenne. De plus si la variation des métriques basées sur les erreurs de prédiction  $ep$  pour de telles orientations est faible tout comme la variation de l'énergie, les variations de l'EAM et l'EQM basées sur les erreurs de prédiction modifiées  $epm$  sont quant à elles beaucoup plus fortes.

Trois comportements sont alors observés. Dans un premier cas, les minimums des différentes métriques diffèrent peu ou prou avec les minimums de l'énergie moyenne (figures 3.5, 3.12 et 3.13).

Dans un deuxième cas, les minima des métriques basées sur l' $epm$  diffèrent de ceux de l'énergie moyenne, mais l'énergie moyenne ainsi que visuellement le contenu de la sous-bande des hautes fréquences spatiales selon l'orientation de filtrage reconstruite ne varie que très faiblement (figures 3.6, 3.7, 3.8, 3.9, 3.10). Quant à la valeur de l'orientation qui minimise les métriques basée sur l' $epm$ , elle donne une meilleure estimation de l'orientation du contour représenté, particulièrement observable sur les figures 3.8, 3.9 et 3.10. Dans ces trois cas l'orientation de filtrage qui minimise l'énergie moyenne de la sous-bande des hautes fréquences dans la direction de filtrage reconstruite est  $90^{\circ}$ . Les mesures se basant sur l'erreur de prédiction modifiée  $epm$  ont permis de privilégier l'orientation de régularité au détriment de l'orientation verticale qui bénéficiait de l'effet d'annulation mutuelle des valeurs des voisins prédites.



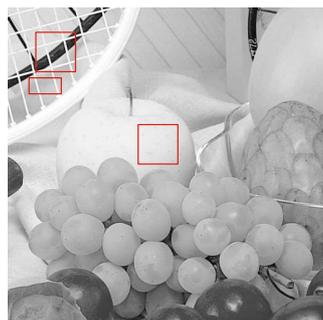
(a) baboon



(b) barbara



(c) boats



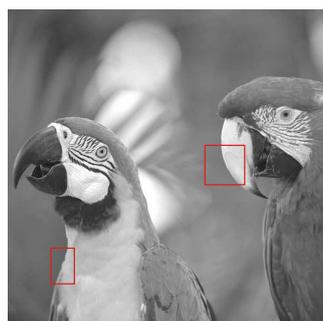
(d) fruits



(e) lena



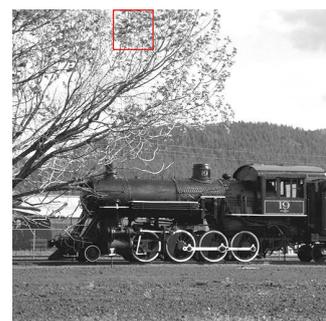
(f) monarch



(g) parrots



(h) plane



(i) train

FIGURE 3.4 – Imagettes choisies au sein des différentes images de test.

Le troisième cas, représenté par la figure 3.11, est un peu particulier car contrairement aux autres exemples, le contour du bec du perroquet est sensiblement courbe. L'orientation qui minimise les métriques basées sur l'*epm* est comprise entre  $90^\circ$  et  $45^\circ$  ce qui est cohérent avec l'orientation « moyenne » du contour. L'orientation qui minimise l'énergie moyenne est quant à elle de  $90^\circ$ . Pour les mêmes raisons que précédemment, l'orientation du contour est mieux estimée par les métriques basées *epm*, par contre ici le contenu de la sous-bande des hautes fréquences spatiales selon la direction de filtrage reconstruite change sensiblement entre les orientations  $90^\circ$  et  $67,5^\circ$ . Sur la figure 3.11-b qui correspond à un filtrage à  $90^\circ$ , la structure relative au morceau du contour d'orientation proche de la verticale n'est pas présente, tandis que la structure relative au morceau du contour d'orientation proche de  $45^\circ$  est bien présente. Au contraire, sur la figure 3.11-c qui correspond à un filtrage à  $67,5^\circ$ , la structure relative au contour du bec du perroquet est présente dans toute la sous-bande reconstruite mais de dynamique plus faible que sur la figure 3.11-b pour le morceau de contour d'orientation proche de  $45^\circ$ .

En conclusion, pour des contenus mono-orientés, l'orientation qui minimise les métriques basées sur les erreurs de prédiction modifiées est une meilleure estimation de l'orientation réelle de ce contour qui ne correspond pas nécessairement avec l'orientation de filtrage qui minimise l'énergie de la sous-bande des hautes fréquences dans la direction de filtrage reconstruite. Ceci est positif du point de vue de l'estimation de l'orientation du contenu à filtrer. Toutefois on peut ne pas être satisfait que une meilleure estimation de l'orientation d'un contour conduise à une énergie moyenne des pixels de la sous-bandes des hautes fréquences non minimale. On peut être rassuré par le fait que la variation d'énergie est faible, même si ce n'est qu'un simple corollaire de l'inefficacité de l'interpolation du  $\theta$ -redressement.

**3.2.2.1.2 Contenu multi-orientés :** Les mesures basées sur les erreurs de prédiction modifiées permettent une estimation plus précise de l'orientation de contours d'un bloc d'image. Nous allons appliquer maintenant ces mesures sur des blocs représentant plusieurs contenus à différentes orientations. Les résultats sont présentés sur les figures 3.14, 3.15, 3.16 et 3.17.

Le premier bloc représente principalement deux mâts, l'un orienté selon un angle proche mais plus petit que  $90^\circ$ , l'autre selon un angle proche mais plus petit que  $135^\circ$ . Les différentes mesures sont minimisées globalement pour une orientation proche de  $90^\circ$ , par contre aucun minimum même local n'apparaît pour des orientation plus proches de  $135^\circ$  que de  $90^\circ$  (voir figure 3.14). Un ou deux minima locaux peuvent être observés à proximité de  $0^\circ$  selon les métriques. Ces minimums locaux pourraient-ils être attribués aux cordages d'orientation proche de l'horizontalité ou pourrait-il être attribué à un quelconque bruit ?

Le bloc extrait de l'image **fruits** contient distinctement trois contenus orientés. L'un plutôt vertical, d'orientation inférieure à  $90^\circ$ , et deux plutôt horizontaux, d'orientation inférieure à  $0^\circ$  pour le premier, et proche de  $-45^\circ$  pour le second. Les résultats présentés sur la figure 3.15 font au contraire état dans le pire des cas de deux minima aux alentours de  $100^\circ$  et un seul à  $0^\circ$ .

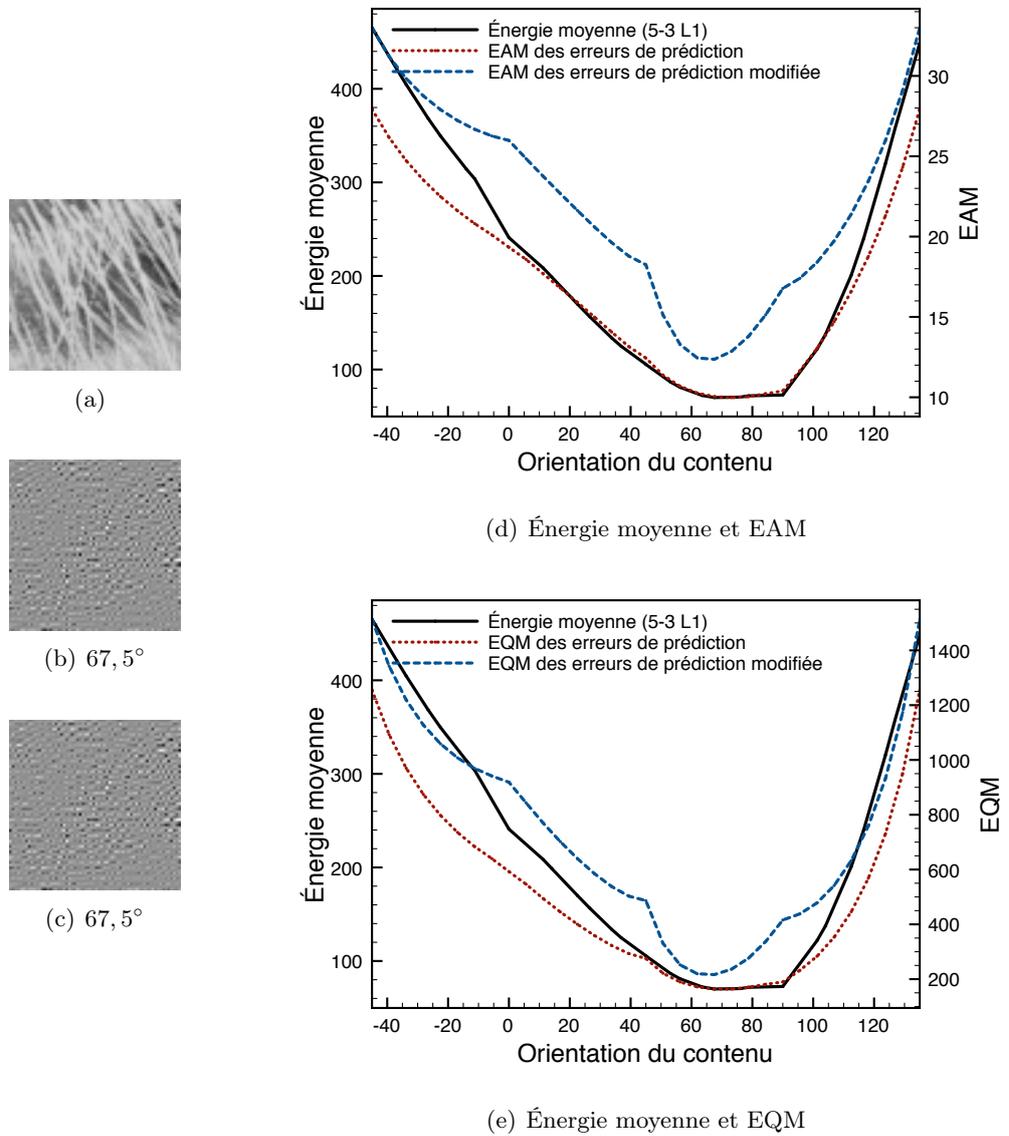


FIGURE 3.5 – Énergie de l'imagette `baboon_mouth_hair` (a) de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ[$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

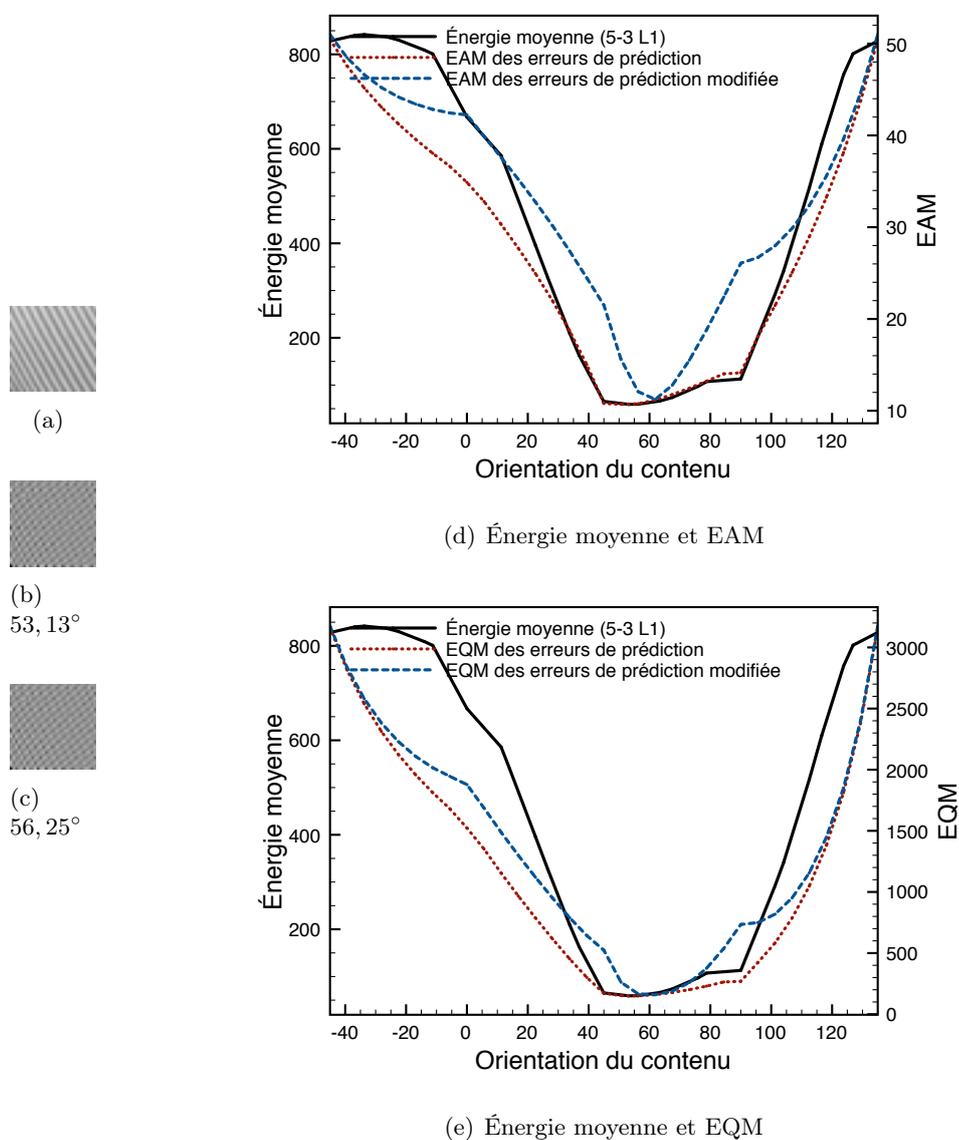


FIGURE 3.6 – Énergie de l'imagette `barbara_pants` (a) de dimension  $32 \times 32$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement, pour trois techniques de redressement, en fonction de l'orientation de filtrage sur  $[-45^\circ, 135^\circ[$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.



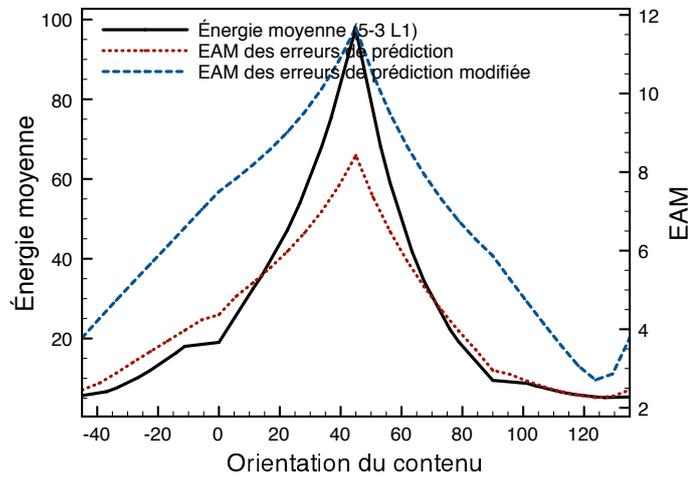
(a)



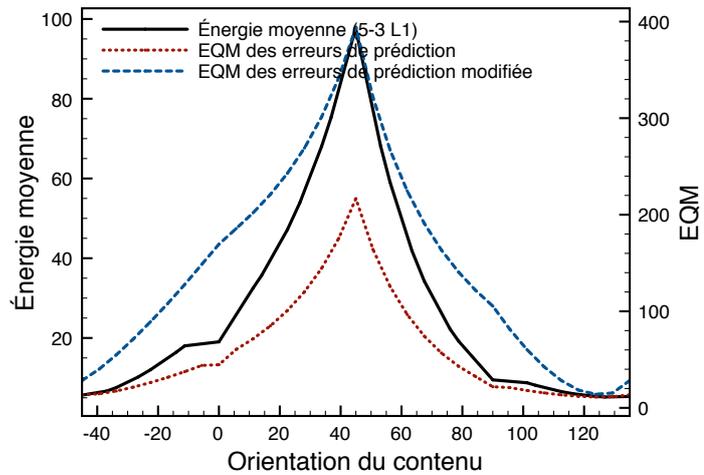
(b)  
126, 87°



(c)  
123, 75°



(d) Énergie moyenne et EAM



(e) Énergie moyenne et EQM

FIGURE 3.7 – Énergie de l'imagette `boats_mast` (a) de dimension  $40 \times 40$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

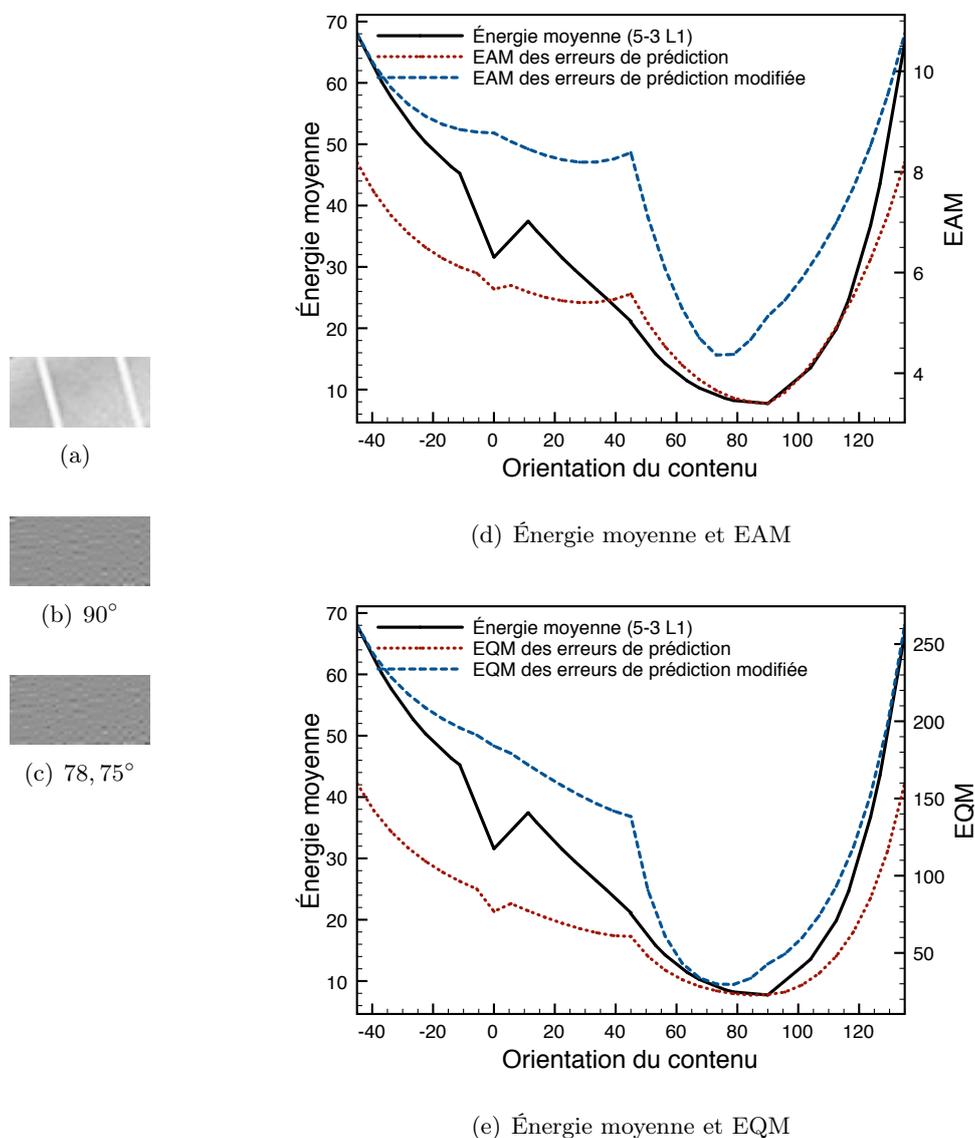


FIGURE 3.8 – Énergie de l'imagette `fruits_racket` (a) de dimension  $52 \times 26$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

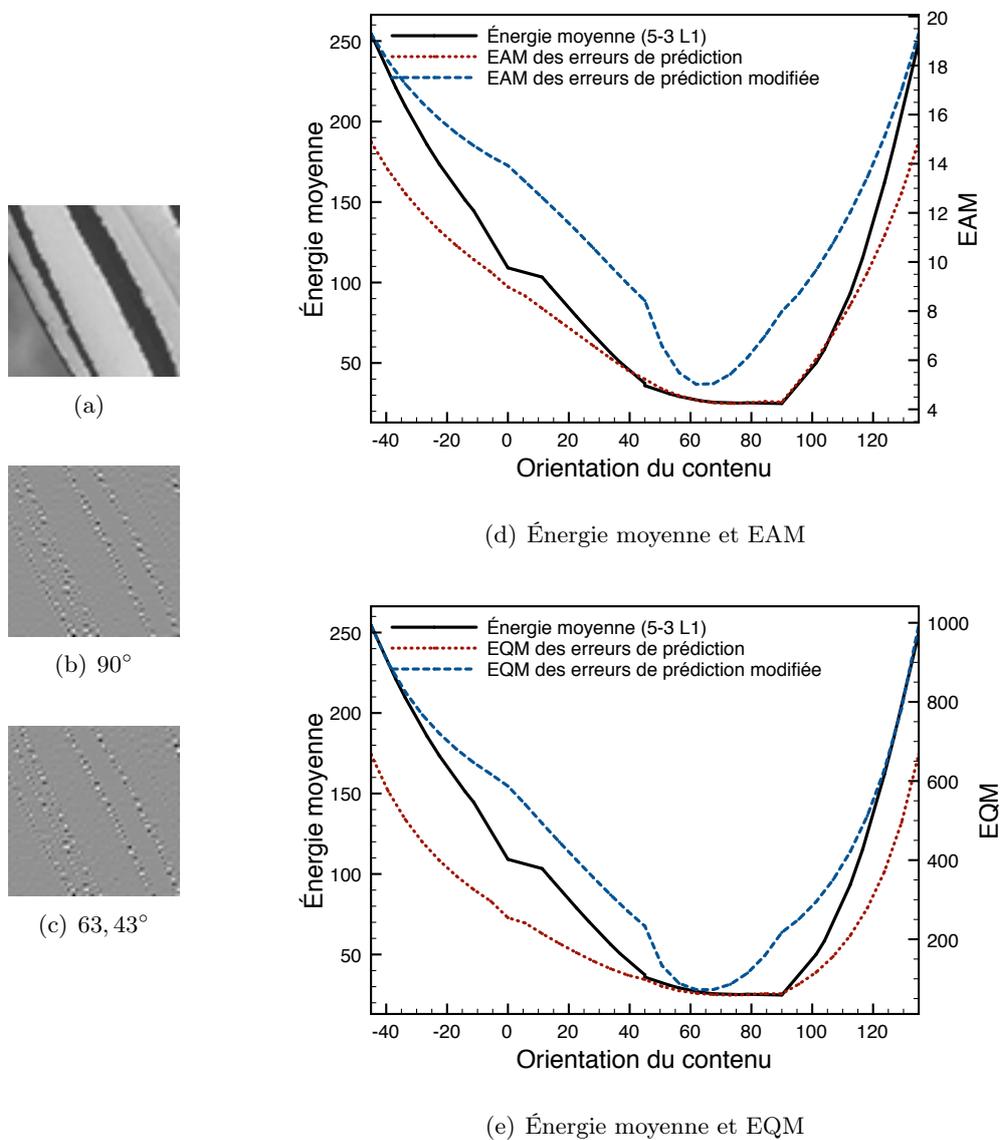


FIGURE 3.9 – Énergie de l'imagette `monarch_wing` (a) de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

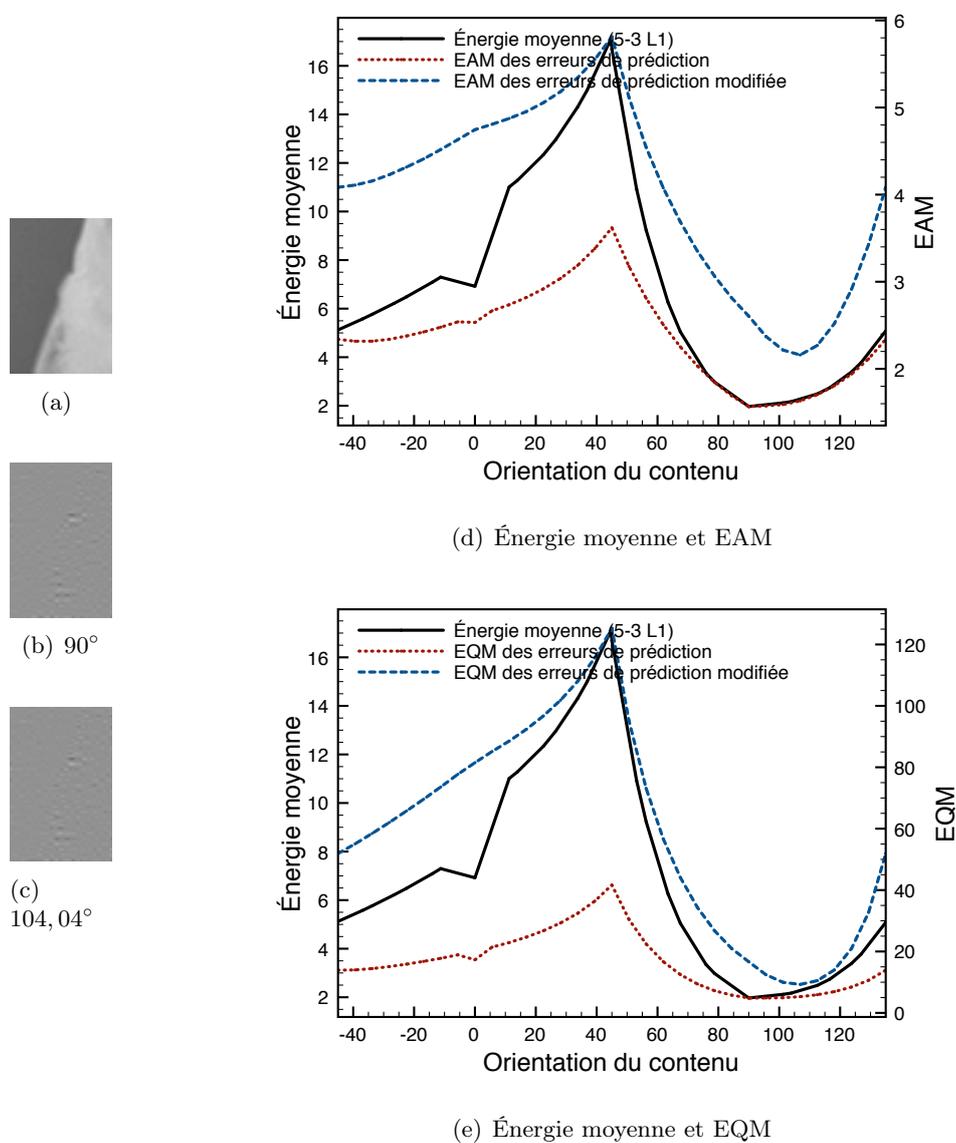


FIGURE 3.10 – Énergie de l’image `parrots_fether` (a) de dimension  $38 \times 58$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l’orientation de filtrage, et mesures d’erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l’EAM des erreurs de prédiction respectivement non modifiée et modifiée.

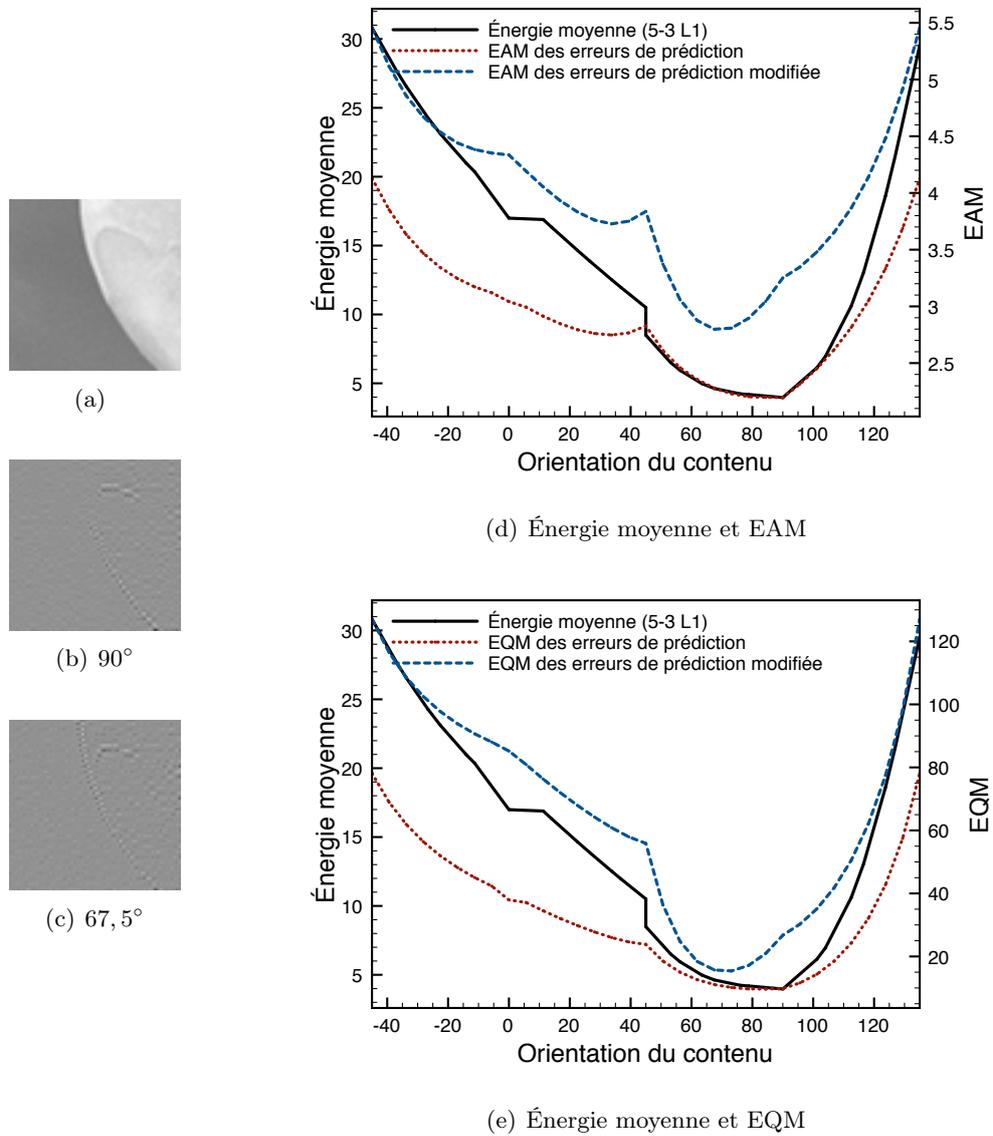


FIGURE 3.11 – Énergie de l'imagette `parrots_nose` (a) de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

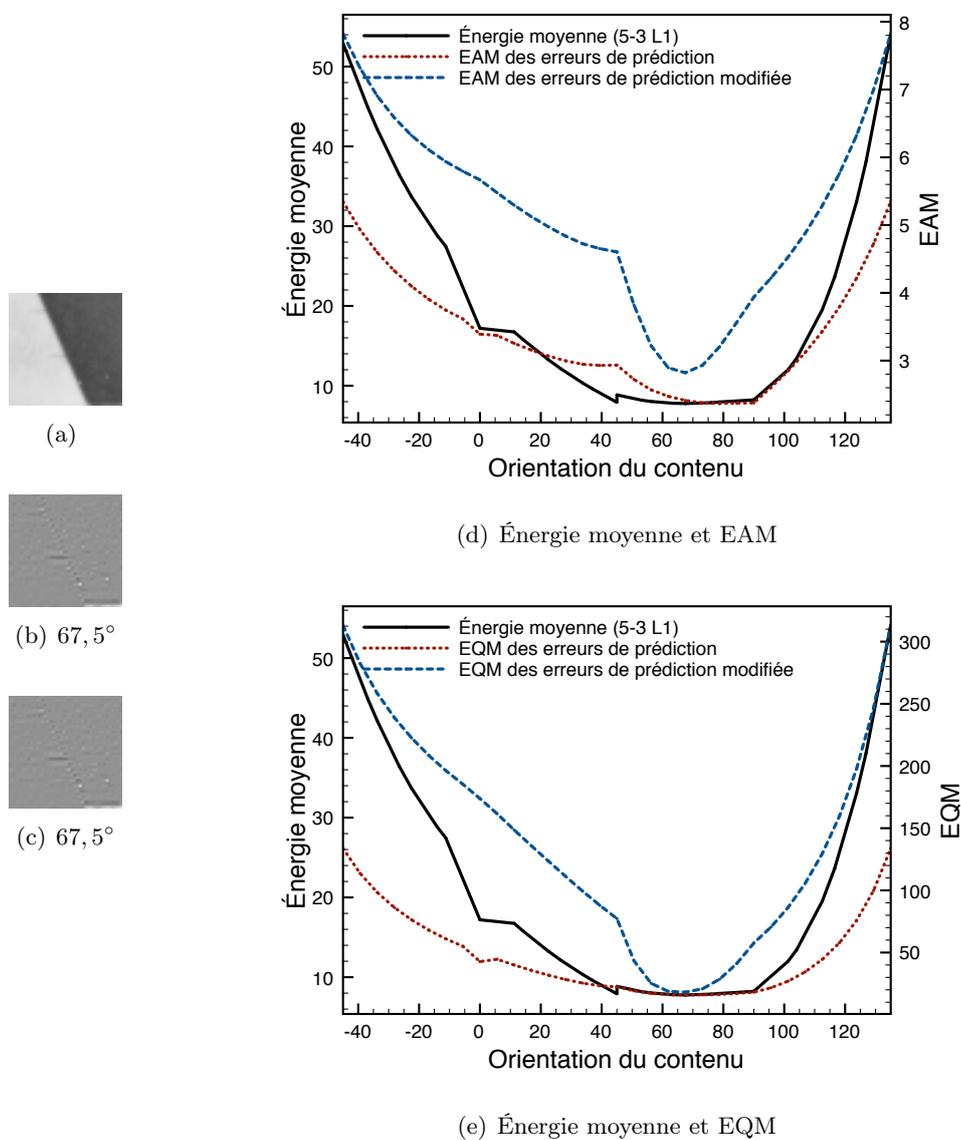


FIGURE 3.12 – Énergie de l’image `plane_wing_back` (a) de dimension  $42 \times 42$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l’orientation de filtrage, et mesures d’erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l’EAM des erreurs de prédiction respectivement non modifiée et modifiée.

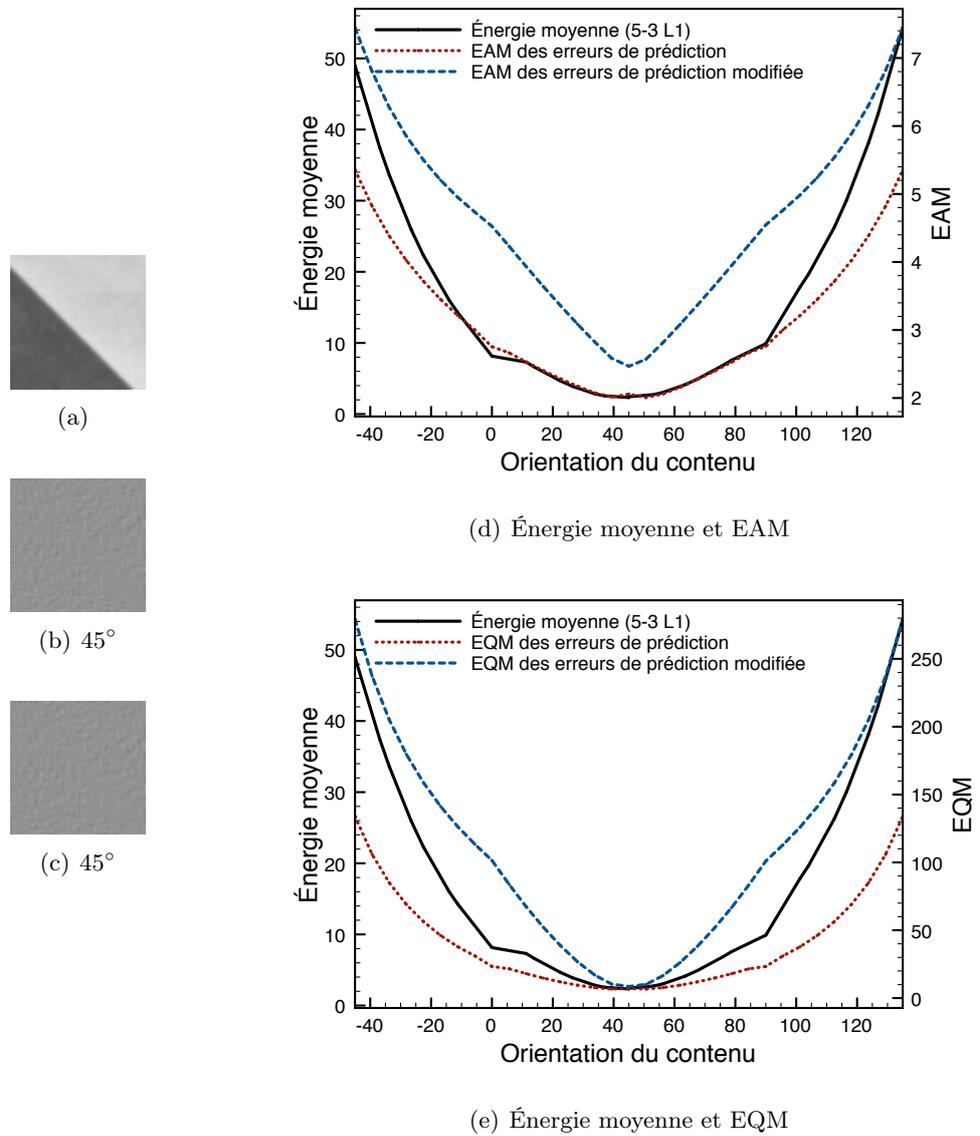


FIGURE 3.13 – Énergie de l'imagette `plane_wing_front` (a) de dimension  $50 \times 50$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ[$ . (b) et (c) représentent les sous-bandes reconstruites des hautes fréquences dans la direction de filtrage qui minimise l'EAM des erreurs de prédiction respectivement non modifiée et modifiée.

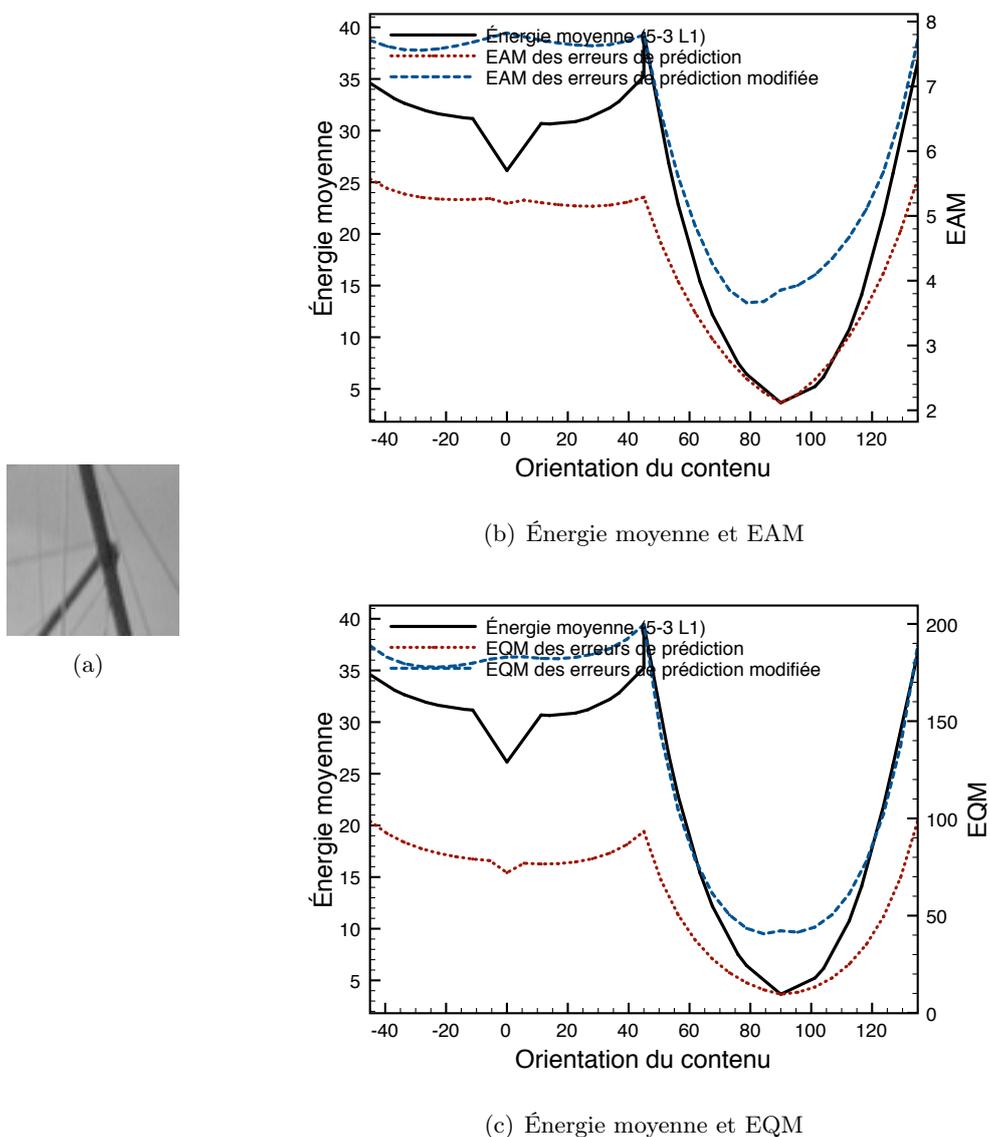
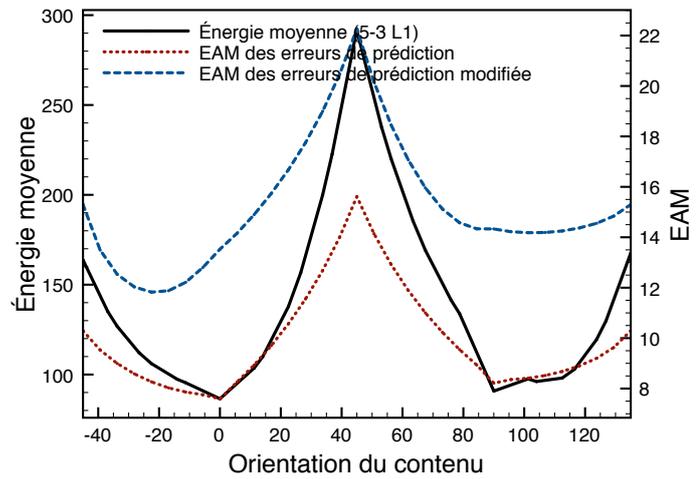


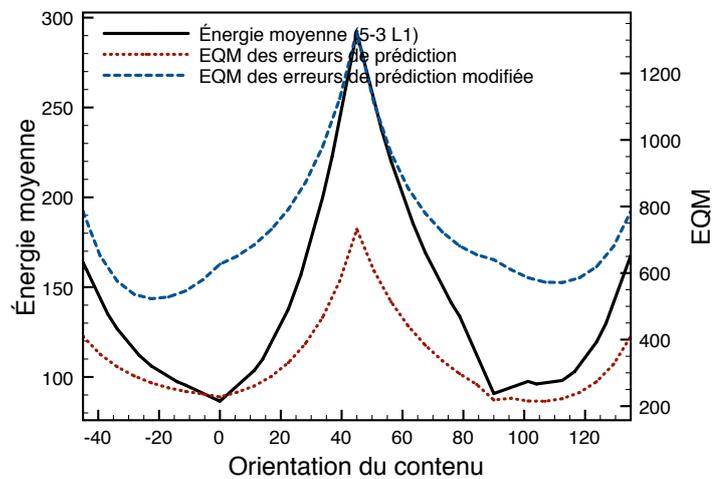
FIGURE 3.14 – Énergie de l’image `boats_masts` de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l’orientation de filtrage, et mesures d’erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .



(a)



(b) Énergie moyenne et EAM



(c) Énergie moyenne et EQM

FIGURE 3.15 – Énergie de l'imagette `fruits_rackets` de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .

Le même type d'observations peut être effectué sur les résultats présentés sur les figures 3.16 et 3.17. Le nombre ou la valeur des minima des différentes mesures ne correspondent pas aux orientations des différents contenus.

**3.2.2.1.3 Contenus isotropes :** L'étude de contenus ne présentant *a priori* pas de direction privilégiée, permet de mettre en évidence les résultats, parfois inattendus, obtenus pour les contenus multi-orientés précédemment étudiés. Les figures 3.18 et 3.19 représentent des zones relativement uniformes, avec une légère structure horizontale pour le nuage de la première figure, et des structures ponctuelles sur la peau de pomme pour la seconde. La figure 3.20 représente un feuillage que l'on peut considérer comme une zone fortement texturée. Les résultats pour ces trois blocs sont similaires, mis à part les valeurs des mesures : très faibles pour les zones uniformes, élevées pour la zone texturée.

Contrairement à ce qu'on pourrait souhaiter, les valeurs prises par les différentes mesures varient en fonction de l'orientation. Deux facteurs permettent de douter du fait que ces variations viendraient seulement de la nature géométrique des images. D'une part les variations sont d'amplitude plus forte pour la zone texturée que pour les zones uniformes, donc on peut imaginer que ces variations soient fonction de la norme du gradient. D'autre part les trois résultats présentent une certaine similarité dans la structure même de ces variations. On peut dès lors imaginer un biais introduit par la transformée elle-même.

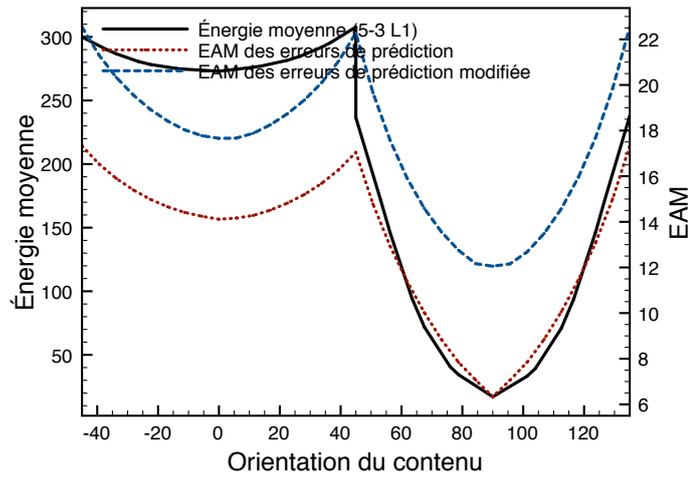
**3.2.2.1.4 Mesures isotropes** L'étude des différentes mesures ou de l'énergie moyenne des éléments de la sous-bandes des hautes fréquences spatiales selon la direction de filtrage reconstruite sur des contenus isotropes a permis la mise en évidence d'un biais introduit par ces mesures elles-mêmes ou le  $\theta$ -redressement. Ce biais est dû à l'interpolation et introduit trois biais différents.

**Énergie des réponses impulsionnelles des filtres d'ondelettes orientées variable selon l'orientation :** Comme il a été énoncé au chapitre précédent, l'énergie des réponses impulsionnelles des filtres d'ondelettes orientées correspondant à la technique de  $\theta$ -redressement varie en fonction de l'orientation de filtrage. Elle varie plus précisément en fonction de l'interpolation. Plus la pondération de l'interpolation linéaire est proche de  $(\frac{1}{2}, \frac{1}{2})$ , i.e.  $\theta = \arctan \frac{1}{2}, -\arctan \frac{1}{2}, \frac{\pi}{2} - \arctan \frac{1}{2}$ , ou  $\frac{\pi}{2} + \arctan \frac{1}{2}$ , plus l'énergie de la réponse impulsionnelle sera faible (la somme des éléments de la réponse impulsionnelle restant constante).

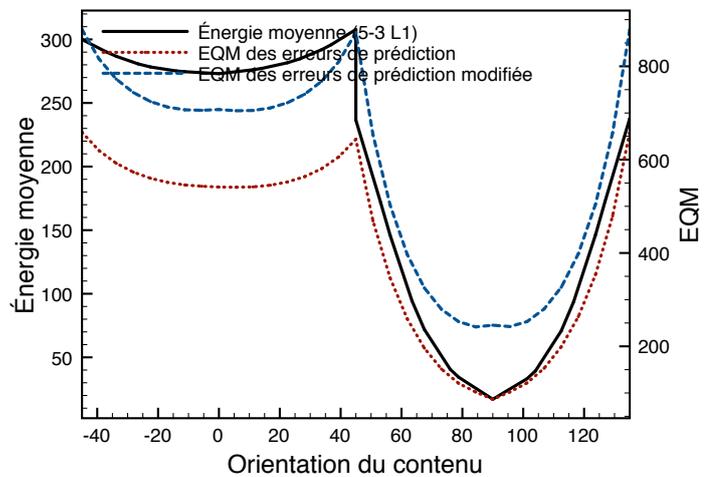
**Erreur d'interpolation :** L'erreur d'interpolation est, dans le cas de l'interpolation linéaire, statistiquement maximale pour une pondération  $(\frac{1}{2}, \frac{1}{2})$ . Elle est potentiellement d'autant plus forte que la variation locale est élevée. À titre d'illustration, pour les signaux sinusoïdaux de la figure 3.21, l'*EAM* et l'*EQM* des erreurs de prédiction  $ep$  et  $epm$  sont bien maximales pour les orientations  $\arctan(-\frac{1}{2}), \arctan(\frac{1}{2}), \frac{\pi}{2} - \arctan(\frac{1}{2})$



(a)



(b) Énergie moyenne et EAM

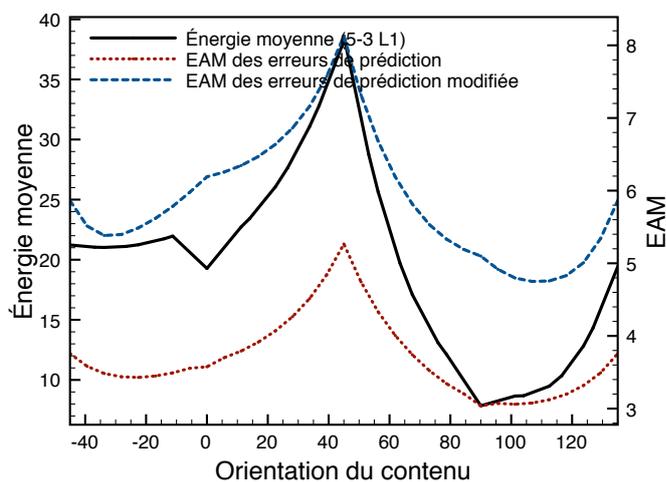


(c) Énergie moyenne et EQM

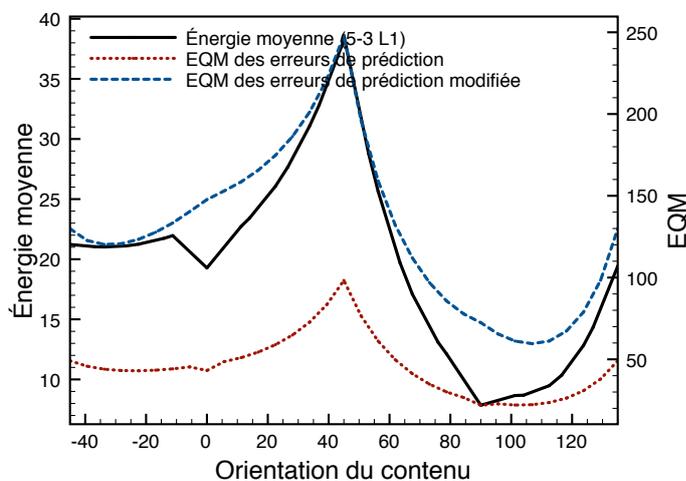
FIGURE 3.16 – Énergie de l'imagette `boats_text` de dimension  $32 \times 32$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .



(a)



(b) Énergie moyenne et EAM



(c) Énergie moyenne et EQM

FIGURE 3.17 – Énergie de l'imagette `lena_hat` de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .

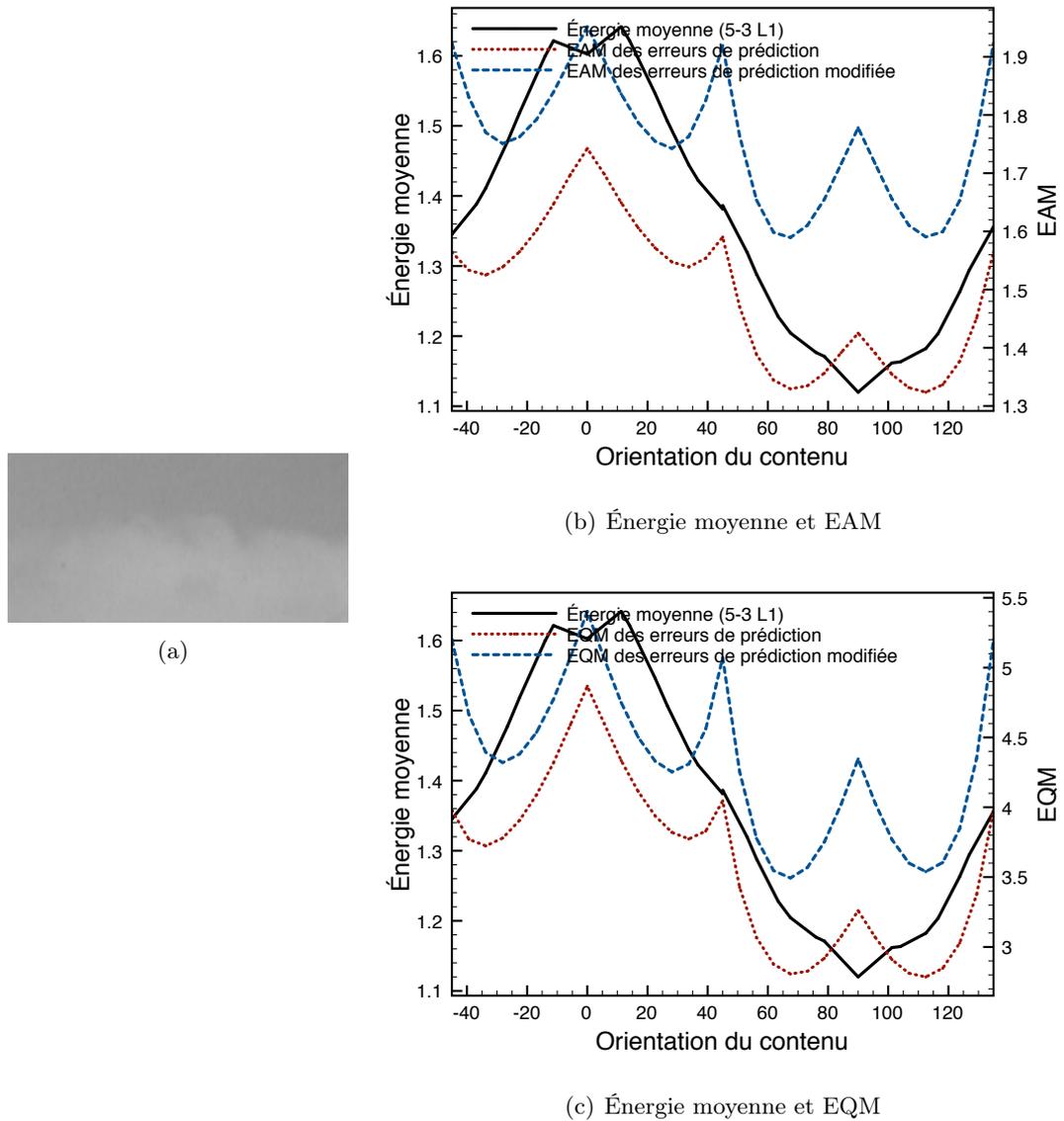
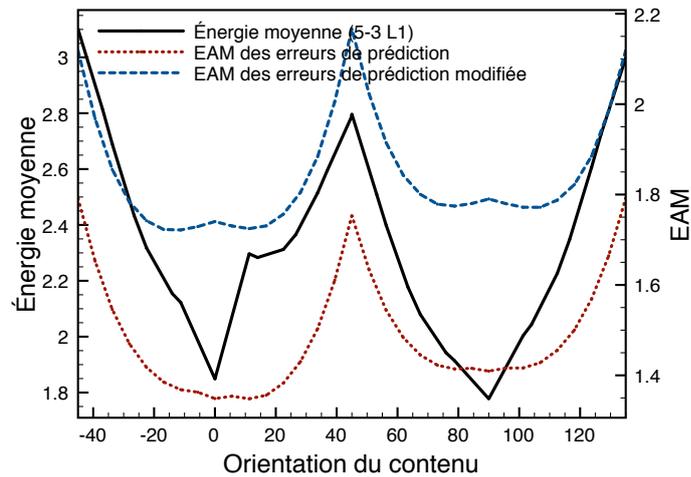


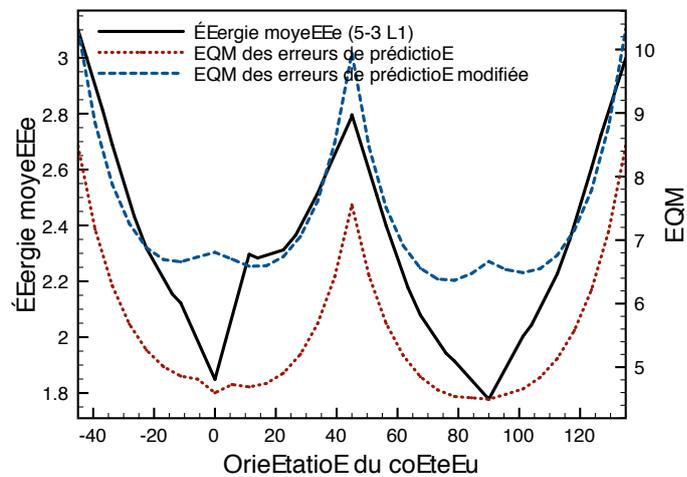
FIGURE 3.18 – Énergie de l’image `boats_sky_and_clouds` de dimension  $128 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l’orientation de filtrage, et mesures d’erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .



(a)



(b) Énergie moyenne et EAM

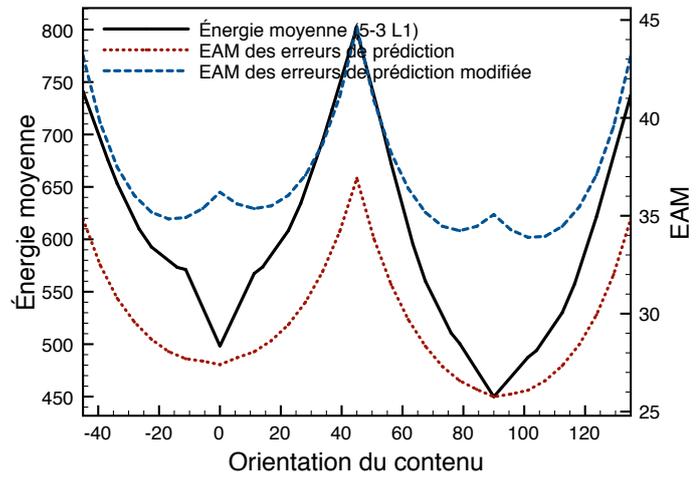


(c) Énergie moyenne et EQM

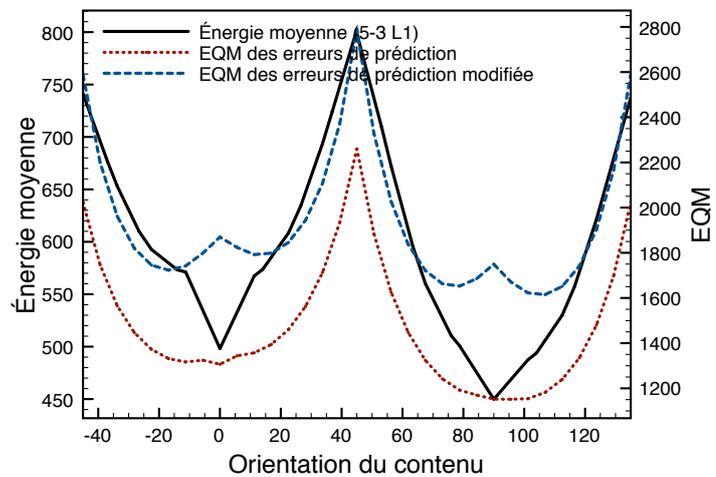
FIGURE 3.19 – Énergie de l'imagette `fruits_apple` de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .



(a)



(b) Énergie moyenne et EAM



(c) Énergie moyenne et EQM

FIGURE 3.20 – Énergie de l'imagette `train_leafs` de dimension  $64 \times 64$ , reconstruite à partir des coefficients des hautes fréquences spatiales uniquement en fonction de l'orientation de filtrage, et mesures d'erreurs de prédiction selon 32 orientations uniformément réparties sur  $[-45^\circ, 135^\circ]$ .

et  $\frac{\pi}{2} + \arctan\left(-\frac{1}{2}\right)$ , et leur valeur est bien décroissante avec l'augmentation de la période des signaux sinusoïdaux.

**Lissage du filtre d'interpolation :** Le filtre d'interpolation linéaire introduit dans les différentes mesures un filtrage passe-bas variant en fonction de la pondération de l'interpolation, et donc par conséquent de l'orientation considérée. L'effet de lissage est alors maximal pour la pondération  $\left(\frac{1}{2}, \frac{1}{2}\right)$ . Sur la figure 3.22, l'*EAM* et l'*EQM* des erreurs de prédiction *ep* et *epm* sont calculées pour deux images uniformes de niveau de gris 128 sur lesquelles a été ajouté un bruit blanc gaussien de moyenne nulle et de variance 2,5 pour la première image et 8 pour la deuxième. Les *PSNR* de ces deux images, avec comme référence une image uniforme de niveau de gris 128, sont respectivement 40.02 dB et 29,92 dB. La valeur des différentes mesures est minimale pour des pondérations  $\left(\frac{1}{2}, \frac{1}{2}\right)$  de l'interpolation, c'est-à-dire pour les orientations  $\arctan\left(-\frac{1}{2}\right)$ ,  $\arctan\left(\frac{1}{2}\right)$ ,  $\frac{\pi}{2} - \arctan\left(\frac{1}{2}\right)$  et  $\frac{\pi}{2} + \arctan\left(-\frac{1}{2}\right)$ . La variation des valeurs des métriques est d'autant plus forte que le bruit est important.

**3.2.2.1.5 Conclusion :** L'objectif premier des mesures étudiées est de déterminer l'orientation de filtrage d'une région d'une image, c'est-à-dire l'orientation qui minimise l'énergie moyenne des éléments de la sous-bande des hautes fréquences spatiales dans la direction de filtrage reconstruite, avec un coût calculatoire nettement moindre. Les mesures basées sur les erreurs de prédiction *ep* répondent bien à cet objectif au travers du minimum global de celles-ci. L'objectif second, mais pas secondaire, est de déterminer l'orientation de filtrage qui minimise le mieux l'énergie relative à la structure orientée la plus significative d'une région. La définition même d'un critère répondant à cet objectif n'est pas aisée. Les mesures étudiées se heurtent à un premier paradoxe de la technique de  $\theta$ -redressement : même pour une région mono-orientée, l'orientation de filtrage qui minimise l'énergie moyenne des hautes fréquences ne correspond pas nécessairement à l'orientation perçue du contour de cette région. Cependant un filtrage selon cette orientation perçue n'introduit pas une augmentation importante de cette même énergie. Il est alors concevable de déterminer les orientations de filtrage d'une région par détermination des orientations perçues dans une telle région.

Les mesures basées sur les erreurs de prédiction modifiées *epm* offrent une meilleure estimation de l'orientation perçue de régions mono-orientées comparées aux mesures basées sur les simples erreurs de prédiction *ep*. Toutes ces mesures sont par contre inefficaces pour déterminer les orientations privilégiées, que ce soit leur valeur ou leur nombre. L'interpolation utilisée introduit plusieurs biais qui rendent incomparables les valeurs des mesures entre elles pour des orientations requérant des pondérations différentes dans l'interpolation. En effet l'erreur d'interpolation et le filtrage passe-bas de l'interpolation ont des effets antagonistes sur la valeur des différentes métriques en fonction de l'orientation.

Ces deux effets impliquent qu'une mesure pour une orientation  $\theta$  n'est pas comparable avec une mesure pour une orientation  $\theta'$  dans le cas où  $\theta$  et  $\theta'$  n'utilisent pas la même pondération dans l'interpolation. L'effet atténuateur de bruit prédomine dans

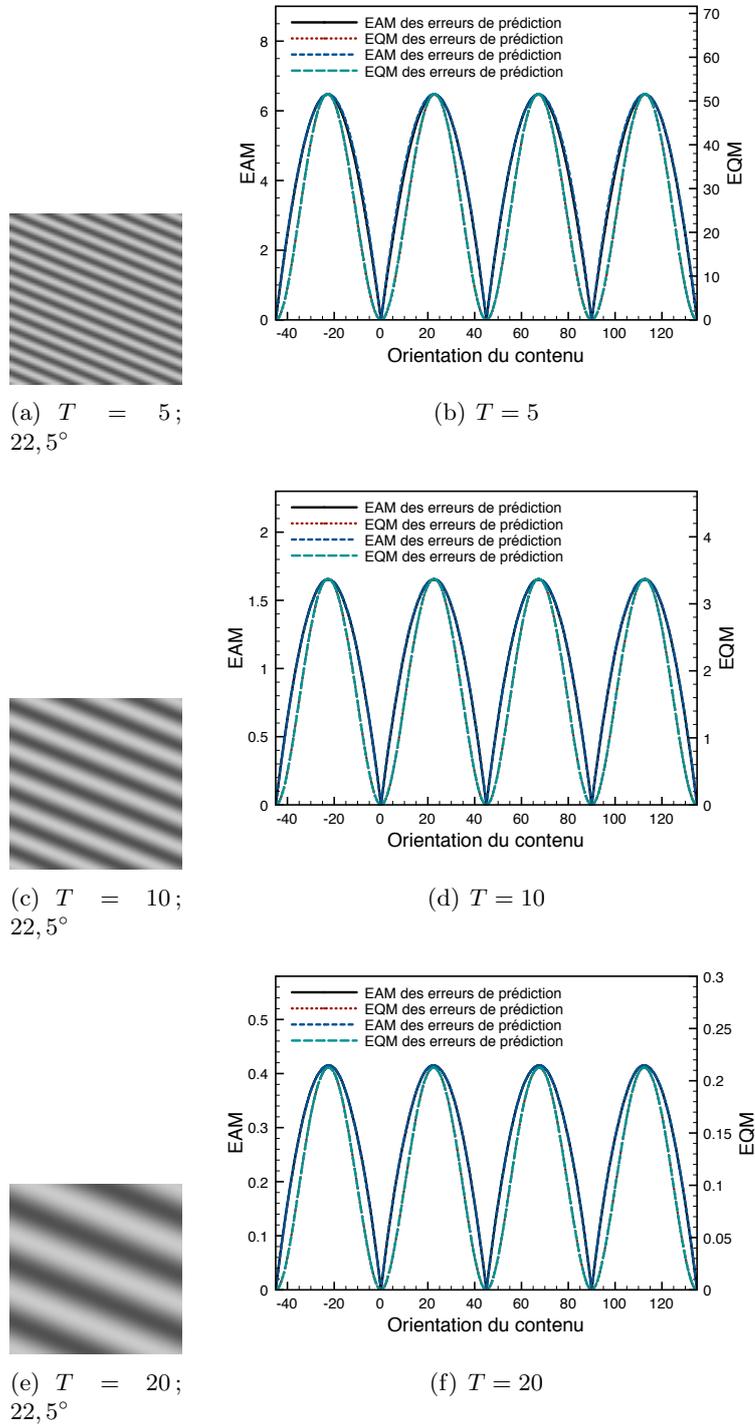
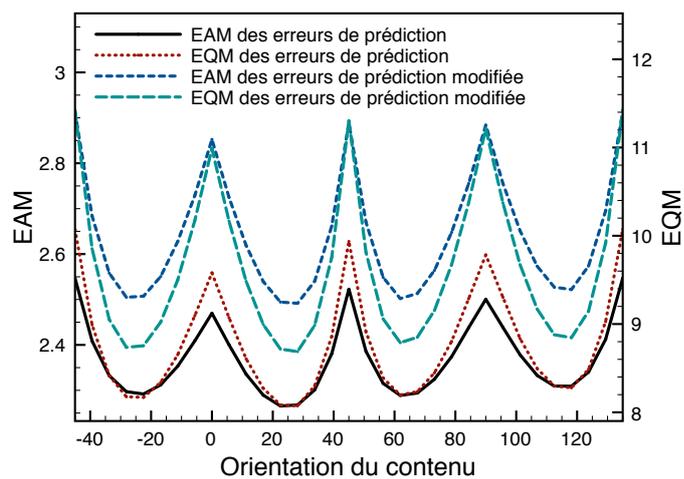
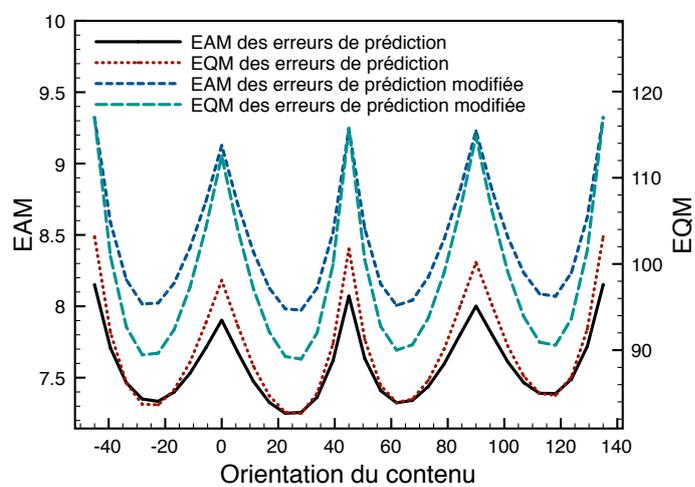


FIGURE 3.21 – Influence de l'erreur d'interpolation en fonction de l'orientation de signaux sinusoidaux sur l'EAM et EQM des erreurs de prédiction  $ep$  et  $epm$  calculées selon l'orientation de ces signaux.

(a)  $PSNR = 40,02dB$ (c)  $PSNR = 40,02dB$ (b)  $PSNR = 29,92dB$ (d)  $PSNR = 29,92dB$ FIGURE 3.22 – Influence du bruit blanc gaussien sur l' $EAM$  et  $EQM$  des erreurs de prédiction  $ep$  et  $epm$ .

les zones uniformes ou texturées sans orientation privilégiée, et provoque donc non pas une réponse uniforme en fonction de l'orientation comme souhaitée, mais une réponse comme celle présentée figure 3.22. L'influence de l'erreur d'interpolation a sur une image naturelle pour effet en pratique d'introduire de faux minima locaux où de déplacer des minimums locaux correspondant réellement à une orientation privilégiée.

### 3.2.2.2 Estimation par modélisation stochastique de l'incertitude sur la mesure de l'orientation du gradient

Afin de déterminer les orientations privilégiées d'une région d'une image, il convient donc de définir à présent une mesure qui soit le plus possible indépendante de l'orientation considérée de façon à pouvoir comparer d'une part les mesures faites selon diverses orientations mais aussi à avoir une réponse constante pour des régions sans orientation particulière. De plus la mesure doit pouvoir différencier l'orientation correspondant à un contour très présent et/ou très contrasté d'un contour court et faiblement contrasté.

La nouvelle mesure s'appuie sur la mesure de l'orientation et de la norme du gradient calculé grâce aux deux filtres :

$$\begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -a & 0 & a \end{bmatrix} \text{ et } \begin{bmatrix} -a & -b & -a \\ 0 & 0 & 0 \\ a & b & a \end{bmatrix}, \quad (3.11)$$

avec  $a = \frac{2-\sqrt{2}}{4}$  et  $b = \frac{\sqrt{2}-1}{2}$ . Contrairement aux filtres de Sobel où  $a = 1$  et  $b = 2$ , la valeur de la norme du gradient pour un échelon analogique idéalement discrétisé de niveaux minimum 0 et maximum 255 orienté selon  $0^\circ$  est la même que pour  $45^\circ$ , à savoir  $\frac{255}{2}$ . La nouvelle mesure s'inspire des travaux de [Vil05] sur l'incertitude de la mesure du gradient. Si le gradient  $\nabla I$  n'est pas le vecteur nul, la probabilité que la direction orthogonale de  $\nabla I(x, y)$  soit  $\theta$  est modélisée par la loi normale, de moyenne  $\theta_I + \frac{\pi}{2}$ , et d'écart-type  $\frac{\sigma}{\|\nabla I\|}$  :

$$P(x, y, \cdot) \sim \mathcal{N}\left(\theta_I + \frac{\pi}{2}, \frac{\sigma}{\|\nabla I(x, y)\|}\right),$$

où  $\sigma$  est l'écart-type de l'erreur sur le gradient,  $\theta_I$  est l'orientation du gradient mesurée par arctan. On a donc :

$$P(x, y, \theta) = \frac{1}{K} \cdot e^{-\frac{1}{2}\left(\text{dist}^2(\theta, \theta_I(x, y) + \frac{\pi}{2}) \cdot \frac{\|\nabla I(x, y)\|^2}{\sigma^2}\right)}. \quad (3.12)$$

avec  $K$  une constante de normalisation, et  $\theta$  une orientation mesurée parmi  $N$  orientations.

Si  $\nabla I = \vec{0}$ ,  $P(x, y, \cdot)$  est une loi uniforme, c'est à dire :

$$P(x, y, \theta) = \frac{1}{N}. \quad (3.13)$$

La mesure  $M_1$ , fonction de l'orientation  $\theta$ , est définie par la moyenne des lois de probabilité  $P(x, y, \theta)$  sur un bloc  $\mathcal{B}$  :

$$M_1(\theta) = \frac{1}{M_{max}} \frac{1}{card(\mathcal{B})} \sum_{(x,y) \in \mathcal{B}} P(x,y,\theta) \quad (3.14)$$

Cette mesure est normalisée par rapport à la valeur maximale atteinte lorsque pour tous les éléments du bloc  $\mathcal{B}$  la norme du gradient est maximale, et que l'orientation du gradient est constante, soit :

$$M_{max} = \frac{1}{card(\mathcal{B})} \sum_{(x,y) \in \mathcal{B}} P_{max}(x,y,\theta) \quad (3.15)$$

$$= \frac{1}{card(\mathcal{B})} \sum_{(x,y) \in \mathcal{B}} \frac{1}{K_{max}} e^{-\frac{1}{2} dist^2(\theta, \theta_I + \frac{\pi}{2}) \cdot \frac{\|\nabla I\|_{max}^2}{\sigma^2}} \quad (3.16)$$

$$= \frac{1}{card(\mathcal{B})} \sum_{(x,y) \in \mathcal{B}} \frac{1}{K_{max}} e^{-\frac{1}{2} dist^2(\theta, \theta_I + \frac{\pi}{2}) \cdot \frac{255^2}{(2\sigma)^2}} \quad (3.17)$$

Cette mesure étant calculée pour un ensemble discret de  $N$  orientations, la probabilité  $P_{max}(x,y,\theta)$  est maximale pour  $\theta = \theta_I + \frac{\pi}{2}$  avec  $\theta_I + \frac{\pi}{2}$  une des orientations pour laquelle la mesure est calculée. Posons que cette orientation est égale à 0. Alors la valeur maximale de la mesure  $M_1$  devient :

$$M_{max} = \frac{1}{card(\mathcal{B})} \sum_{(x,y) \in \mathcal{B}} \frac{1}{K_{max}} \quad (3.18)$$

$$= \frac{1}{K_{max}} \quad (3.19)$$

La constante de normalisation  $K_{max}$  est quant à elle définie par :

$$K_{max} = \sum_{i=0}^{N-1} e^{-\frac{1}{2} dist^2(0, i \frac{\pi}{N}) \cdot \frac{\|\nabla I\|_{max}^2}{\sigma^2}} \quad (3.20)$$

$$= 1 + 2 \sum_{i=0}^{\frac{N}{2}} e^{-\frac{1}{2} (i \frac{\pi}{N})^2 \cdot (\frac{255}{2\sigma})^2} + e^{-\frac{1}{2} (\frac{\pi}{2} \cdot \frac{255}{2\sigma})^2} \quad (3.21)$$

Les figures 3.23 à 3.32 représentent les résultats pour la métrique  $M_1$  pour différents contenus. La valeur de l'écart-type caractérisant le bruit sur la mesure du gradient est fixée à 2, 5. Hormis pour des textures orientées (3.23 et 3.24) où la mesure est opportune pour la majorité des pixels de l'image ; pour les autres contenus, les pixels pour lesquels la norme du gradient est faible sont prépondérants, ce qui provoque un effet de tassement vers une loi uniforme de la métrique  $M_1$ . C'est pourquoi une second mesure, notée  $M_2$ , propose de pallier à ce problème.

Les pixels dont la norme du gradient est inférieure à un certain seuil  $\tau$ , ici pour les résultats présentés 10% de la norme maximale ( $\frac{255}{2}$ ), ne sont pas pris en compte dans la mesure. La mesure  $M_2$  est donc définie simplement par :

$$M_2(\theta) = \frac{1}{M_{max}} \frac{1}{\text{card}(\{\mathcal{B}/\|\nabla I(x, y)\| \geq \tau\})} \sum_{(x, y) \in \{\mathcal{B}/\|\nabla I(x, y)\| \geq \tau\}} P(x, y, \theta) \quad (3.22)$$

Pour les figures 3.23 et 3.24, l'orientation qui maximise la métrique  $M_2$  donne une bonne estimation de l'orientation perçue de la texture. Pour la première on retrouve l'orientation estimée par les précédentes méthodes. Par contre pour l'image représentant la texture du pantalon de **Barbara**, l'estimation de l'orientation perçue est meilleure qu'avec les méthodes précédentes. La mesure  $M_2$  pour la première texture est plus étalée que la seconde ce qui correspond bien aux caractéristiques de ces textures. De la même façon, la mesure  $M_2$  correspondant à la figure 3.25 est relativement étalée car l'orientation du contour de l'image est courbe. On peut donc imaginer caractériser la rectilinéarité d'un contour à partir de cette métrique.

Les figures 3.26, 3.27 et 3.29 présentent plusieurs orientations privilégiées. La mesure  $M_2$  permet de les identifier correctement. Les sous-bandes reconstruites des hautes fréquences spatiales selon ces orientations privilégiées sont présentées pour ces images. L'idée est de se rendre compte de l'intérêt qui peut exister à opérer un filtrage selon une direction ne correspondant pas à celle qui minimise l'énergie moyenne.

Les figures 3.30 et 3.31 concernent des images qui ne présentent pas d'orientation privilégiée. Ces dernières peuvent même être considérées comme uniformes. Le premier critère utilisé ici pour les classifier est simple : si aucun pixel de l'image n'a de norme du gradient supérieure au seuil  $\tau$ , alors l'image est classifiée comme uniforme. Toutefois ce critère pose un problème majeur dans le cas où par exemple un faible nombre de pixels ont une norme supérieure à  $\tau$ . Or dans ce cas, seuls ces pixels participent à la mesure. Par conséquent la mesure est très sensible aux structures non significatives qui peuvent émerger au dessus du seuil fixé. La solution adoptée ici est de considérer que pour qu'une mesure soit valide, un certain quorum de pixels participant à la mesure doit être atteint. Pour les résultats présentés, ce quorum a été choisi comme représentant la dimension moyenne de l'image analysée :

$$\text{quorum} = \sqrt{\text{hauteur} \times \text{largeur}} \quad (3.23)$$

Cette valeur a été choisie pour que typiquement un contour fort sur un fond uniforme qui traverse toute l'image soit pris en compte. Étant donnée la nature des filtres de gradient utilisés, c'est-à-dire centrés, tout contour entraînera au moins des réponses à ces filtres en deux pixels pour chaque pixel du contour. On peut donc imaginer que le quorum pourrait être plutôt fixé à deux fois la dimension moyenne de l'image. La valeur initiale de ce seuil est tout de même conservé car il permet une discrimination suffisante.

La figure 3.32 représente la mesure pour une zone très texturée ne présentant pas d'orientation privilégiée. Pour cette zone texturée, et à l'opposé des zones uniformes, le quorum est parfaitement atteint. Une mesure valide est donc calculée, pour laquelle un ou plusieurs maxima existent. Afin de déterminer si ces maxima correspondent à une orientation privilégiée, une mesure de distance moyenne entre  $M_2$  et une loi uniforme est calculée :

$$d = \sum_{\theta \in \Theta} \left| M_2(\theta) - \frac{K_{max}}{N} \right| \quad (3.24)$$

Comme on peut l'observer sur la figure 3.32, cette distance vaut environ 0,00252 alors que pour les contenus mono-orientés ou multi-orientés la distance est plus de dix fois supérieure. Un seuil sur cette distance permettrait donc de distinguer les contenus orientés des contenus texturés isotropes. Le choix de ce seuil n'est pas évident. La figure 3.28 illustre bien ce propos. La valeur de la distance  $d$  pour cette image est de 0,00999. Faut-il considérer cette image comme une zone présentant deux orientations privilégiées ( $0^\circ$  et  $90^\circ$ ) correspondant à une analyse générale d'une suite de caractères alignés horizontalement, ou alors faut-il considérer cette image plus en détail pour y intégrer les orientations des formes des lettres et donc y voir plutôt une région très texturée sans vraiment de direction privilégiée. Un tel choix de seuil devrait dépendre par conséquent de la différence de traitement appliqué à ces différentes classes.

Cette analyse est appliquée par bloc de taille fixe sur des images. Pour chaque bloc nous disposons alors de deux informations : la classe structurale du bloc considéré et l'orientation privilégiée (ou l'horizontale pour les classes uniformes ou textures isotropes) au sein de ce bloc. À partir de cette segmentation en bloc, deux arbres quaternaires, l'un contenant l'information de classe structurale, l'autre d'orientation, sont formés par fusion des blocs de même classe et même orientation. Ces deux arbres ont par conséquent la même structure. Seule la valeur de l'information portée par les feuilles diffère. Des exemples de ces arbres sont donnés en annexe E pour des blocs de taille minimale  $16 \times 16$ .

**3.2.2.2.1 Conclusion** La sous-section 3.2.2.1 a montré que des mesures basées sur le calcul de l'erreur de prédiction ne permettent pas de détecter les orientations privilégiées d'une région d'une image. Ces mesures, faisant intervenir une interpolation linéaire, introduisent différents biais qui rendent les mesures pour différentes orientations incomparables.

Une mesure éloignée de la transformée utilisée, présentée dans la sous-section 3.2.2.2, basée sur une simple mesure du gradient et prenant en compte l'incertitude de cette mesure permet quant à elle de déterminer si une région présente des orientations privilégiées ou non. Dans ce dernier cas, il peut s'agir d'une zone uniforme ou d'une zone texturée anisotrope. Dans le premier cas, les différents maxima de la mesure proposée correspondent aux orientations privilégiées de la région considérée.

Un seuillage des pixels basé sur la norme du gradient ainsi que sur le nombre de pixels de norme supérieure à un tel seuil permet de classifier les zones uniformes. La distance moyenne absolue entre la mesure et une loi uniforme permet de distinguer les zones texturées isotropes des zones présentant des orientations privilégiées. Le nombre de maxima de la mesure permet dans ce dernier cas de différencier les zones mono-orientées des zones multi-orientées.

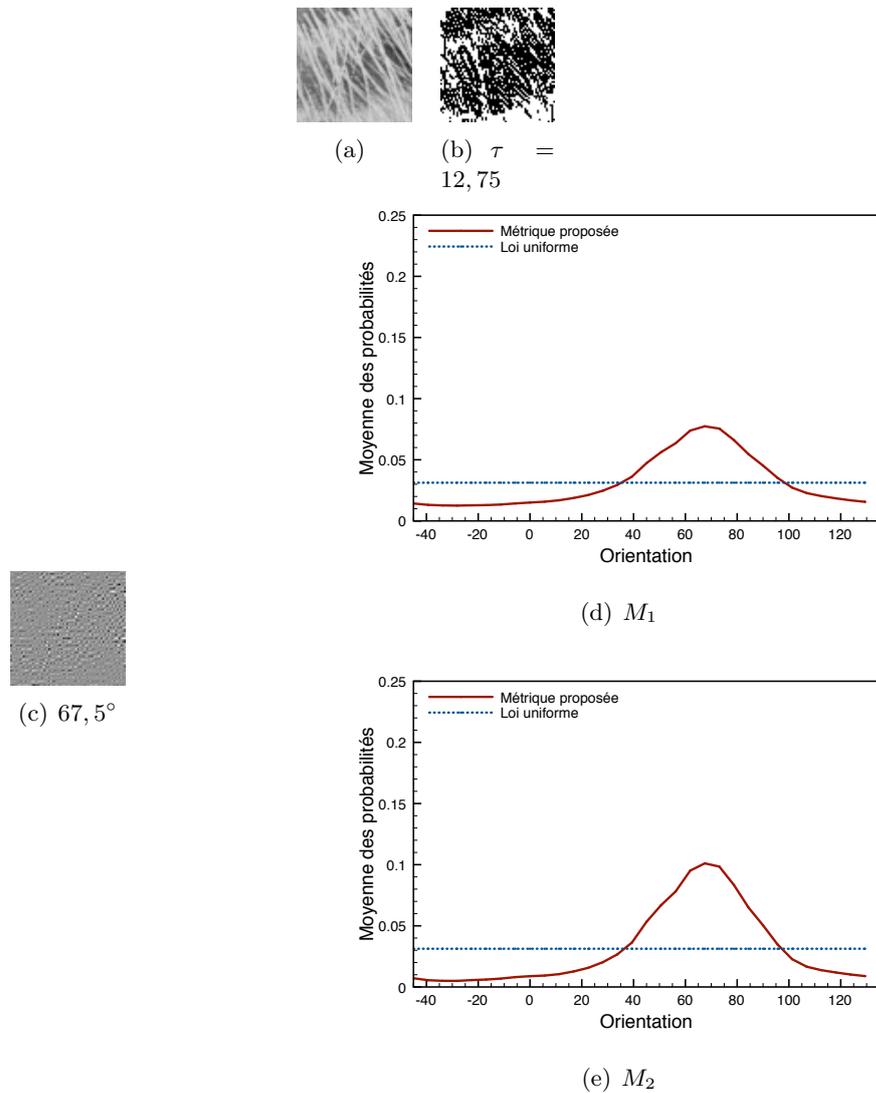


FIGURE 3.23 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `baboon_mouth_hair` de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . L'orientation qui maximise  $M_2$  est  $67,5^\circ$ , et le pourcentage de pixels participant à cette mesure est de 59,11% (le quorum est atteint : 2421/64). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0262.

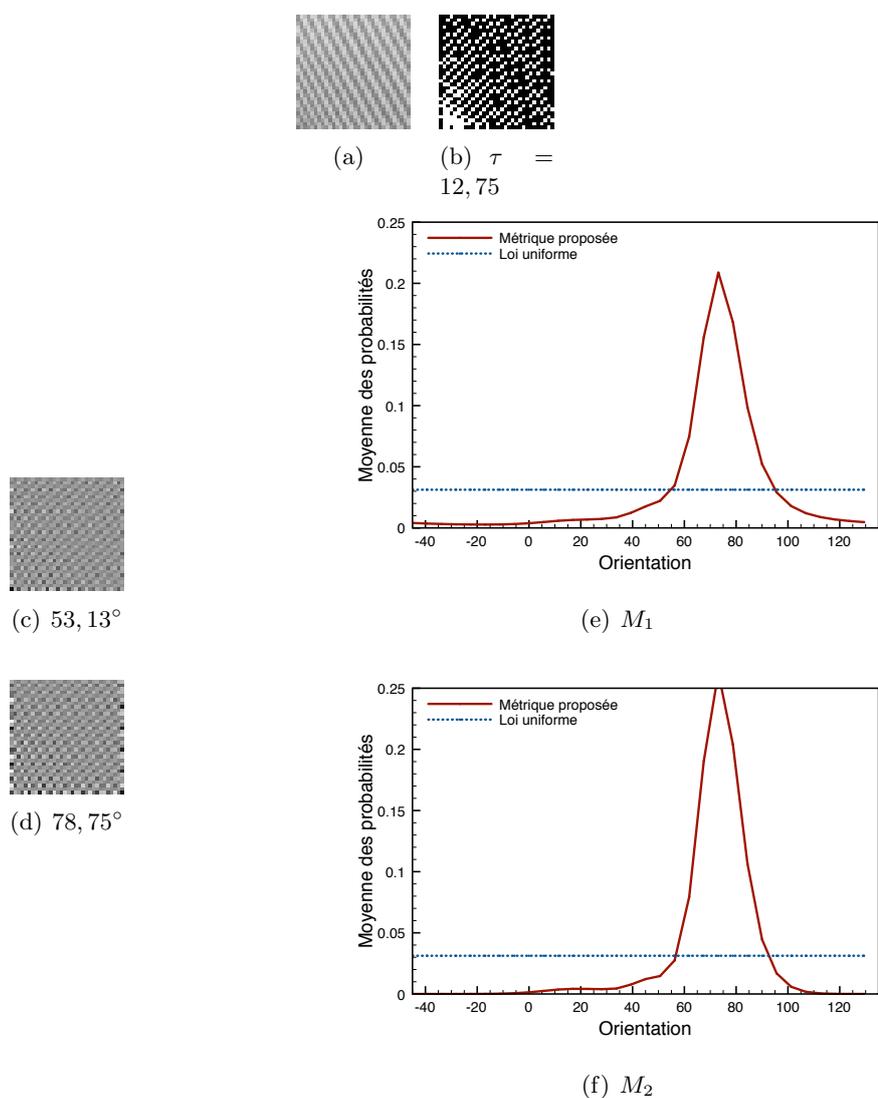


FIGURE 3.24 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `barbara_pants` de dimension  $32 \times 32$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . L'orientation qui maximise  $M_2$  est  $78,75^\circ$ , et le pourcentage de pixels participant à cette mesure est de 69,82% (le quorum est atteint : 715/32). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0438.

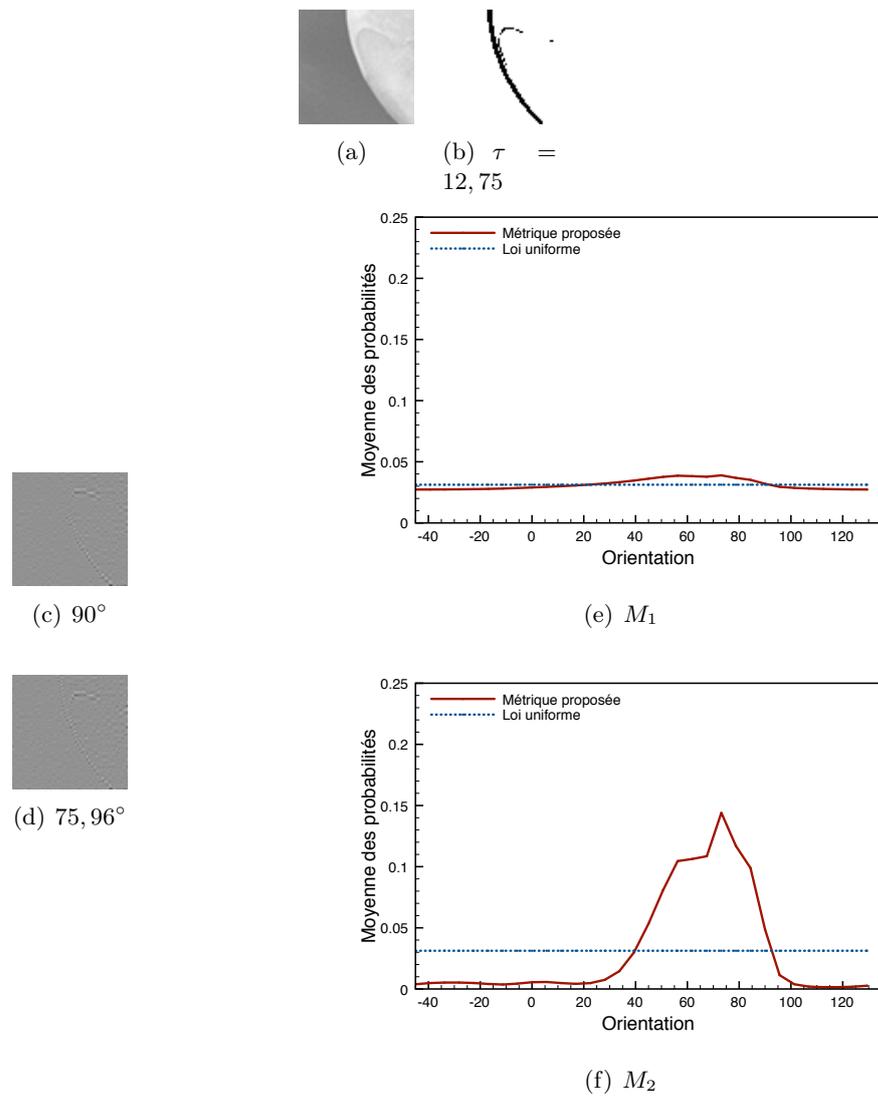


FIGURE 3.25 – Mesures de  $M_1$  et  $M_2$  pour l'imagette **parrots\_nose** de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . L'orientation qui maximise  $M_2$  est  $73, 125^\circ$ , et le pourcentage de pixels participant à cette mesure est de 5, 10% (le quorum est atteint : 209/64). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0363.

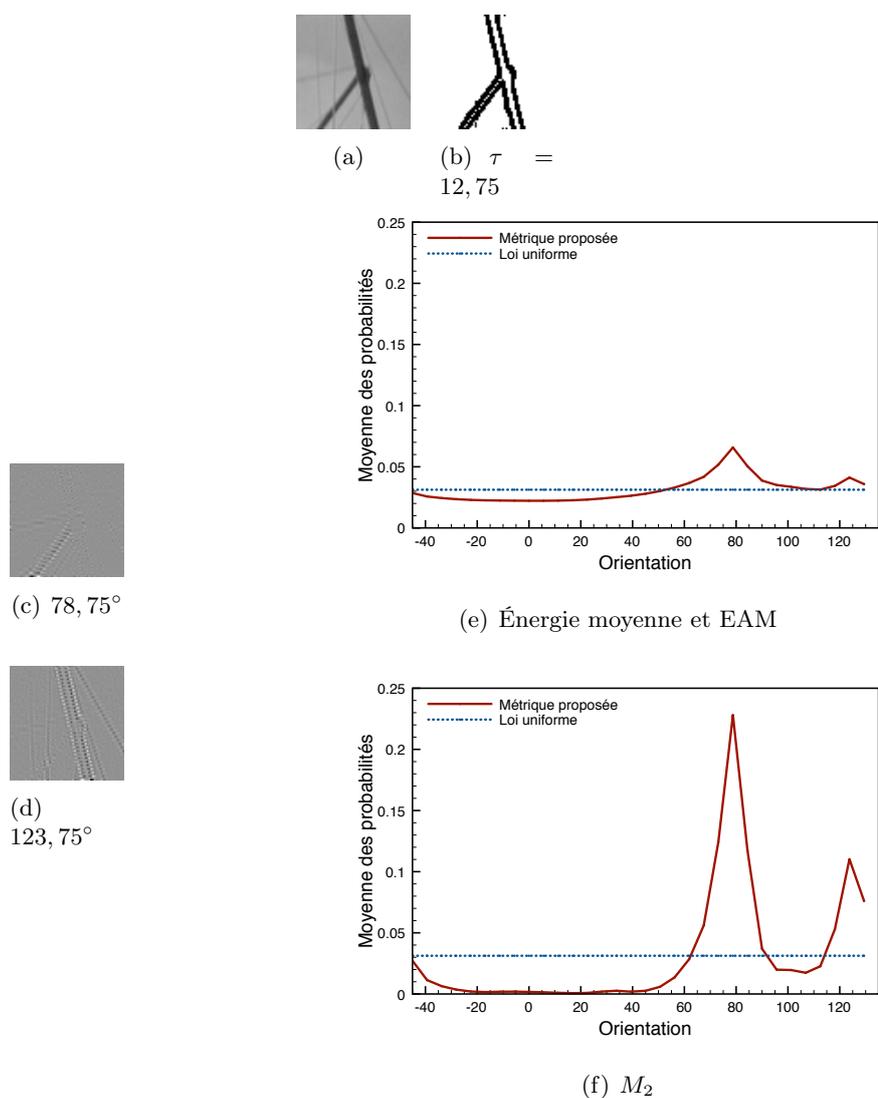


FIGURE 3.26 – Mesures de  $M_1$  et  $M_2$  pour l’image `boats_masts` de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Les orientations qui maximisent  $M_2$  sont  $78,75^\circ$  puis  $123,75^\circ$ , et le pourcentage de pixels participant à cette mesure est de 13,40% (le quorum est atteint :  $549/64$ ). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0345.

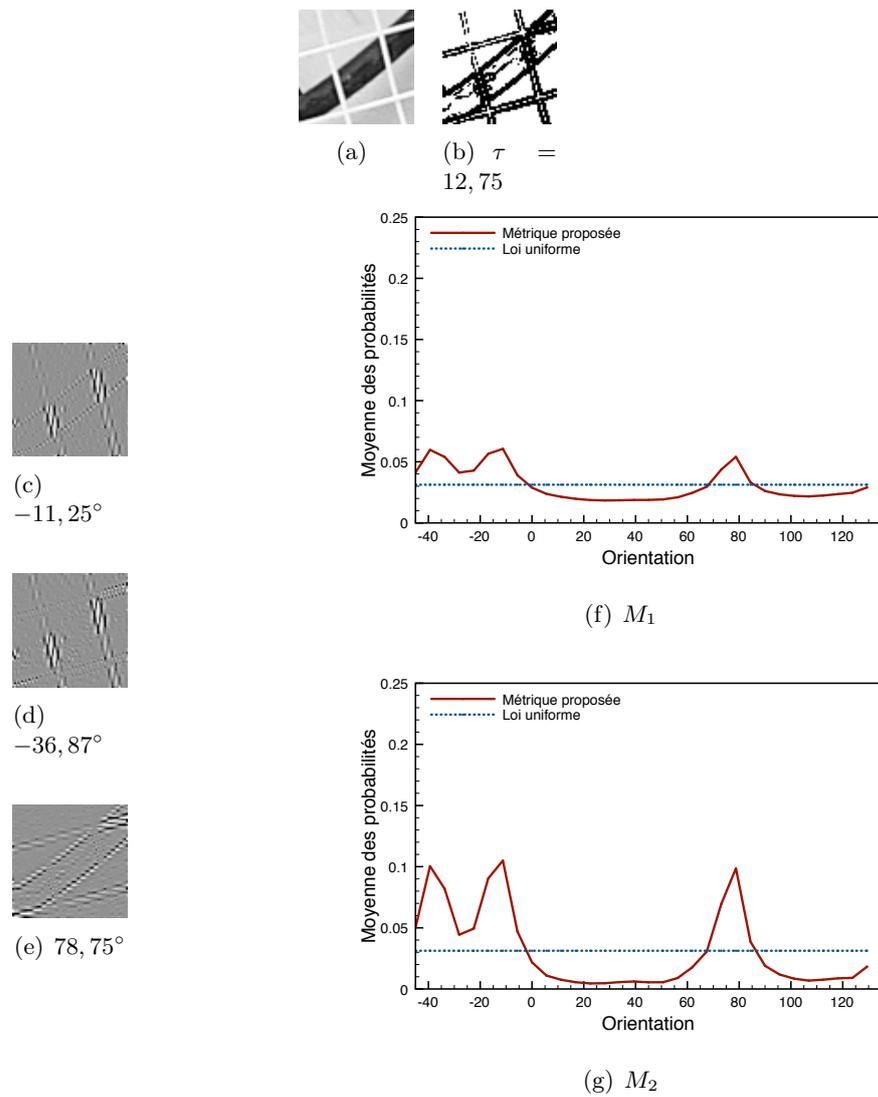


FIGURE 3.27 – Mesures de  $M_1$  et  $M_2$  pour l'imagette **fruits\_rackets** de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Les orientations qui maximisent  $M_2$  sont  $-11,25^\circ$ ,  $-39,375^\circ$ , puis  $78,75^\circ$ , et le pourcentage de pixels participant à cette mesure est de 34,86% (le quorum est atteint : 1428/64). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0270.

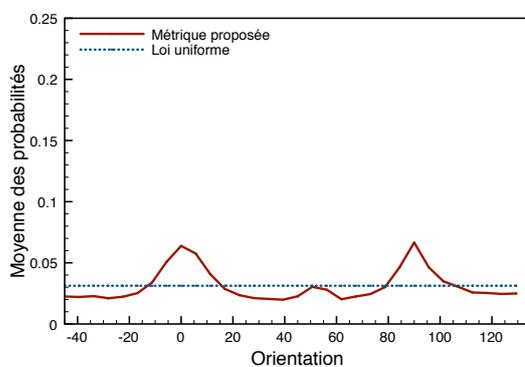
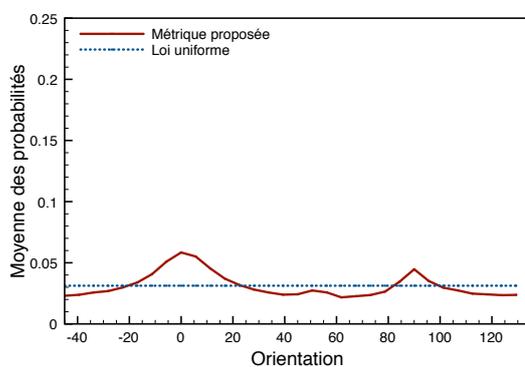
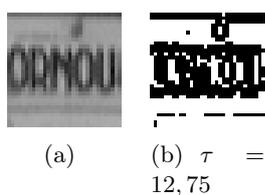


FIGURE 3.28 – Mesures de  $M_1$  et  $M_2$  pour l’image `boats_text` de dimension  $32 \times 32$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Les orientations qui maximisent  $M_2$  sont  $0^\circ$  puis  $90^\circ$ , et le pourcentage de pixels participant à cette mesure est de 48,93% (le quorum est atteint :  $501/32$ ). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,00999.

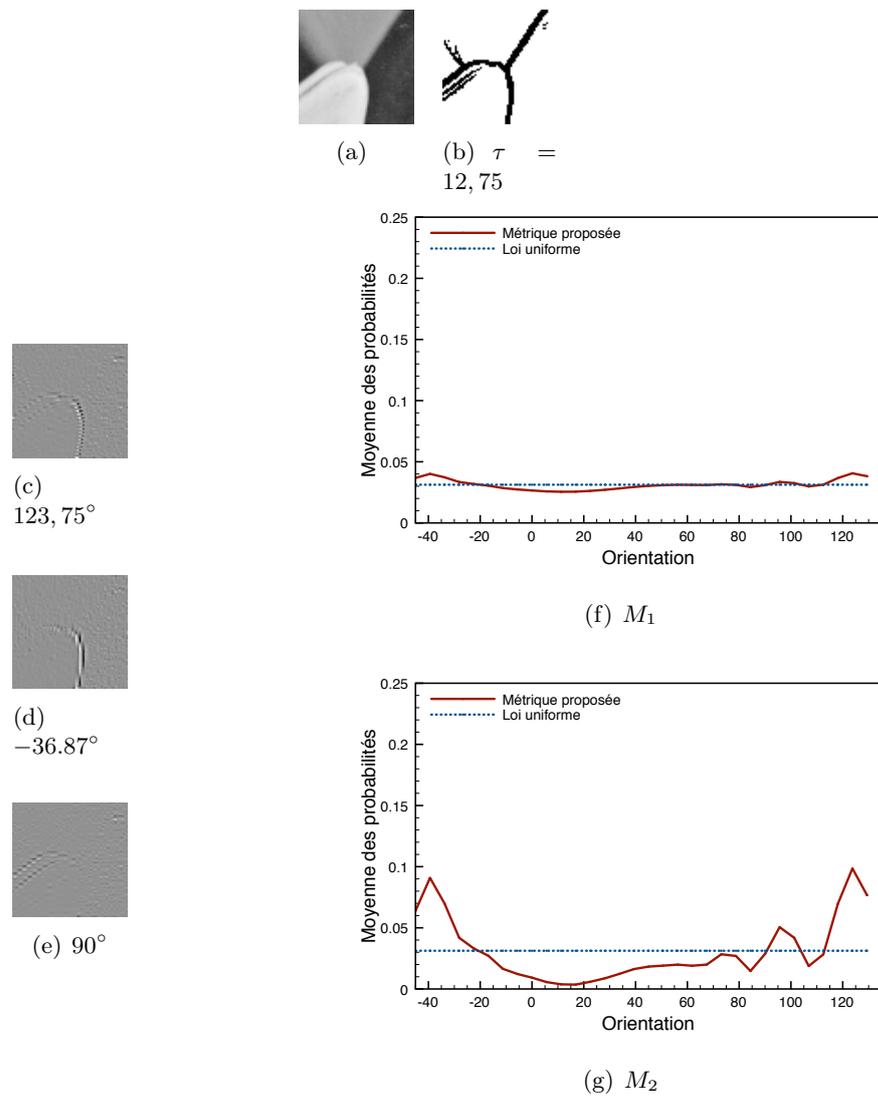


FIGURE 3.29 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `lena_hat` de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Les orientations qui maximisent  $M_2$  sont  $123,75^\circ, -39,375^\circ$  puis  $95,625^\circ$ , et le pourcentage de pixels participant à cette mesure est de 12,33% (le quorum est atteint : 505/64). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,0203.



(a)

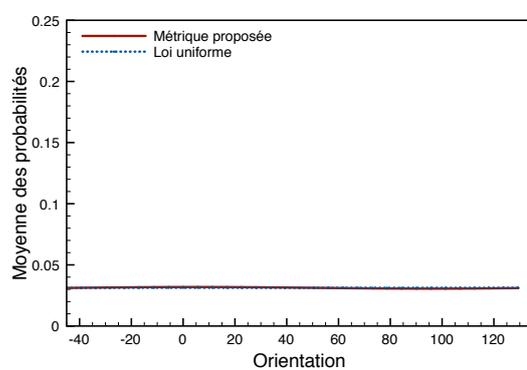
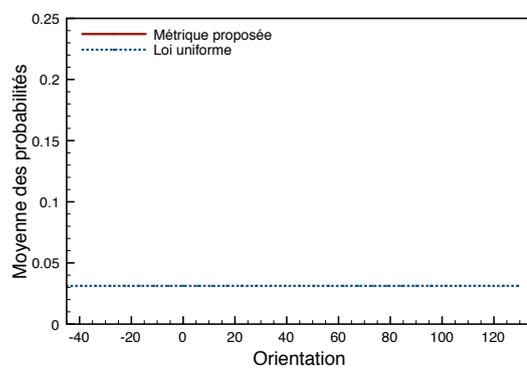
(b)  $\tau = 12,75$ (c)  $M_1$ (d)  $M_2$ 

FIGURE 3.30 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `boats_sky_and_clouds` de dimension  $128 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Le pourcentage de pixels participant à cette mesure est de 0% (le quorum n'est pas atteint : 0/91). L'image est donc classifiée comme uniforme.

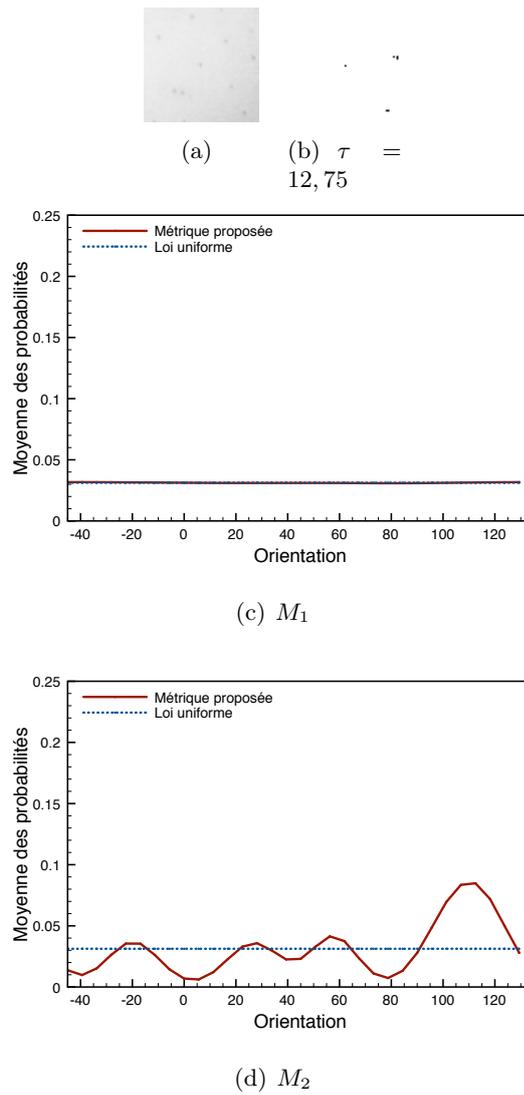
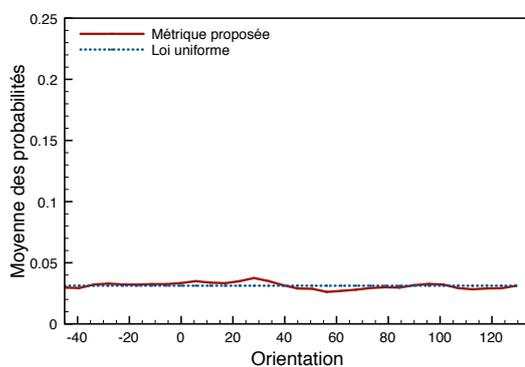


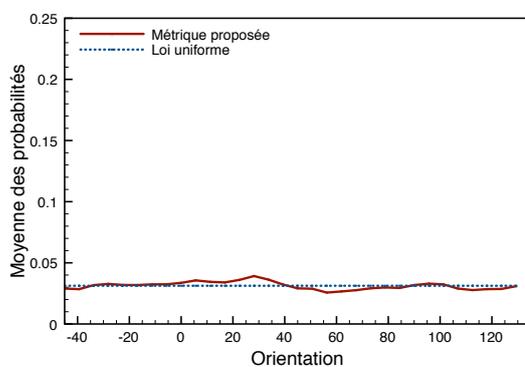
FIGURE 3.31 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `fruits_apple` de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Le pourcentage de pixels participant à cette mesure est de 0,15% (le quorum n'est pas atteint : 6/64). L'image est donc classifiée comme uniforme.



(a) (b)  $\tau = 12,75$



(c)  $M_1$



(d)  $M_2$

FIGURE 3.32 – Mesures de  $M_1$  et  $M_2$  pour l'imagette `train_leafs` de dimension  $64 \times 64$  selon 32 orientations uniformément réparties sur  $[-\frac{\pi}{4}, \frac{3\pi}{4}]$ . Les orientations qui maximisent  $M_2$  sont  $0^\circ$  puis  $90^\circ$ , et le pourcentage de pixels participant à cette mesure est de 48,93% (le quorum est atteint : 501/32). La distance absolue moyenne entre  $M_2$  et une loi uniforme est de 0,00252.

## Conclusion

L'estimation de l'orientation de filtrage a été envisagée sous deux angles différents. D'une part, depuis le chapitre précédent, selon le principe de la minimisation de l'énergie de la sous-bande des hautes fréquences spatiales selon la direction de filtrage reconstruite. Que cette orientation corresponde ou non à l'orientation perçue du contenu filtré, c'est celle qui minimise globalement l'énergie. D'autre part, l'estimation de la géométrie au travers de l'extraction de contour ou par la définition de nouvelles mesures plus ou moins proches de la transformée appliquée, donnent quant à elles une orientation de filtrage qui peut ne pas correspondre à l'orientation qui minimise l'énergie moyenne. Ce peut être contre productif pour des contenus mono-orientés, cela peut introduire un intérêt pour des contenus multi-orientés. Pour le cas des contenus mono-orientés, le caractère contre productif est très limité dans le cadre d'un  $\theta$ -redressement. Par contre le caractère instable de la technique de Mojette- $(p, q)$ -redressement rend l'usage de telles mesures risqué.

L'estimation des orientations de filtrage à partir des contours extraits, ainsi que l'estimation des orientations privilégiées à l'aide de la nouvelle mesure  $M_2$  ont de plus permis une classification des régions de l'image en quatre classes : les zones uniformes, les zones texturées isotropes, les zones mono-orientées et les zones multi-orientées.



# Conclusion

Parmi les transformées directionnelles existantes, nous avons choisi de nous intéresser aux transformations adaptatives implémentées par schéma de redressement. Pour cela nous avons étudié trois techniques de redressement pour implanter des transformées en ondelettes orientées :

le  $(p, q)$ -**Réveilles-redressement** (ou 8-connexe- $(p, q)$ -redressement) qui correspond à une transformée en ondelettes le long de lignes discrètes 8-connexes ;

le  $(p, q)$ -**Mojette-redressement** (ou  $(p, q)$ -connexe- $(p, q)$ -redressement) qui correspond à une transformée en ondelettes sur une grille discrète ré-échantillonnée ;

le  $\theta$ -**redressement** où la valeur des échantillons est interpolée pour les étapes de prédiction et de mise à jour. L'interpolation linéaire a été ici considérée.

La première méthode n'a jamais donnée satisfaction quant à sa capacité à minimiser l'énergie de la sous-bande reconstruite des hautes fréquences dans la direction de filtrage. Les deux autres techniques étudiées peuvent être mises en compétition en fonction du contenu à traiter. Le  $(p, q)$ -Mojette-redressement a démontré son efficacité pour divers contenus synthétiques représentant des structures rectilignes et pour certains contours naturels particulièrement rectilignes et particulièrement contrastés. Le  $\theta$ -redressement, permet dans certains cas un gain dans la minimisation de l'énergie de la sous-bande des hautes fréquences dans la direction de filtrage reconstruite par rapport aux orientations de filtrage ne nécessitant pas d'interpolation.

Le choix des orientations de filtrage n'est pas trivial avec la méthode utilisant l'interpolation car l'énergie des filtres équivalents aux diverses orientations de filtrage possibles varie en fonction de l'orientation. Ainsi les dynamiques des signaux des sous-bandes produites pour différentes orientations de filtrage ne sont pas directement comparables. Comparer ces sous-bandes après reconstruction est alors possible mais d'autant plus coûteux. De plus nous ne souhaitons pas nécessairement déterminer l'orientation qui minimise globalement l'énergie mais l'orientation qui permet un traitement adapté à l'orientation principale de l'image (ou d'une région de l'image), même si cela s'écarte d'une approche purement codage. Deux méthodes ont été alors proposées. Une première méthode basée sur une extraction des contours significatifs de l'image. Après chaînage et estimation de l'orientation des éléments de ces contours, les régions de l'image découpée en blocs, sont classifiées, et pour chaque bloc une orientation de filtrage correspondant, si elle existe, à une orientation dominante de la région est assignée. La deuxième méthode est basée sur la mesure de l'orientation du gradient, en modélisant l'incertitude

de cette mesure en tout point par une loi gaussienne. Ces mesures pour chaque bloc sont alors cumulées et le nombre de maxima ainsi que l'écart avec une loi uniforme sont utilisés pour classer comme pour la première méthode ces différents blocs.

Ces deux méthodes fournissent à la fois une classification des régions de l'image découpée en blocs et un choix d'orientation de filtrage pour ceux-ci. Ces deux informations servent de base à la définition d'une transformée en ondelettes orientées adaptée au contenu de l'image.

La plupart des implémentations des transformées en ondelettes utilisant les techniques de redressement précédemment évoquées [CG07, DWW<sup>+</sup>07, WZS06] mettent en oeuvre la même décomposition polyphase pour l'ensemble de l'image. Un redressement basé colonne est donc appliqué pour tous les blocs de l'image, puis un redressement basé ligne, ou inversement. Cela implique que le filtrage selon la première comme la seconde dimension de l'image ne peut faire intervenir que des orientations dont l'implantation correspond à la décomposition polyphase considérée.

Pour la transformée proposée ici, la décomposition polyphase peut varier d'un bloc à l'autre. Le choix de la première orientation de filtrage n'est donc pas dépendante de cette décomposition. Au contraire, c'est la décomposition polyphase qui dépend du choix de l'orientation de filtrage pour chaque région de l'image. La transformée que nous proposons est donc définie par sa première orientation de filtrage.

Le choix de cette orientation est définie par l'une de nos deux méthodes de classification et d'estimation des orientations privilégiées. Ces deux méthodes introduisent quatre classes qui caractérisent l'activité des régions de l'image : uniforme, mono-orientée, multi-orientée et texturée de façon isotrope. Ces classes ne désignent pas exactement le même contenu selon la méthode employée, et donc ne conduisent pas à la même segmentation (1 : méthode basée extraction de contours, ou 2 : méthode basée mesure du gradient) :

#### **zone uniforme**

1. zone ne contenant aucun élément de contour extrait. Par conséquent soit aucun contour de longueur supérieure à un certain seuil n'a été extrait, soit aucune norme du gradient mesurée en un point de la zone ne dépassait un certain seuil ;
2. zone où moins d'un certain quorum d'échantillons n'ont leur norme du gradient supérieure à un certain seuil.

#### **zone mono-orientée**

1. zone où le pourcentage d'éléments de contours dont l'orientation appartient à une même classe d'orientation est supérieur à un seuil haut ;
2. zone où la présence d'une orientation privilégiée unique (un seul maximum) ne semble pas être due au bruit (écart suffisant avec une loi uniforme).

#### **zone multi-orientée**

1. zone où le pourcentage d'éléments de contours dont l'orientation appartient à une même classe d'orientation est supérieur à un seuil bas ;

2. zone où la présence d'au moins deux orientations privilégiées ne semble pas être due au bruit.

#### zone texturée de façon isotrope

1. zone où le pourcentage d'éléments de contours dont l'orientation appartient à une même classe n'est pas supérieur à un seuil bas ;
2. zone où aucune orientation privilégiée n'est détectée (écart avec une loi uniforme trop faible).

Pour chacune de ces classes, l'orientation de filtrage prend une signification particulière, et les contenus des sous-bandes produites après filtrage peuvent être interprétés différemment.

Pour les zones uniformes ou les textures isotropes, aucune orientation de filtrage n'est préconisée par l'algorithme d'estimation des orientations privilégiées. Le choix de l'orientation de filtrage peut dès lors être arbitraire, par exemple  $0^\circ$  ou  $90^\circ$  pour correspondre à une transformée en ondelettes plus conventionnelle, ou dépendre d'orientations de filtrage de régions voisines.

Pour les zones mono- et multi-orientées, l'orientation de filtrage est choisie comme l'orientation dominante de la région considérée. L'objectif est alors de tirer parti de la régularité le long de cette direction pour minimiser localement l'énergie de la sous-bande des hautes fréquences produite. Pour cette raison, cette direction de filtrage est nommée direction de régularité. Dans l'idéal, pour une zone mono-orientée, une telle sous-bande ne fera pas apparaître la structure du contour ou de la texture orientés selon cette direction, mais uniquement le bruit le long de ces structures. De même pour une zone multi-orientée concernant les structures relatives à la direction de régularité, par contre la sous-bande des hautes fréquences après le premier filtrage contiendra les structures qui correspondent aux autres orientations privilégiées de la région.

Nous proposons finalement une représentation basée sur une décomposition en ondelettes selon deux directions qui s'adaptent localement de telle sorte que la première orientation de filtrage corresponde à l'orientation dominante localement. Des filtres différents peuvent être appliqués selon chacune des deux orientations de filtrage, permettant ainsi un traitement anisotropique. Comme la transformée en ondelettes 2D classique, notre représentation produit quatre sous-bandes pour un niveau de décomposition donné :

**BF** les basses fréquences selon les deux directions de filtrage ;

**HFReg** les hautes fréquences selon la direction de régularité et basses fréquences selon la seconde direction de filtrage ;

**HFOrtho** les hautes fréquences selon la seconde direction de filtrage et basses fréquences selon la direction de régularité ;

**HFRes** les hautes fréquences selon les deux directions de filtrage.

Idéalement, les structures considérées dominantes seront uniquement présentes dans la sous-bande *HFOrtho* alors que la sous-bande *HFReg* ne contiendra que le bruit le long de ces structures.

Des exemples de classification d'images ainsi que les orientations de filtrage estimées par la méthode basée sur la mesure du gradient sont présentées à l'annexe E.



Deuxième partie

Quantification et codage



# Introduction

Dans la première partie, nous avons construit une représentation structurelle d'images par transformée locale en ondelettes orientées. Cette représentation consiste pour une image donnée d'une part en deux arbres quaternaires représentant les orientations de filtrage correspondant aux orientations de régularité estimée dans l'image, ainsi que ses classes structurelles, et d'autre part les sous-bandes produites après transformée en ondelettes orientées.

En fonction de chaque classe structurelle, l'interprétation du contenu de chaque sous-bande d'un même niveau est différente. Alors que l'on ne peut traiter les éléments des sous-bandes classés comme uniformes ou texturés isotropes que de façon isotrope justement, il est possible de traiter les éléments des sous-bandes classés comme mono-orientés ou multi-orientés de façon anisotrope.

À cette fin nous proposons d'adapter les lois de quantification des différentes sous-bandes en fonction de la classe structurelle des différentes régions de l'image, pour finalement produire un train binaire après codage comparable avec le coût de codage de la transformée en ondelettes.

Toutefois pour que la comparaison soit complète, il faut prendre en compte le coût de codage de la géométrie de l'image, c'est-à-dire les deux arbres quaternaires de notre représentation, ou alors les contours extraits de l'image qui conduisent à ces deux arbres quaternaire.

Le coût de codage des informations d'orientation et de classification des régions de l'image est l'objet du quatrième chapitre, et l'adaptation des lois de quantification en fonction de la classification est étudiée dans le dernier chapitre.



## Chapitre 4

# Codage des informations d'orientation et de classification des régions

### Introduction

Au chapitre 3, deux méthodes ont été proposées pour établir d'une part les orientations de filtrage utilisées pour l'application d'une transformée en ondelettes orientées implantée par une technique de redressement, ainsi qu'une carte indiquant la classification de l'activité des différentes régions de l'image selon quatre classes : zones uniformes, mono-orientées, multi-orientées ou texturées de façon isotrope. Les deux méthodes aboutissent à la description de telles informations au travers de deux arbres quaternaires (l'un représentant l'orientation de filtrage, l'autre la classe des régions de l'image). Seule la première méthode construit les arbres à partir d'un résultat intermédiaire : les contours extraits de l'image. Pour que l'image soit reconstruite après codage/décodage, les informations supplémentaires doivent elles aussi être transmises au décodeur. Cette partie de la thèse détaille donc cet aspect.

Dans un premier temps, le codage de chaînes d'éléments de contours est étudié et une amélioration dans un contexte multi-résolution est proposée. Le codage des arbres quaternaires est ensuite envisagé, plusieurs implantations sont proposées dans notre cas particulier de codage d'arbres quaternaires partageant une structure commune et des informations complémentaires. Finalement les coûts de codage de ces deux représentations de l'information nécessaire à la reconstruction d'images sont comparés entre eux.

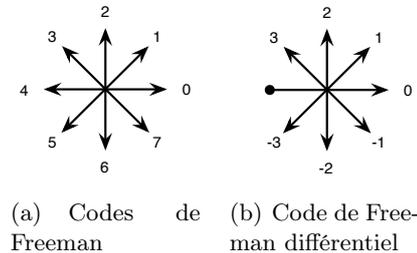


FIGURE 4.1 – Codage de chaînes.

## 4.1 Codage arithmétique adaptatif des contours d'une image en fonction de la résolution inférieure

### Introduction

Le codage de chaînes d'éléments est utilisé pour des applications telles que le codage de contours ou le codage de cartes de segmentation. Freeman a proposé une représentation à huit directions, en utilisant des codes de longueur fixe, pour le codage de chaînes d'éléments 8-connexes. Chaque chaîne est composée d'une tête, le premier élément d'un bout de la chaîne, représentée par ses coordonnées, et d'un chemin représenté par une succession de directions, que nous nommerons aussi « liens », chacune codée de 0 à 7, indiquant la position de l'élément suivant (voir figure 4.1-a). Avec ce système de codage, la densité de probabilité des symboles de direction peut être considérée comme uniforme. Par conséquent le coût de codage pour un chemin quelconque est de l'ordre de 3 bits par lien sans compter les symboles de terminaison.

Pour améliorer ce débit, Freeman a proposé un codage des changements de direction, appelé aussi codage différentiel de Freeman. Chaque chaîne est toujours composée d'une tête et d'un chemin. Mais ce chemin est représenté par une première direction, puis par une succession de changements de direction, chacun codé de  $-3$  à  $3$  (voir figure 4.1-b). Considérant la régularité des chaînes, la densité de probabilité de ces symboles n'est pas uniforme, et donc un codage à longueur variable est adapté, conduisant à un débit typique d'environ 2 bits par lien.

Lu et Dunham [LD91] ont amélioré le codage de Huffman de chaînes d'éléments de contours, en modélisant la structure de la source des symboles par un modèle de Markov de premier et second ordre, initialisé par des statistiques empiriques calculées à partir d'un ensemble de cartes de contours. À partir de cette idée, Chan and Siu [CS95] ont proposé une implantation utilisant un codeur arithmétique adaptatif, où l'initialisation est là faite via les statistiques empiriques calculées sur le message à transmettre. Ces statistiques doivent donc être transmises au décodeur. Un codeur arithmétique adaptatif est aussi retenu par Pateux et Labit [PL97], mais le codeur est alors initialisé par une loi *a priori* telle que les statistiques utilisées par Lu et Dunham. Dans ce cas par contre, le codeur arithmétique adaptatif fait évoluer ces statistiques au fur et à mesure du codage.

Comme décrit dans la section 3.2.1 du chapitre 3, les contours peuvent être direc-

tement extraits de l'image et guider le filtrage orienté appliqué sur cette dernière. Pour une transformée multi-résolution, nous avons par exemple choisi d'extraire les contours de l'image pour chacun de ses niveaux de résolution (sauf le dernier niveau sur lequel aucune transformée n'est appliquée). Ces contours, représentés par des chaînes de coordonnées de pixels, doivent être encodés pour chacun de ces niveaux de résolution.

Ces contours sont donc extraits à chaque résolution de manière indépendante. Certains contours peuvent être présents à pleine résolution, mais pas à plus faible résolution car des détails ne sont plus détectés. À l'inverse, des structures à basse résolution peuvent apparaître qui n'étaient pas significatives à pleine résolution. Par exemple un contour présentant une transition longue dont la norme du gradient est faible ne sera pas détecté à pleine résolution, mais pourra l'être à basse résolution.

Sans aucun *a priori* sur les contours extraits aux différents niveaux de résolution, nous proposons d'améliorer les méthodes de l'état de l'art pour coder les contours extraits d'une image à la résolution  $r_i$ , en utilisant une information disponible au décodeur : l'image décodée à la résolution inférieure  $r_i - 1$ .

L'amélioration apportée par cette information est d'abord étudiée du point de vue du coût entropique. Notre méthode est finalement implantée avec un codeur arithmétique et comparée avec les méthodes existantes.

Dans l'étude présentée, pour chaque image testée, les versions à différentes résolutions de l'image sont calculées en utilisant trois transformées en ondelettes successives utilisant les filtres d'ondelettes 9/7. Pour chacun des trois niveaux de résolution obtenus, les contours sont extraits avec la méthode décrite dans la section 3.2.1 du chapitre 3. Les contours extraits des dix images de tests sur lesquelles l'étude a porté sont présentés à l'annexe D.

#### 4.1.1 Codage arithmétique et modélisation markovienne de codes de Freeman différentiels

Un codeur s'appuyant sur les codes de Freeman ou les codes différentiels de Freeman, peut utiliser les probabilités d'apparition de symboles modélisées par des chaînes de Markov d'ordre  $N$ . Le symbole à coder représentant la direction suivante, ou le changement de direction suivant, à coder est représenté par la variable aléatoire  $A$ . La probabilité de réalisation  $a_i$  de la variable aléatoire  $A$  conditionnellement à un certain passé, noté  $B = \{B_j/0 < j \leq N\}$  et représentant les  $N$  dernières directions ou derniers changements de direction, définit les probabilités d'apparition des symboles à coder. Cette probabilité conditionnelle est notée  $Pr(A = a_i|B)$ . Par exemple, compte tenu de la régularité des contours extraits dans les images naturelles, un changement de direction important est rarement suivi d'autres changements similaires.

Lu et Dunham [LD91] ont obtenu pour le codage de chaînes d'éléments de contours, de meilleurs résultats pour des chaînes de Markov d'ordre 1 et 2 avec un codeur arithmétique qu'avec un codeur de Huffman. Pour cette raison le codage arithmétique est aussi considéré dans notre étude..

Dans l'implantation d'un codeur arithmétique adaptatif [BCK02], les probabilités conditionnelles relatives au modèle de Markov considéré, le contexte, sont représentées

par des compteurs. Ceux-ci sont mis à jour après chaque codage d'un symbole. Le compteur relatif à ce symbole pour un contexte donné est, par exemple, incrémenté afin que la probabilité d'apparition à l'avenir d'un tel symbole, dans le même contexte, soit plus importante.

Le nombre de ces compteurs, noté  $\#B$ , est défini par la combinatoire de  $B$ , le nombre de contextes, et donc dépend de l'ordre de Markov considéré.

Ces compteurs peuvent être initialisés par une loi uniforme ou par une loi *a priori* de telle sorte que le rapport, entre la valeur de chaque compteur et la somme des valeurs de tous les compteurs, représente au mieux la loi considérée. En pratique pour le cas uniforme, tous les compteurs sont initialisés avec une valeur identique  $K$ , en général égale à 1. Plus  $K$  est petit, plus les compteurs, et donc les probabilités conditionnelles du modèle de Markov, s'adapteront rapidement à la statistique du message. L'initialisation à partir d'une loi *a priori*, basée par exemple sur des statistiques empiriques, a pour objectif de réduire le temps d'apprentissage du codeur. Le choix de la somme des valeurs des compteurs,  $K \times \#B$  dans le cas uniforme, répond au compromis suivant : plus cette somme est faible, plus l'adaptation sera rapide mais plus la modélisation des statistiques empiriques seront quantifiées fortement à travers leur représentation par les compteurs.

L'efficacité d'un codeur arithmétique adaptatif est liée justement à sa capacité à s'adapter à la statistique du message à coder. Cette capacité est directement liée au nombre de compteurs nécessaires et à la taille du message à coder. Pour un codeur arithmétique adaptatif initialisé à partir d'une loi uniforme, plus l'ordre de Markov considéré est élevé, plus le message doit être long pour que l'apprentissage ait le temps de s'effectuer et puisse être efficace sur le codage du message. Utiliser une loi *a priori* peut permettre d'accélérer la convergence de l'estimation de la statistique du message si celle-ci est effectivement bien approchée par la loi *a priori*.

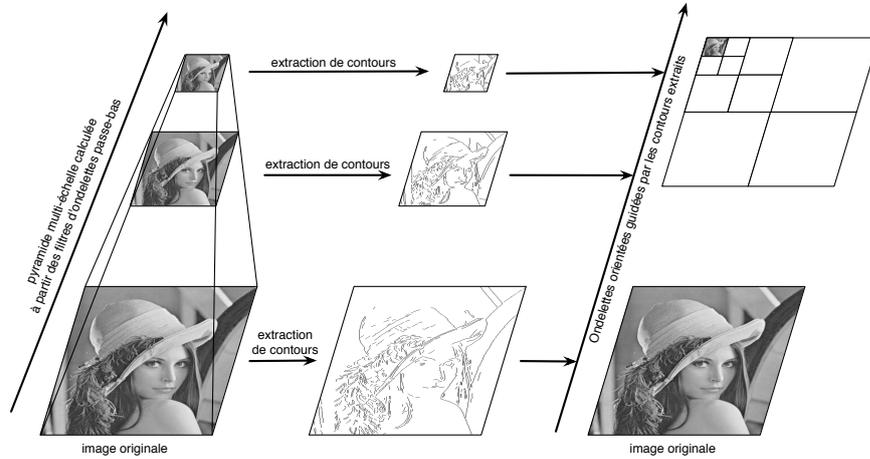
#### 4.1.2 Entropies des probabilités conditionnelles pour un codage de contours d'images à différentes résolutions

Dans un contexte de codage d'images multi-résolution et des contours extraits aux différents niveaux de résolution, l'objectif est d'utiliser l'information déjà disponible au décodeur pour améliorer l'efficacité du codeur, ce qui se traduit par l'amélioration de la prédiction des symboles à coder.

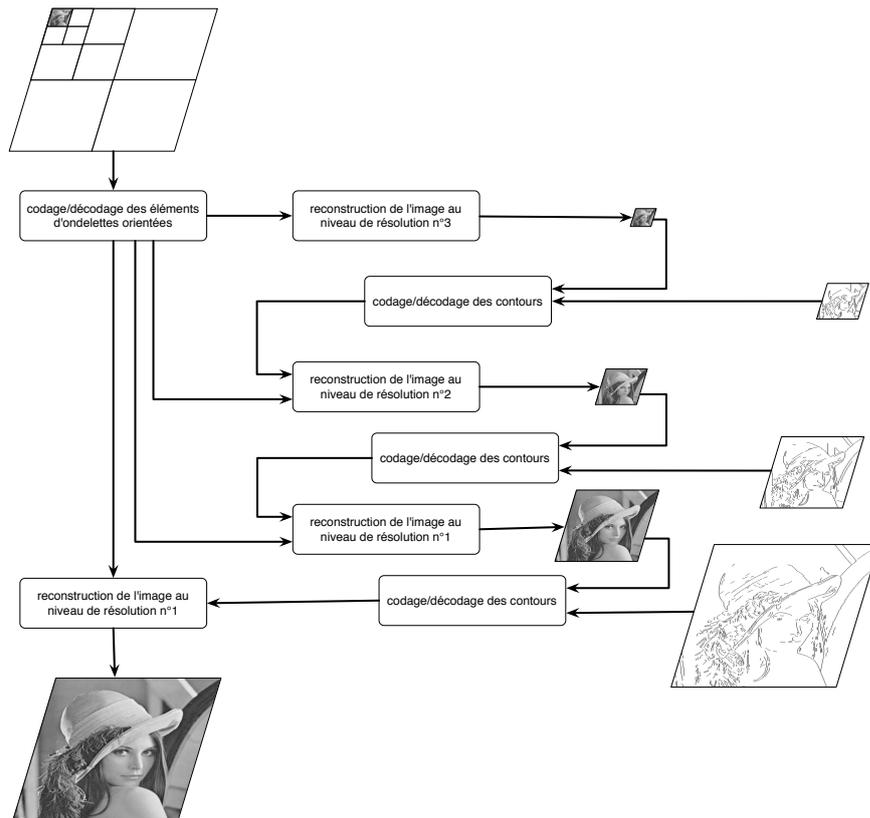
Notre première idée était d'utiliser les contours décodés à une résolution plus faible pour prédire la position des contours à coder à la résolution courante. Toutefois comme certains détails n'apparaissent pas aux résolutions inférieures, la prédiction est parfois impossible. De plus même si l'information existe, elle est parfois trop « quantifiée » pour être utilisable.

Pour ces raisons nous avons préféré utiliser comme information l'image décodée à la résolution inférieure. La figure 4.2 représente les différentes étapes de codage et de décodage des contours extraits à différents niveaux de résolution en utilisant l'information de l'image décodée à la résolution inférieure à la résolution d'extraction des contours.

Plus précisément l'orientation du gradient au sein de cette dernière image, qui offre une bonne indication de l'orientation des contours à la résolution courante, est comparée



(a) Extraction des contours à chaque niveau de résolution obtenu par une transformée en ondelettes 2D, suivi d'une transformée en ondelettes orientées.



(b) Codage et décodage des contours et des éléments d'ondelettes orientées.

FIGURE 4.2 – Schéma de codage et décodage des contours extraits utilisés dans un contexte de transformée en ondelettes orientées.

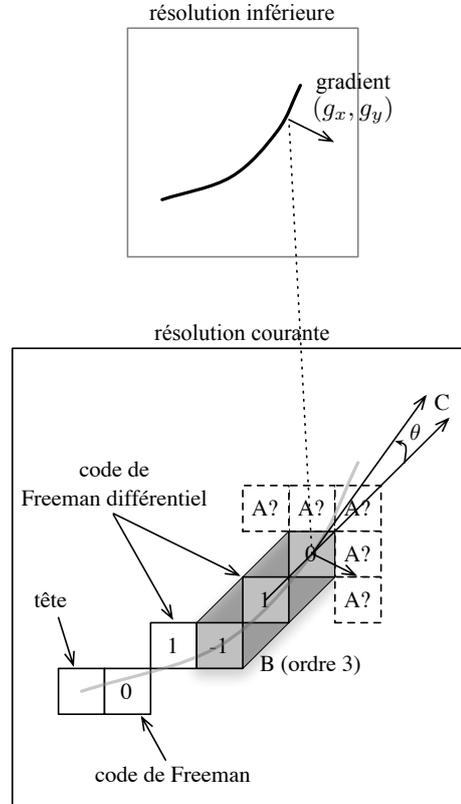


FIGURE 4.3 – *A* : le prochain code différentiel de Freeman à coder/décoder, *B* : les  $N$  derniers codes différentiels de Freeman codés/décodés, *C* : l'écart entre la direction orthogonale au gradient à la résolution inférieure et la direction donnée par les deux derniers pixels codés/décodés.

à l'orientation donnée par les deux derniers éléments de contours décodés. Soit  $(P_{ix}, P_{iy})$  les coordonnées du dernier élément de contour décodé, et  $\theta_{P_i} \in [-\pi, \pi]$  la direction orthogonale à l'orientation du gradient au point  $(\lfloor \frac{P_{ix}}{2} \rfloor, \lfloor \frac{P_{iy}}{2} \rfloor)$  de l'image décodée à la résolution inférieure. L'information est alors disponible en tous points de l'image. Notons  $C$  l'écart entre  $\theta_{P_i}$  et l'orientation de la droite  $(P_{i-1}, P_i)$  (voir figure 4.3).

L'entropie moyenne, sur trois niveaux de résolution, des probabilités conditionnelles du lien suivant en considérant le passé à la résolution courante ( $Pr(A|B)$ ), l'information à la résolution inférieure ( $Pr(A|C)$ ) ou les deux ( $Pr(A|B \cap C)$ ), pour les dix images de test considérées est représentée sur la figure 4.4 pour des ordres de Markov de zéro à trois.

Comme déjà observé dans la littérature, plus l'ordre de Markov est élevé, plus l'entropie est faible. Concernant l'utilisation de l'information représentée par  $C$ , on constate bien que l'entropie de  $A$  conditionnellement à  $B$  et  $C$  est en moyenne plus faible que celle de  $A$  conditionnellement à  $C$ , qui elle-même est en moyenne plus faible que l'en-

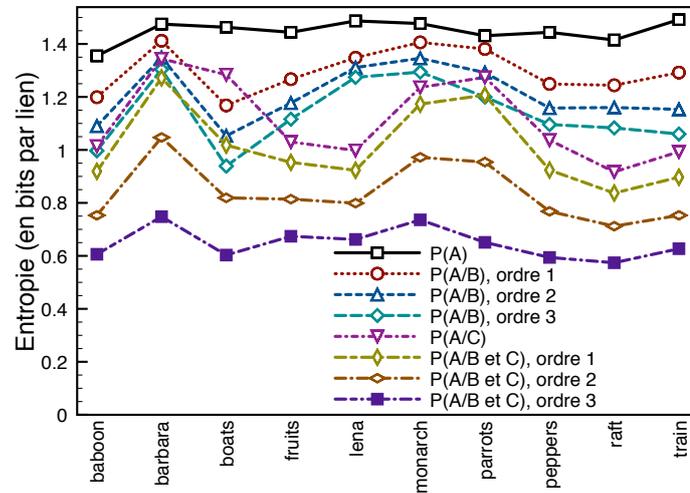


FIGURE 4.4 – Entropie moyenne, sur trois niveaux de résolution, des probabilités conditionnelles du lien suivant en considérant le passé à la résolution courante ( $P(A|B)$ ), l'information à la résolution inférieure ( $P(A|C)$ ) ou les deux ( $P(A|B \cap C)$ ) pour les dix images de test considérées.

trope de  $A$  conditionnellement à  $B$ . Ces résultats encouragent l'implantation de ces probabilités conditionnelles dans un codeur arithmétique adaptatif.

#### 4.1.3 Codage arithmétique adaptatif de contours d'images à différentes résolutions

Trois raisons font que le codage arithmétique adaptatif des contours d'images à différentes résolutions ne pourra pas atteindre le minimum entropique calculé dans la section précédente :

- l'entropie est calculée connaissant toutes les statistiques des contours, alors que le codeur arithmétique utilise uniquement les statistiques des éléments de contours déjà codés ;
- plus la dimension du contexte est importante, plus longue sera la phase d'apprentissage ;
- le calcul de l'entropie ne prend pas en compte le codage des symboles de terminaison de chaînes, ainsi que le coût de codage nécessaire aux éléments de têtes des contours.

En l'occurrence, pour le codage des têtes des chaînes de contours, les coordonnées de ces têtes, triées pour être balayées ligne par ligne et de gauche de droite, sont codées de façon différentielle avec un codeur arithmétique adaptatif d'ordre 1. Les premiers successeurs de ces têtes (éléments d'indice 1), représentés par un code de Freeman (non différentiel donc) sont aussi codés avec un codeur arithmétique adaptatif d'ordre 1. Les résultats présentés dans cette section ne concerneront toutefois que le codage des chemins (éléments d'indice supérieur ou égal à 2) car le codage des premiers éléments

est le même pour toutes les techniques comparées entre elles ici. Nous devons préciser d'autres aspects techniques :

Plus la dimension du contexte est importante, plus l'adaptation du codeur est lente. Par conséquent l'efficacité du codeur s'en ressent. Il convient donc de limiter au maximum le nombre de valeurs possibles que peuvent prendre les réalisations de  $B$  et  $C$ .

Par construction, les contours que nous extrayons sont tels que deux éléments de contours 8-connexes ne peuvent pas avoir de voisin 8-connexe en commun. Les changements de direction possibles de la chaîne sont alors restreints à l'ensemble  $\{-1, 0, 1\}$ .

L'information que représente  $C$  est quant à elle continue, il convient donc de la quantifier. La quantification doit donc être suffisamment précise pour être efficace, mais aussi suffisamment grossière pour limiter le nombre de compteurs mis en jeu, et donc l'entropie.

La quantification la plus adaptée à ce compromis est tout d'abord étudiée. Puis les différents modèles de probabilités conditionnelles sont mis en œuvre avec un codeur arithmétique adaptatif initialisé à partir d'une loi uniforme. Finalement une initialisation *a priori* utilisant les statistiques des contours codés aux résolutions inférieures sera également testée.

#### 4.1.3.1 Quantification de $C$

Sur la figure 4.5, le débit moyen du codage des contours extraits en utilisant les informations  $B$  et  $C$ , pour les dix images de test sur trois niveaux de résolution, est représenté en fonction du pas de quantification de  $C$  en considérant une loi de quantification uniforme. L'ordre de Markov de  $B$  varie de zéro à trois. Pour un ordre égal à zéro, cela revient à n'utiliser que l'information venant de la résolution inférieure  $C$ . Dans ce cas, le meilleur pas de quantification, celui qui minimise le débit, est obtenu pour  $20^\circ$ . À partir de là, plus l'ordre de Markov augmente, plus le pas de quantification minimisant le débit augmente pour compenser l'augmentation du nombre de compteurs utilisés par le codeur. Toutefois, les résultats montrent que globalement, en comparant les débits obtenus en utilisant différents ordres de Markov pour  $B$ , le débit minimal est déjà obtenu pour un ordre de Markov nul. Ceci est contraire à nos attentes considérant les entropies précédemment calculées, mais s'explique aisément par l'augmentation considérable de la combinatoire des valeurs prises par  $B$  et  $C$ .

#### 4.1.3.2 Comparaison avec les méthodes de l'état de l'art

Les résultats du codage arithmétique adaptatif initialisé avec une loi uniforme utilisant les probabilités conditionnelles de  $A$  étant donnés  $B$ ,  $C$  et  $B \cap C$  sont présentés figure 4.6 pour chaque image de test. Les résultats pour  $Pr(A|B)$  représentant les méthodes de l'état de l'art [PL97, CS95], sont présentes avec un ordre de Markov pour  $B$  variant de zéro à trois. Exceptés les cas des images **baboon** et **barbara**, la méthode utilisant comme contexte l'information de l'image à la résolution inférieure décodée ( $Pr(A | C)$ ) surpasse les méthodes utilisant comme contexte l'information des éléments de contours précédemment codés/décodés ( $Pr(A|B)$ ). Ces deux exceptions peuvent être

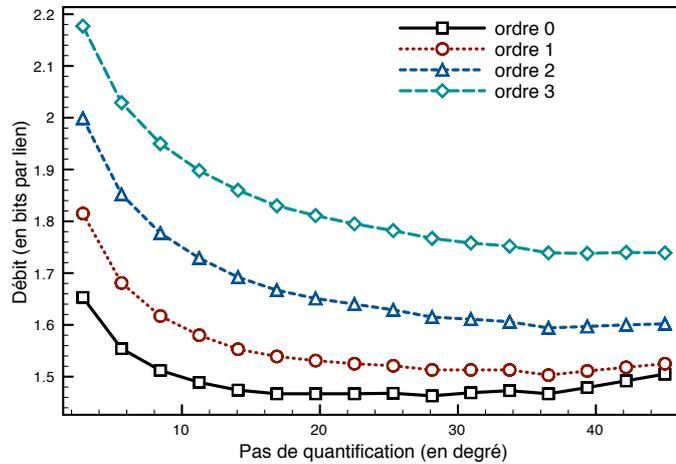


FIGURE 4.5 – Débit moyen du codage des chemins des contours extraits pour les dix images de test sur trois niveaux de résolution, utilisant les probabilités conditionnelles  $Pr(A | B \cap C)$ , où l'ordre de Markov de  $B$  varie de 0 à 3, en fonction du pas de quantification de  $C$ .

expliquées par le fait que de nombreux contours à pleine résolution (la fourrure du singe et la texture du pantalon de Barbara) ne sont pas présents à la résolution inférieure. Dès lors l'information d'orientation du gradient à cette résolution est inutile.

Contrairement aux résultats de l'étude entropique précédente, le fait d'utiliser tous les contextes simultanément ( $B \cap C$ ) ne permet pas de diminuer plus le coût de codage qu'en utilisant seulement le contexte  $B$  ou  $C$ . Ceci s'explique par une dilution des contextes : le nombre élevé de contextes différents fait tendre les lois de probabilités conditionnelles de  $A$  étant donné  $B \cap C$  vers une loi uniforme.

#### 4.1.3.3 Initialisation du codeur arithmétique avec une loi *a priori*

De façon à améliorer l'apprentissage du codeur, nous proposons de tirer encore parti de l'approche multi-résolution de la représentation utilisée. À cette fin le codeur est initialisé à partir des statistiques observées calculées pour les chemins des contours extraits aux résolutions inférieures. En pratique, les compteurs du codeur ne sont tout simplement pas réinitialisés (remis à un) entre le codage des différents niveaux de résolution. Les résultats d'une telle initialisation comparés à l'approche uniforme sont présentés sur la figure 4.7 et une légère amélioration y apparaît.

## Conclusion

Dans le contexte de codage de contours extraits à différentes résolutions d'une image. Une méthode pour coder les chaînes d'éléments de contours à un niveau de résolution donné, et en utilisant l'information disponible aux résolutions inférieures, a été

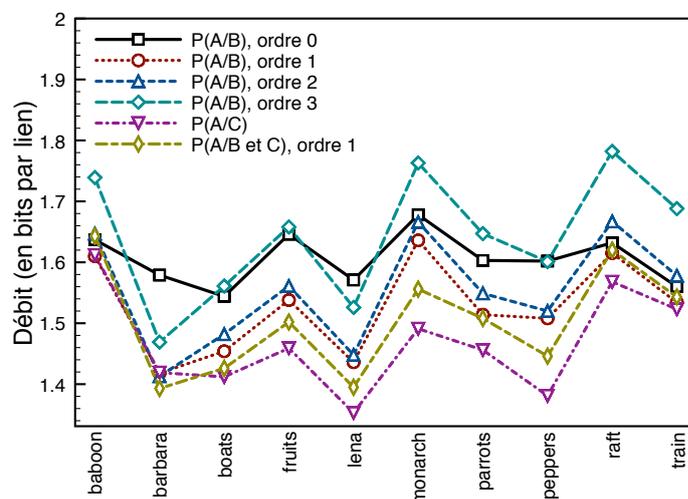


FIGURE 4.6 – Codage arithmétique adaptatif initialisé avec une loi uniforme utilisant les probabilités conditionnelles de  $A$  étant donné  $B$ ,  $C$  et  $B \cap C$  pour chaque image de test.

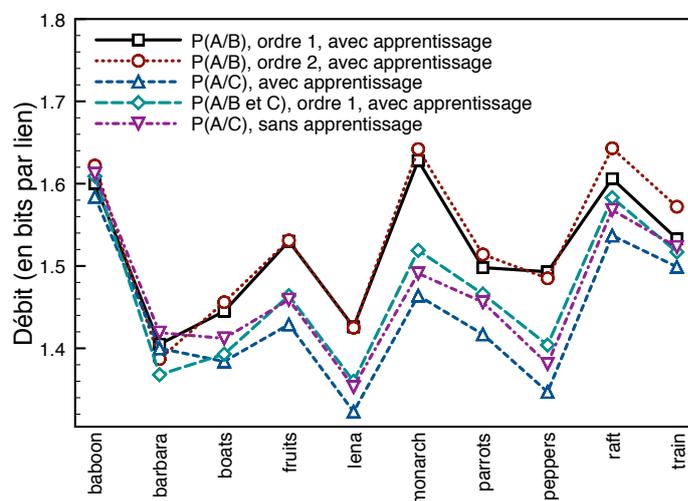


FIGURE 4.7 – Comparaison du codage arithmétique adaptatif des chemins des contours extraits sur les dix images des tests en utilisant comme probabilités conditionnelles  $Pr(A | C)$  initialisé à partir d'une loi uniforme,  $Pr(A | B \cap C)$  et  $Pr(B)$  tous deux initialisés à partir des statistiques des chemins des contours codés/décodés aux résolutions inférieures.

proposée[JRB07b].

En effet l'écart entre la direction orthogonale à l'orientation du gradient à la résolution inférieure, et l'orientation d'un morceau du contour déjà codé/décodé à la résolution courante, est corrélé avec la position du prochain élément du contour à coder/décoder.

Une étude préliminaire de l'entropie des probabilités conditionnelles en présence a permis de confirmer l'intérêt d'une telle approche.

En pratique l'implantation par un codeur arithmétique adaptatif de notre méthode surpasse les méthodes de l'état de l'art utilisant un modèle de chaîne de Markov d'ordre 1 à 3. Toutefois la combinaison de notre méthode en collaboration avec un modèle de Markov n'a pas été concluant pour des problèmes de dilution de contexte. Malgré cela la diminution mesurée du débit avec notre méthode par rapport à un modèle de chaîne de Markov d'ordre 1, est en moyenne de 3% sur une moyenne de dix images et atteint jusqu'à 8%.

Notre méthode est toutefois moins efficace que la méthode de l'état de l'art dans certains cas où une structure extraite à la résolution courante n'est pas présente à la résolution inférieure. Pour ces cas de figure, il pourrait être intéressant de mettre en compétition les deux méthodes au lieu de les utiliser en combinaison.

## **4.2 Codage arithmétique adaptatif des arbres quaternaires définissant les orientations de filtrage et les classes de structure**

Aussi bien pour la méthode d'estimation des orientations de filtrage et de classification basée sur une extraction préalable de contours (section 3.2.1), que pour la méthode basée sur une estimation de l'orientation du gradient (section 3.2.2.2), les arbres quaternaires décrivant ces deux informations partagent, par construction, la même structure. Dans cette section, trois implantations du codage de ces arbres quaternaires sont proposées et décrites succinctement, afin de pouvoir les comparer dans la section suivante avec le coût de codages des chaînes de contours de la méthode basée sur une extraction de contours.

Ces trois méthodes s'appuient sur un codage arithmétique adaptatif. La structure de l'arbre précisant si un noeud est terminal (feuille) ou non est codée en fonction de la profondeur dans l'arbre des noeuds considérés. Ainsi la probabilité qu'un noeud soit une feuille est directement liée à sa profondeur dans l'arbre. Dans nos implantations, ces probabilités sont initialisées par des lois uniformes, et s'adaptent au fur et à mesure du codage/décodage. Ces probabilités pourraient être initialisées, par exemple, en augmentant la probabilité qu'un noeud soit une feuille plus on s'éloigne de la racine de l'arbre. Dans les résultats qui seront présentés, cette propriété sera prise en compte par l'adaptabilité du codeur seulement, et non à son initialisation.

Les valeurs des feuilles de l'arbre sont elles aussi codées arithmétiquement, avec un contexte de passé d'ordre 1.

Dans une première implantation, les valeurs d'orientation de filtrage et de classe d'activité sont codées indépendamment les unes des autres avec comme contexte les

valeurs de la dernière feuille codée.

Dans les deux implantations suivantes, nous tenons compte du fait que pour des régions de l'image classifiées comme « uniforme » ou « texture isotrope », une orientation de filtrage est arbitrairement fixée (la direction selon l'horizontale). Ainsi pour ces feuilles, la valeur de l'orientation de filtrage n'est pas transmise.

Dans la première de ces deux implantations, le contexte de passé est celui précédemment décrit (les valeurs de la dernière feuille codée). Toutefois cette dernière feuille codée n'est pas nécessairement proche dans l'arbre de la feuille à coder, le parcours de l'arbre s'effectuant prioritairement de la racine vers les feuilles, puis seulement ensuite, pour les quatre fils d'un nœud, de haut vers le bas et de gauche à droite (voir figure 4.8-b).

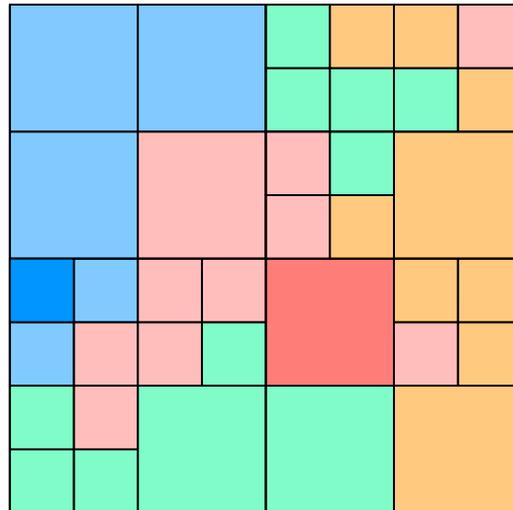
La troisième implantation utilise alors comme contexte l'information des feuilles déjà codées/décodées jouxtant la feuille à coder/décoder. À cette fin, une carte des classes et des orientations de filtrage déjà codées/décodées est mise à jour au fur et à mesure du codage/décodage. Par exemple, pour une feuille des arbres quaternaires correspondant à une région de l'image de coordonnées  $(tlx, tly)$  (coordonnées du coin haut-gauche (*top-left*) du bloc), on considère les valeurs de classe et d'orientation situées aux coordonnées d'abscisse  $tlx - 1$  ou d'ordonnée  $tly - 1$  déjà codées/décodées. La valeur la plus représentée est alors choisie comme valeur contextuelle (voir figure 4.8-c).

### 4.3 Comparaison du coût de codage entre les deux méthodes

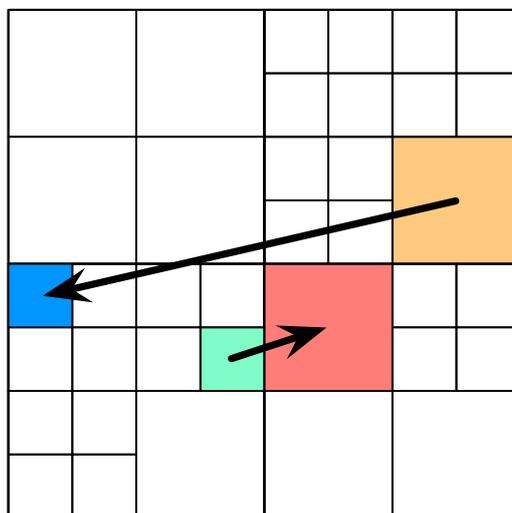
Nous comparons ici le coût de codage des éléments de contours extraits d'une image, avec le coût de codage des structures en arbre calculées à partir de ces mêmes contours. Le passage d'une description de l'information, les contours, à une autre, les deux arbres quaternaires, n'est évidemment pas réversible. L'information est quantifiée plus ou moins finement via les paramètres que sont les seuils utilisés pour la classification des blocs de l'image, et surtout via la taille minimale possible de ces blocs qui correspond aussi à la profondeur maximale des arbres quaternaires. Les résultats proposés dans le tableau 4.1 ont été obtenus avec les paramètres suivants :

- $\theta$ -redressement avec les filtres d'ondelettes 5/3 et les filtres de prédiction/mise à jour  $L1$  (filtres à deux échantillons unitaires), 16(+2) orientations de filtrage pour la première dimension du redressement réparties uniformément (+2 correspond aux orientations  $-45$  et  $45$  pouvant être implantées de deux façons différentes) ;
- seuils utilisés pour la norme du gradient des éléments de contours : 5% et 10% de la norme maximale du gradient ;
- taille minimale des contours extraits : 8 éléments de contours ;
- pourcentages d'éléments de contours appartenant à la même classe d'orientation utilisées : 75% et 50%.

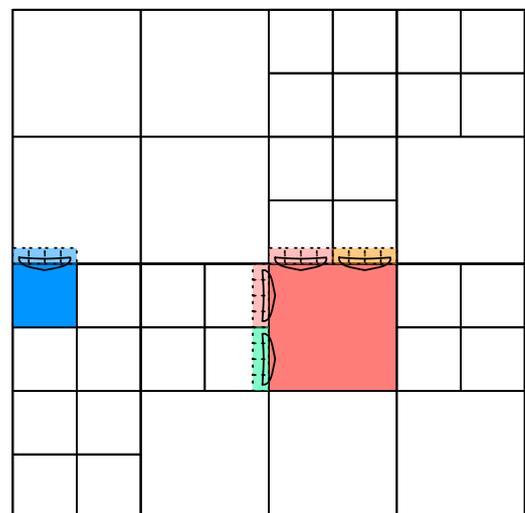
Trois observations peuvent être faites à propos des résultats présentés dans le tableau 4.1.



(a) Exemple d'arbre quaternaire des classes structurelles.



(b) Contexte : dernière feuille décodée.



(c) Contexte : voisinage spatial dont les feuilles ont déjà été décodées.

FIGURE 4.8 – Différents contextes utilisés pour le codage arithmétique des arbres quaternaires.

Tout d'abord concernant le coût de codage des éléments de contours, on observe que le coût de codage des chemins des chaînes de contours, c'est-à-dire les éléments des chaînes d'indice supérieur à 2 (0 représentant la tête), avoisine le coût de codage des têtes uniquement.

Ensuite concernant le coût de codage des arbres quaternaires, les deux dernières méthodes de codage sont évidemment plus efficaces car, par un choix arbitraire d'orientation de filtrage pour certaines classes d'activités, la quantité d'information à transmettre a été réduite. La troisième méthode est légèrement plus efficace que la deuxième. Ceci s'explique facilement par une meilleure prise en compte du contexte de codage qui fait toujours référence pour la troisième méthode à un voisinage spatial proche, ce qui n'est pas si systématique pour la deuxième méthode.

Enfin même pour une taille minimale des blocs de 4 pixels, le coût de codage des arbres reste toujours inférieur au coût de codage des éléments de contours. Par ailleurs, on peut observer que la variance du coût de codage des éléments de contours est beaucoup plus forte que celle du coût de codage des arbres quaternaires. En outre, un coût plus élevé que les autres pour les éléments de contours ne signifie pas un coût plus élevé pour les arbres associés. Cette non corrélation s'explique aisément par le fait qu'une région présentant plusieurs contours parallèles (exemple théorique), coûtera nécessairement plus cher que si cette région ne contient qu'un seul de ces contours, sans pour autant apporter plus d'information au niveau des arbres quaternaires.

En conclusion le coût de codage des chaînes de contours restent beaucoup plus coûteux que le codage des arbres quaternaires associés. Ceci est dû pour partie à la redondance liée aux contours présentant une même orientation dans une même région de l'image. On perçoit facilement que ceci peut aussi compliquer le lien à établir entre coût de codage des contours et distorsion de l'image reconstruite.

## Conclusion

À l'aide d'un codeur arithmétique adaptatif, l'orientation du gradient donnée à une résolution inférieure d'une image permet de diminuer le coût de codage des chaînes d'éléments de contour par rapport à une approche basée uniquement sur l'information des  $N$  derniers éléments de contours codés/décodés. Un couplage des deux approches ne permet d'amélioration à cause d'un effet de dilution de contexte.

Les contours extraits d'une image permettent de classifier les régions de l'image et de déterminer l'orientation principale au sein de ces régions qui correspond à l'orientation utilisée ensuite pour le filtrage de la région. Ces deux informations, représentées par des arbres quaternaires, peuvent être codées en lieu et place des contours, là aussi avec un codage arithmétique adaptatif basé sur une exploitation de la cohérence spatiale des nœuds de l'arbre précédemment codés/décodés.

Une comparaison des deux solutions de codage permet de conclure à un coût bien moindre du codage des arbres quaternaires représentant les classes structurelles et les orientations de filtrage par rapport au codage en amont des éléments de contours déterminant ces arbres.

image	Codage des chaînes d'éléments de contour				Codage des arbres quaternaires					
	tête	suc tête	chemin	total contour	8 × 8			4 × 4		
					arbres 1	arbres 2	arbres 3	arbres 1	arbres 2	arbres 3
baboon	2266	335	2292	4893	1720	1432	1413	4193	3487	3469
barbara	2056	213	2791	5060	1361	1099	1060	3219	2604	2540
boats	1483	187	2034	3704	1207	987	964	2951	2422	2372
farm	1399	196	1409	3004	1124	912	894	2412	1972	1945
fruits	891	126	1370	2387	1249	1001	993	2692	2187	2160
lake	1210	182	1305	2697	1123	905	889	2726	2233	2206
lena	1218	145	1591	2954	1163	920	895	2848	2280	2227
monarch	1081	166	1570	2817	1088	886	856	3028	2502	2465
parrots	752	106	1012	1870	798	645	629	1789	1466	1439
peppers	750	106	1065	1921	1071	874	862	2230	1822	1799
plane	1008	152	1184	2344	885	735	733	2270	1882	1872
raft	1919	285	2142	4346	1560	1317	1298	3967	3348	3332
train	1988	297	2190	4475	1438	1204	1183	3716	3108	3086
yosemite	1256	184	1212	2652	1118	937	928	2496	2086	2075
coût moyen	1376,93	191,43	1654,79	3223,14	1207,5	989,57	971,21	2895,5	2385,64	2356,21
débit moyen par pixel	0,0420	0,0058	0,05050	0,0984	0,0368	0,0302	0,0296	0,0884	0,0728	0,0719

TABLE 4.1 – Coûts de codage de l'information des contours extraits (« tête » : coût de codage des premiers éléments de chaque chaîne de contour, « suc tête » : coût de codage des premiers successeurs de la tête de chaque chaîne de contour, « chemin » : coût de codage des autres éléments de contours) et de l'information des orientations de filtrage et de classification au travers des arbres quaternaires codés selon les trois méthodes proposées, et des tailles minimales des blocs de 8 × 8 ou 4 × 4. Les coûts de codage sont ici présentés en octets, sauf le débit moyen en bits par pixel de l'image.



## Chapitre 5

# Quantification adaptée à la classification structurelle de l'information

### Introduction

Dans la première partie nous avons présenté des méthodes d'estimation des orientations de filtrage et de classification structurelle des régions d'une image. Le chapitre précédent a permis de comparer le coût de codage de ces différentes méthodes. L'objectif de ce chapitre est maintenant de tirer parti de la méthode d'estimation des orientations de filtrage et de classification qui s'est révélée la moins coûteuse (estimation basée sur la mesure du gradient) pour le procédé de quantification des éléments des différentes sous-bandes produites par la transformée en ondelettes orientées.

À cette fin nous rappellerons les principes de l'application de la quantification scalaire aux coefficients d'ondelettes ainsi que certaines améliorations que permet la norme JPEG 2000. Puis nous exploiterons les caractéristiques de la transformées en ondelettes orientées couplées avec la connaissance a priori du contenu d'une image pour adapter les lois de quantification appliquées aux sous-bandes de la transformée en ondelettes orientées.

### 5.1 Quantification et codage des éléments d'une représentation en ondelettes standard 2D

Dans un système de compression « avec pertes », c'est-à-dire où l'image reconstruite après décodage (même sans erreur de transmission) n'est pas identique à l'originale. L'opération de quantification associe à chaque valeur prise par la donnée à quantifier son représentant le plus proche dans un dictionnaire de taille finie. Ce sont les indices de ces représentants dans le dictionnaire qui sont codés. L'image est reconstruite au décodeur en associant à chaque indice décodé la valeur du représentant correspondant

dans le dictionnaire.

Nous décrirons ici seulement quelques mécanismes basiques de quantification des coefficients de la transformée en ondelettes 2D autorisés dans JPEG 2000[TM01]. Après un bref rappel de la quantification scalaire uniforme, nous inventorierons quelques aspects de JPEG 2000 Part 1 et 2 pour tenir compte des effets de masquage visuel.

### 5.1.1 Quantification scalaire uniforme

La quantification scalaire uniforme est définie par un pas de quantification fixé  $\Delta$ . À toute valeur  $x \in \mathbb{R}$  est associé un indice de quantification  $q \in \mathbb{Z}$  et un représentant  $\hat{x} \in \mathbb{R}$  :

$$q = Q(x) = \text{signe}(x) \cdot \left\lfloor \frac{|x|}{\Delta} \right\rfloor \quad (5.1)$$

$$\hat{x} = Q^{-1}(q) = \begin{cases} 0 & q = 0 \\ \text{sign}(q) \cdot (|q| + \delta) \cdot \Delta & q \neq 0 \end{cases} \quad (5.2)$$

Les valeurs comprises dans l'intervalle  $[-\Delta, \Delta]$  sont quantifiées à zéro. Cet intervalle est appelé « zone morte » [TM01]. La taille de la « zone morte » est ici de  $2\Delta$ , contrairement aux autres distances entre plus proches représentants qui est de  $\Delta$ . La valeur de  $\delta$ , comprise entre 0 et 1, paramètre la valeur des représentants. Typiquement  $\delta = 1/2$ , comme illustré sur la figure 5.1, la valeur d'un représentant est alors choisie au milieu de chaque intervalle  $[k \cdot \Delta, (k + \text{signe}(k)) \cdot \Delta]$ ,  $k \in \mathbb{Z}^*$ . Toutefois, dû à la distribution des sources, on constate que la valeur  $\delta = 0,375$  minimise en pratique l'erreur quadratique moyenne de l'image reconstruite.

La taille de la zone morte peut être agrandie ou réduite. La quantification est alors définie en fonction d'un paramètre  $nz$  en vue d'une zone morte de dimension  $2 \cdot (1 - nz) \cdot \Delta$  au lieu de  $2\Delta$  :

$$q = Q_{nz}(x) = \begin{cases} 0 & |x| < -nz \cdot \Delta \\ \text{signe}(x) \cdot \left\lfloor \frac{|x| + nz \cdot \Delta}{\Delta} \right\rfloor & |x| \geq -nz \cdot \Delta \end{cases} \quad (5.3)$$

$$\hat{x} = Q_{nz}^{-1}(q) = \begin{cases} 0 & q = 0 \\ \text{sign}(q) \cdot (|q| - nz + \delta) \cdot \Delta & q \neq 0 \end{cases} \quad (5.4)$$

Si  $nz$  est positif, alors la taille de la zone morte est réduite, sinon elle est au contraire agrandie. Marcellin *et al* [MLB<sup>+</sup>02] indique que dans JPEG2000, une valeur d'environ 0,25 pour  $nz$  (une zone morte de mesure  $1,5\Delta$ ) permet entre autre une légère diminution de l'erreur quadratique moyenne (EQM) de l'image reconstruite.

### 5.1.2 Prise en compte des effets de masquage

Le terme de masquage visuel est employé pour désigner une propriété du système visuel humain : le seuil différentiel de visibilité d'un stimulus n'est pas fixe, il est lié

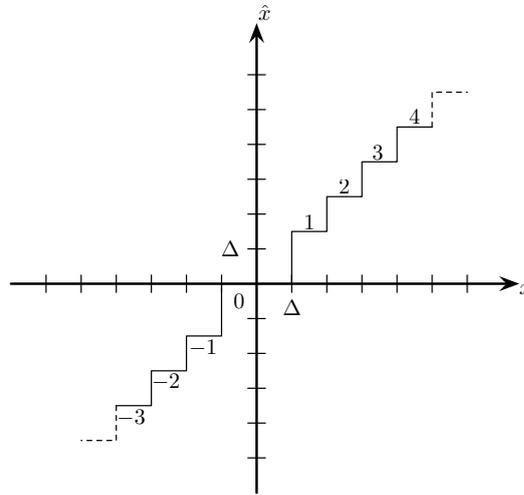


FIGURE 5.1 – Loi de quantification scalaire uniforme avec zone morte (valeur du pas de quantification  $\Delta$ ).

à la valeur moyenne locale du signal d'image environnant le stimulus, appelé signal masquant. Il dépend aussi d'autres caractéristiques du signal masquant, induisant généralement une augmentation du seuil différentiel de visibilité.

Dans le contexte de la quantification, on peut considérer que le signal masquant est l'image originale, et le signal masqué correspond aux erreurs introduites par la quantification. La capacité à estimer correctement les effets de masquage au sein d'une image a pour objectif de maximiser la réduction de l'information à transmettre tout en minimisant les distorsions perçues. Cela peut correspondre entre autre à déterminer le pas de quantification le plus grand qui n'introduit pas de changement perceptible dans l'image.

Plusieurs solutions existent dans JPEG 2000 pour prendre en compte ces effets. Nous nous intéresserons tout d'abord à deux solutions faisant appel à un modèle simple où le signal masquant est considéré comme n'appartenant qu'à un domaine fréquentiel restreint en un site donné sans tenir compte des autres signaux environnants. Le potentiel de masquage d'un tel signal dépend à la fois de sa fréquence et de son amplitude.

Une première solution consiste donc à pondérer les pas de quantification d'un signal à une fréquence donnée en fonction de la fonction de sensibilité au contraste (CSF : *Contrast Sensitivity Function*). Un exemple d'une telle fonction est donnée figure 5.2. Cette fonction indique la sensibilité aux variations de contrastes pour une fréquence spatiale donnée. L'œil humain est ainsi par exemple plus sensible aux variations d'amplitude d'un signal moyenne fréquence, que haute fréquence.

Les fonctions de sensibilité au contraste étant définies en fonction de la fréquence spatiale en cycles par degré, elles dépendent de la distance à laquelle sont visualisés les signaux. Pour pouvoir être appliquées aux coefficients de la transformée en ondelettes il convient donc tout d'abord de définir à quelle distance sera perçue l'image

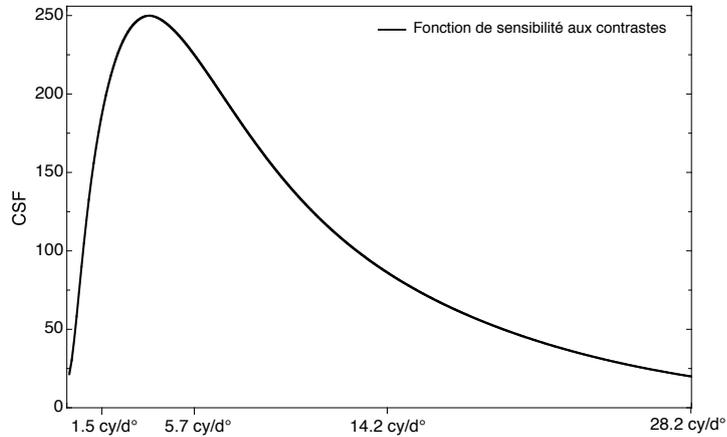


FIGURE 5.2 – Exemple de fonction de sensibilité aux contrastes.

codée/décodée. Ensuite une valeur extraite de la fonction de sensibilité au contraste est associée à chaque sous-bande de la transformée en ondelettes. Les méthodes pour définir une telle valeur sont variées (sensibilité moyenne sur la bande fréquence, sensibilité maximale, etc.). Ainsi différentes tables de pondérations peuvent être trouvées dans la littérature pour différentes distances de visualisation (exemples tables 5.1 tirées de [TM01]).

D'autre part, il a été déterminé expérimentalement [ZDL02] que l'effet de masquage dans une décomposition en ondelettes est une fonction monotone croissante (monotone si on ne considère pas l'effet de facilitation) de l'amplitude du coefficient d'ondelettes. Ainsi plus l'amplitude d'un coefficient d'ondelettes est élevée, plus l'effet de masquage est important, et donc plus le bruit de quantification peut être élevé sans introduire d'artefact visuel. On peut donc se permettre une quantification plus grossière pour les valeurs élevées de coefficients d'ondelettes. Toutefois, plutôt que d'appliquer directement une quantification non uniforme sur les valeurs des coefficients d'ondelettes, on peut appliquer par exemple une fonction puissance ( $y = x^\alpha, 0 < \alpha \leq 1$ ) sur les valeurs des coefficients d'ondelettes puis quantifier avec une loi de quantification uniforme. Une valeur conseillée de l'exposant est  $\alpha = 0,7$ .

Toutefois cette modélisation du masquage n'est valable que si un coefficient d'ondelettes représente un seul et unique signal en terme de fréquences spatiales. Or la décomposition en ondelettes est telle que ce n'est pas nécessairement le cas.

La figure 5.3 met en évidence les fréquences spatiales mises en jeu dans chaque sous-bande de la transformée en ondelettes. Le pire exemple est donné par les sous-bandes de hautes fréquences à la fois selon l'horizontale et la verticale qui contiennent des détails diagonaux qui correspondent à des signaux orientés autour de  $-45^\circ$  ( $135^\circ$ ) et  $45^\circ$ . Un coefficient d'ondelettes d'une sous-bande diagonale peut représenter une superposition

Pondérations CSF pour une distance de visualisation de 1000 pixels.

Niveau	CH	CV	CD
0	0,756353	0,756353	0,573057
1	0,998276	0,998276	0,996555
2	1,000000	1,000000	1,000000
3	1,000000	1,000000	1,000000
4	1,000000	1,000000	1,000000

Pondérations CSF pour une distance de visualisation de 1700 pixels.

Niveau	CH	CV	CD
0	0,307191	0,307191	0,108920
1	0,861593	0,861593	0,742342
2	1,000000	1,000000	1,000000
3	1,000000	1,000000	1,000000
4	1,000000	1,000000	1,000000

Pondérations CSF pour une distance de visualisation de 3000 pixels.

Niveau	CH	CV	CD
0	0,038487	0,038487	0,003075
1	0,410628	0,410628	0,182760
2	0,921045	0,921045	0,848324
3	1,000000	1,000000	1,000000
4	1,000000	1,000000	1,000000

TABLE 5.1 – Pondérations CSF pour trois distances de visualisation.

de signaux dont certains sont orientés autour de  $-45^\circ$  et d'autres autour de  $45^\circ$ . Or, pour le système visuel humain, les signaux orientés à  $45^\circ$  ne masquent pas ceux qui sont orientés selon  $-45^\circ$  et inversement. Par conséquent, une superposition de tels signaux antagonistes entraînant une valeur de coefficient d'ondelettes élevée est considérée à tort comme très masquante par le modèle ci-dessus. L'effet masquant est donc sur-estimé, la quantification résultante est finalement trop grossière : les distorsions seront visibles.

Le fait qu'un signal d'une orientation donnée ne masque pas un autre signal de fréquence angulaire trop éloignée se retrouve aussi pour les autres sous-bandes à cause de l'empiètement des signaux plutôt diagonaux sur les sous-bandes des détails horizontaux et verticaux, et réciproquement.

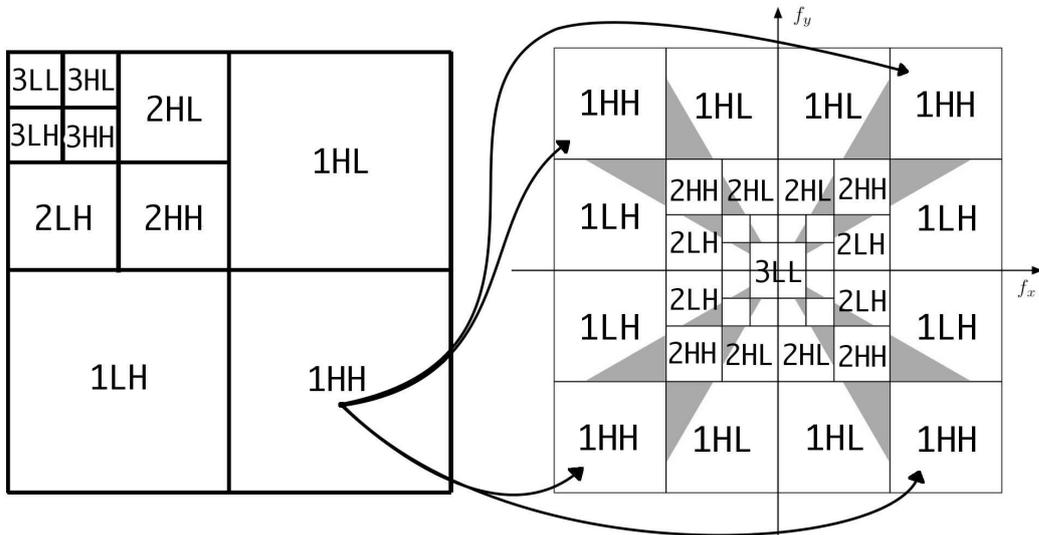


FIGURE 5.3 – À gauche : la décomposition en ondelettes 2D en sous-bandes représentées dans un espace spatio-fréquentiel. À droite : la décomposition en ondelettes dans l'espace fréquentiel.

En fait, fondamentalement, l'effet de masquage est généré par les caractéristiques du signal masquant sur une extension spatiale plus importante que celle du stimulus sur lequel porte son effet. Dans le « point-wise extended masking », chaque coefficient [ZDL00] va donc être quantifié en fonction de sa propre valeur, mais aussi en fonction des coefficients d'ondelettes dans un voisinage causal considérant que plus l'activité locale de l'image est importante, plus l'effet de masquage est important. Après l'application d'une fonction puissance comme décrite ci-dessous :

$$x_i \rightarrow y_i = \text{sign}(x_i) \cdot |x_i|^\alpha, \quad (5.5)$$

le résultat est normalisé par un facteur de masquage spatial en fonction de l'amplitude des signaux voisins :

$$z_i = y_i / \left(1 + \frac{a}{|\phi_i|} \sum_{k \text{ voisin de } i} |\hat{x}_k|^\beta\right) \quad (5.6)$$

où  $|\phi_i|$  est la taille du voisinage causal,  $a$  est un facteur de normalisation avec une valeur constante égale à  $(10000/2^{\text{bitdepth}-1})^\beta$  et  $\text{bitdepth}$  le nombre de bits par pixel de l'image d'origine,  $\hat{x}_k$  est la valeur quantifiée des coefficients voisins. Les paramètres  $\beta$  et  $N$  déterminent le degré de masquage visuel. Ensuite, une loi de quantification uniforme est appliquée sur  $z_i$ .

Cette méthode permet de distinguer les coefficients de grandes amplitudes qui se situent dans une zone texturée de ceux qui se situent dans une structure de contours. Toutefois son inconvénient est de ne pas respecter le principe de scalabilité car la quantification d'un élément dépend de ses voisins causaux décodés. Voisins qui peuvent donc avoir des valeurs différentes selon que la reconstruction est plus ou moins complète. Une méthode pour contrer ce phénomène consiste à ne retenir que les  $n$  bits les plus significatifs de chaque coefficient voisin pour calculer le facteur de masquage visuel pour ainsi augmenter les chances de disposer des mêmes valeurs au codage et au décodage.

## 5.2 Impact d'une modification de la quantification selon la classe de structure

Notre objectif est de tirer parti des connaissances a priori que nous avons de la classification des régions de l'image ainsi que des propriétés de filtrage de la transformée en ondelettes orientées. Pour cela, nous comparons la transformée en ondelettes classique 2D sur 5 niveaux de décomposition avec une transformée en ondelettes orientées sur 5 niveaux de décomposition, implantée par technique de redressement utilisant une interpolation linéaire au quart de pixel maximum. Pour la transformation en ondelettes orientée, les orientations de filtrage sont déterminées à pleine résolution pour une taille de bloc de  $16 \times 16$  pixels avec la méthode décrite dans la section 3.2.2.2 du chapitre 3, et les orientations retenues sont utilisées pour les 3 premiers niveaux de décomposition. Pour les deux derniers niveaux, une transformation en ondelettes classiques 2D est appliquée. Les deux transformées, orientées ou non, sont implantées par redressement avec les coefficients :

$$\begin{aligned} \alpha &\approx -1,586134342 \\ \beta &\approx -0,05298011854 \\ \gamma &\approx 0,8829110762 \\ \delta &\approx 0,4435068522 \\ \zeta &\approx 1,149604398 \end{aligned}$$

correspondant aux filtres d'ondelettes 9/7, et avec les filtres de prédiction et de mise à jour  $L1$  ( $\begin{pmatrix} 1 & 1 \end{pmatrix}$ ). Toutefois, pour que les filtres d'analyse d'ondelettes en question soient normalisés à 1, les valeurs des éléments des deux transformées en ondelettes sont

divisées par deux après chaque niveau de décomposition. Mais tout d'abord une transformation affine est appliquée sur l'image originale pour que son intervalle de valeurs soit compris entre  $-\frac{1}{2}$  et  $\frac{1}{2}$  :  $x \mapsto \frac{x}{255} - \frac{1}{2}$ .

Afin de quantifier les éléments des transformées en ondelettes, une base de pas de quantification en fonction de l'orientation de filtrage sera proposée. Puis nous étudierons les lois de probabilité des éléments de chaque sous-bande de la transformée en ondelettes orientées considérée par rapport à la transformée en ondelettes classique. Finalement, nous pondérerons les pas de quantification de base en fonction de la classe de structure de façon à adapter la quantification au contenu structurel des images.

### 5.2.1 Base de pas de quantification pour la transformée en ondelettes orientées

Le pas de la loi de quantification appliquée sur une sous-bande est défini comme étant inversement proportionnel à la racine carrée de l'énergie de l'image reconstruite avec uniquement un élément non nul égal à 1 dans cette sous-bande [TM01]. Les pas de quantification pour les différentes sous-bandes sont donc réglés mutuellement en fonction de la valeur de cette énergie et modulés par un facteur général de quantification noté  $\lambda$ .

Ces pas de quantification dépendent par conséquent des filtres d'ondelettes utilisés pour la synthèse, et en particulier de la somme des carrés des éléments de ces filtres. Considérant la transformée en ondelettes orientée présentée ci-dessus, c'est-à-dire avec interpolation linéaire, bien que les sommes des coefficients des filtres de synthèse restent constantes en fonction de l'orientation de filtrage, les sommes des coefficients élevés au carré varient. Par conséquent, le pas de quantification utilisé pour un coefficient d'ondelettes dépend des orientations de filtrage qui ont été utilisées pour aboutir à l'obtention d'un tel coefficient.

La valeur du pas de quantification pour une sous-bande en fonction de ces orientations de filtrage est obtenue par reconstruction d'une transformée où tous les éléments de toutes les sous-bandes sont nuls sauf un élément de la sous-bande considérée qui est unitaire. Les pas de quantification pour toutes les sous-bandes sont calculés une fois pour toute, mais doivent être stockés. Afin de limiter la combinatoire sur le nombre de pas de quantification, à calculer et à stocker, nous faisons l'approximation que le pas de quantification d'un coefficient d'ondelettes orientées  $c(x, y)$  au niveau de décomposition  $l$  ne dépend que des orientations de filtrage utilisées pour filtrer les éléments de coordonnées  $(2^k \cdot x, 2^k \cdot y)$ ,  $k \in \{0, \dots, l-1\}$ . Ainsi nous négligeons le fait que le coefficient d'ondelettes orientées  $c(x, y)$  au niveau  $l$  dépend de coefficients voisins des éléments de coordonnées  $(2^k \cdot x, 2^k \cdot y)$  au niveau  $k$  qui peuvent avoir été filtrés avec une orientation différente.

Le choix d'utiliser pour les trois premiers niveaux de décomposition la même carte des orientations de filtrage fait que le pas de quantification d'un coefficient d'ondelettes orientées de coordonnées  $(x, y)$  au niveau de décomposition  $l$  dépend alors uniquement de l'orientation de filtrage du bloc  $16 \times 16$  contenant le pixel de coordonnées  $(2^l \cdot x, 2^l \cdot y)$  à pleine résolution.

Les pas de quantification ainsi calculés sont donnés dans le tableau 5.2. Pour chaque

orientation		4				3			2			1			0		
1	2	$\Delta_{EE}$	$\Delta_{OE}^4$	$\Delta_{EO}^4$	$\Delta_{OO}^4$	$\Delta_{OE}^3$	$\Delta_{EO}^3$	$\Delta_{OO}^3$	$\Delta_{OE}^2$	$\Delta_{EO}^2$	$\Delta_{OO}^2$	$\Delta_{OE}^1$	$\Delta_{EO}^1$	$\Delta_{OO}^1$	$\Delta_{OE}^0$	$\Delta_{EO}^0$	$\Delta_{OO}^0$
(1, -1)	(0, 1)	0, 0292	0, 0294	0, 0281	0, 0300	0, 0586	0, 0597	0, 0573	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481
(1, - $\frac{3}{4}$ )	(0, 1)	0, 0292	0, 0292	0, 0281	0, 0294	0, 0586	0, 0592	0, 0568	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(1, - $\frac{1}{2}$ )	(0, 1)	0, 0293	0, 0292	0, 0283	0, 0289	0, 0586	0, 0581	0, 0579	0, 119	0, 127	0, 112	0, 25	0, 269	0, 241	0, 493	0, 530	0, 448
(1, - $\frac{1}{4}$ )	(0, 1)	0, 0294	0, 0291	0, 0286	0, 0287	0, 0586	0, 0572	0, 0585	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(1, 0)	(0, 1)	0, 0295	0, 0291	0, 0291	0, 0288	0, 0586	0, 0586	0, 0581	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481
(1, $\frac{1}{4}$ )	(0, 1)	0, 0294	0, 0291	0, 0286	0, 0287	0, 0586	0, 0572	0, 0585	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(1, $\frac{1}{2}$ )	(0, 1)	0, 0293	0, 0292	0, 0283	0, 0289	0, 0586	0, 0581	0, 0579	0, 119	0, 127	0, 112	0, 25	0, 269	0, 241	0, 493	0, 530	0, 448
(1, $\frac{3}{4}$ )	(0, 1)	0, 0292	0, 0292	0, 0281	0, 0294	0, 0586	0, 0592	0, 0568	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(1, 1)	(0, 1)	0, 0292	0, 0294	0, 0281	0, 0300	0, 0586	0, 0597	0, 0573	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481
(1, 1)	(1, 0)	0, 0292	0, 0281	0, 0294	0, 0300	0, 0597	0, 0586	0, 0573	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481
( $\frac{3}{4}$ , 1)	(1, 0)	0, 0292	0, 0281	0, 0292	0, 0294	0, 0592	0, 0586	0, 0568	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
( $\frac{1}{2}$ , 1)	(1, 0)	0, 0293	0, 0283	0, 0292	0, 0289	0, 0581	0, 0586	0, 0579	0, 119	0, 127	0, 112	0, 25	0, 269	0, 241	0, 493	0, 530	0, 448
( $\frac{1}{4}$ , 1)	(1, 0)	0, 0294	0, 0286	0, 0291	0, 0287	0, 0572	0, 0586	0, 0585	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(0, 1)	(1, 0)	0, 0295	0, 0291	0, 0291	0, 0288	0, 0586	0, 0586	0, 0581	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481
(- $\frac{1}{4}$ , 1)	(1, 0)	0, 0294	0, 0286	0, 0291	0, 0287	0, 0572	0, 0586	0, 0585	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(- $\frac{1}{2}$ , 1)	(1, 0)	0, 0293	0, 0283	0, 0292	0, 0289	0, 0581	0, 0586	0, 0579	0, 119	0, 127	0, 112	0, 25	0, 269	0, 241	0, 493	0, 530	0, 448
(- $\frac{3}{4}$ , 1)	(1, 0)	0, 0292	0, 0281	0, 0292	0, 0294	0, 0592	0, 0586	0, 0568	0, 119	0, 121	0, 117	0, 25	0, 255	0, 251	0, 494	0, 509	0, 465
(-1, 1)	(1, 0)	0, 0292	0, 0281	0, 0294	0, 0300	0, 0597	0, 0586	0, 0573	0, 120	0, 120	0, 120	0, 25	0, 250	0, 258	0, 494	0, 494	0, 481

TABLE 5.2 – Pas de quantifications en fonction de l'orientation de filtrage.

sous-bande et chaque niveau sont donnés les pas de quantification de base en fonction du couple d'orientations de filtrage. La première orientation correspond à l'orientation de régularité choisie, définie au quart de pixel, la seconde orientation correspond à l'orientation horizontale ou verticale la plus éloignée. Les ordres de grandeur des pas de quantification pour un niveau donné ne sont pas fondamentalement différents d'une sous-bande à l'autre car les filtres d'ondelettes sont normalisés. D'un niveau à l'autre on observe par contre une variation d'un facteur 2. Les pas de quantifications  $\Delta_{OE}^k$  et  $\Delta_{EO}^k$  peuvent être évidemment différents l'un de l'autre pour un niveau  $k$  en fonction des orientations de filtrage comme expliqué ci-dessus. Ces pas sont différents aussi pour les deux derniers niveaux de la décomposition bien qu'une transformée en ondelettes classique soit appliquée pour ces niveaux. En effet une sous-bande d'un de ces niveaux est reconstruite de niveau en niveau décroissant, et donc en utilisant la transformée en ondelettes orientées pour les trois niveaux de plus haute résolution.

### 5.2.2 Histogrammes des coefficients d'ondelettes orientées

Les lois de probabilité des coefficients d'ondelettes orientées ont été étudiées en fonction de chacune des quatre classes structurelles : uniforme, mono-orientée, multi-orientée et texturée isotrope. À cette fin, les histogrammes des coefficients d'ondelettes et des coefficients d'ondelettes orientées calculés sur une base de dix images (voir annexe C) sont donnés figures 5.4 à 5.6 pour chaque classe de structure et pour les trois premiers niveaux de décomposition. Le pas de quantification utilisé pour l'obtention de ces histogrammes est fixé à 0,001 pour toutes les sous-bandes et pour tous les niveaux. Cette valeur a été choisie pour fournir un pas suffisamment faible. Elle correspond à un dixième de la moyenne sur les dix images de l'écart-type des coefficients d'ondelettes le plus faible calculé pour chaque niveau et chaque sous-bande. En l'occurrence ici, cette valeur est obtenue pour la sous-bande diagonale du premier niveau de décomposition.

Sur chaque figure sont représentés les histogrammes des indices de quantification des coefficients de la transformée en ondelettes (TO) et de la transformée en ondelettes orientée (TOO) des trois sous-bandes : OE, EO et OO. Pour les régions correspondant aux classes uniformes et textures isotropes, OE correspond pour les deux transformées aux coefficients des détails verticaux (CV), EO aux coefficients des détails horizontaux (CH) et OO aux coefficients des détails diagonaux (CD). Pour ces deux classes, l'orientation de filtrage de la transformée en ondelettes orientées est fixée à l'horizontale. Les résultats observés sont toutefois différents et dûs à l'influence des blocs voisins classifiés différemment. Ceci reste vrai pour les deux autres classes pour la transformée en ondelettes (TO). Par contre pour la transformée en ondelettes orientée (TOO), OE correspond aux hautes fréquences selon la direction de régularité et basses fréquences selon la direction complémentaire (noté HFReg), EO correspond à l'inverse (noté HFOrtho), et OO correspond aux hautes fréquences selon les deux directions (noté HFRes).

L'étude de ces figures permet d'aboutir à deux observations. La première concerne la dynamique générale de toutes les sous-bandes pour les deux transformées en fonction de la classe structurelle. Sans surprise par exemple, la distribution des probabilités pour les régions classifiées comme uniformes est fortement plus resserrée que pour les régions

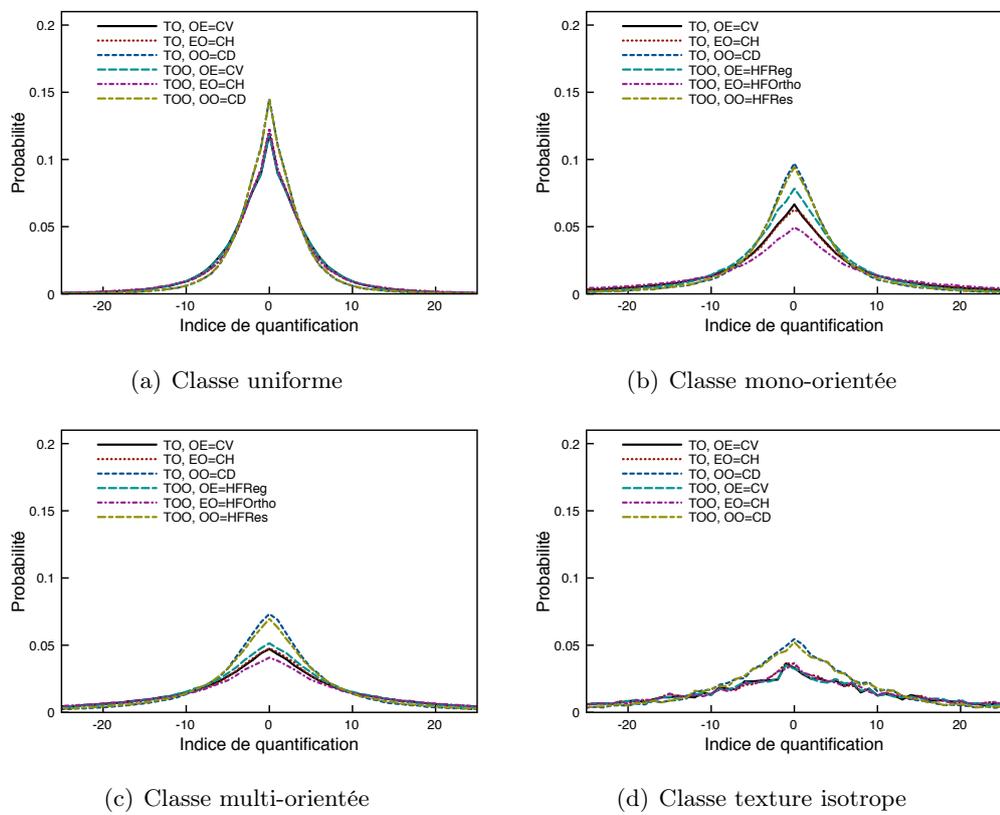
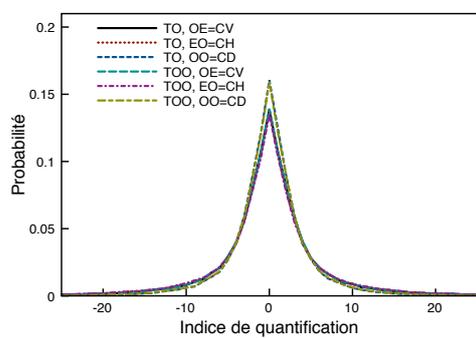
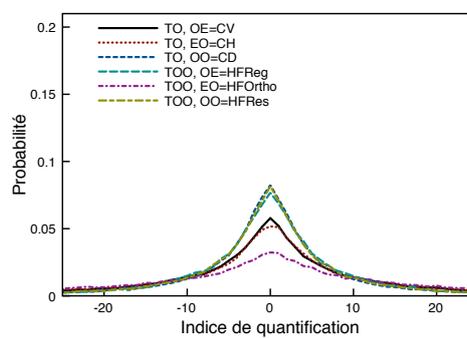


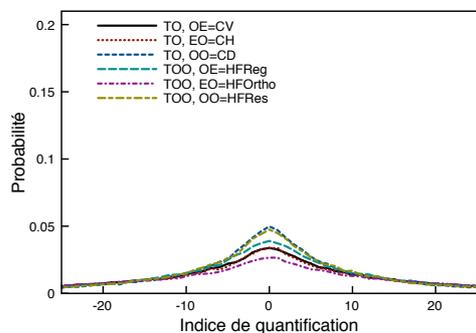
FIGURE 5.4 – Histogrammes des coefficients d'ondelettes pour le premier niveau de décomposition (ensemble de la base de dix images).



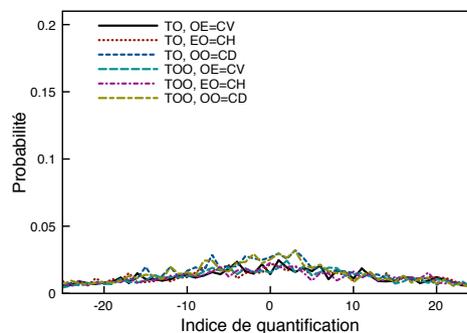
(a) Classe uniforme



(b) Classe mono-orientée



(c) Classe multi-orientée



(d) Classe texture isotrope

FIGURE 5.5 – Histogrammes des coefficients d'ondelettes pour le deuxième niveau de décomposition (ensemble de la base de dix images).

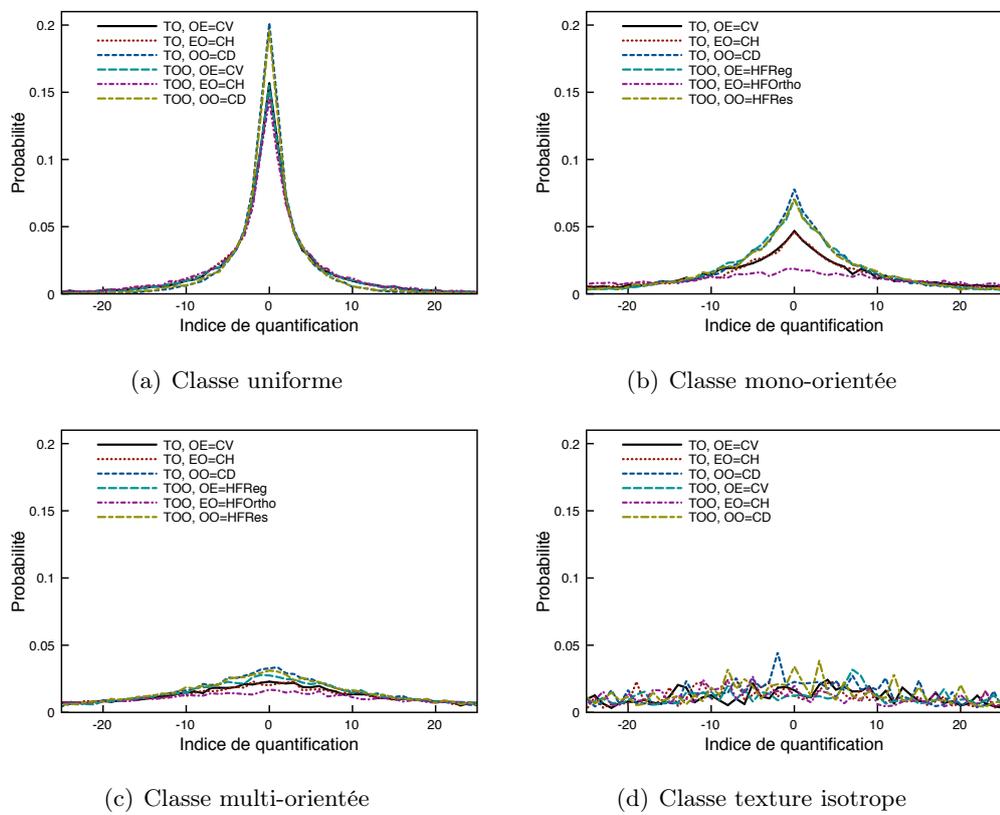


FIGURE 5.6 – Histogrammes des coefficients d'ondelettes pour le troisième niveau de décomposition (ensemble de la base de dix images).

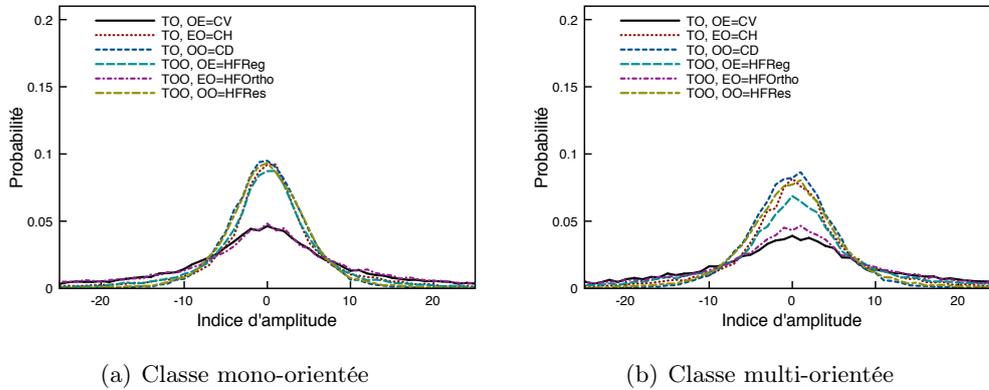


FIGURE 5.7 – Histogrammes des coefficients d'ondelettes pour le premier niveau de décomposition de l'image *boats*.

classifiées comme textures isotropes. Cette première observation peut être exploitée pour adapter le pas de quantification en fonction des classes structurales.

La seconde observation concerne la différenciation qui apparaît entre les sous-bandes HFReg et HFOrtho pour les régions classifiées mono-orientées et multi-orientées par rapport au comportement des sous-bandes CH et CV. Si les distributions concernant les sous-bandes CH et CV de la transformée en ondelettes (TO) sont très proches, les distributions concernant les sous-bandes HFReg et HFOrtho de la transformée en ondelettes orientée (TOO) diffèrent systématiquement de la même manière. Nous avons déjà observé que l'énergie de la sous-bande HFReg était inférieure à celle de HFOrtho. Les distributions des probabilités pour les deux classes concernées donnent une vision plus complète de la répartition des valeurs des coefficients d'ondelettes. Il convient ici de rappeler que ces résultats sont issus de moyennes sur une base de dix images. En effet, si l'on s'intéresse aux distributions concernant une image donnée, on est susceptible d'observer la chose suivante : si l'image contient plus de contours orientés verticalement que horizontalement (par exemple l'image *boats*), la distribution des sous-bandes CH et CV seront sensiblement différentes, la distribution de CH se rapprochant de celle de HFReg, celle de CV de HFOrtho (voir figure 5.7).

Toutefois pour une image quelconque, le codage après transformée en ondelettes ne dispose pas de cette information et ne peut donc considérer que le cas moyen où les distributions des probabilités pour les sous-bandes CH et CV sont considérées identiques. Ces deux sous-bandes sont donc quantifiées de la même façon avec la transformée en ondelettes. Par contre, la différenciation qui apparaît entre les distributions moyennes des coefficients de la transformée en ondelettes orientée des sous-bandes HFReg et HFOrtho va permettre un traitement différent de ces deux sous-bandes.

Ces résultats laissent à penser qu'une politique différente de quantification selon chaque classe structurale peut permettre de tirer partie de la classification des régions.

Les régions, telles que les zones uniformes, où la distribution des coefficients est très resserrée, peuvent être quantifiées plus finement pour éviter par exemple l'apparition

de zones d'aplat (apparition de faux contours).

Au contraire, les régions telles que les zones texturées peuvent être quantifiées plus grossièrement et ainsi profiter de l'effet masquage de ce type de zone.

En ce qui concerne les régions mono-orientées ou multi-orientées, appliquer le même raisonnement reviendrait donc à considérer qu'il faut quantifier plus finement les hautes fréquences selon la direction de régularité, et plus grossièrement les hautes fréquences selon la direction complémentaire. Toutefois nous verrons qu'en présence d'un contour contrasté, l'information de hautes fréquences selon la direction de régularité, d'énergie faible comparée à celle des hautes fréquences selon la direction complémentaire, peut être négligée en partie, ceci dépendant du niveau de décomposition de la transformée.

La politique de quantification des différentes sous-bandes de la transformée en ondelettes orientée repose donc sur deux idées :

- tenir compte de l'aspect statistique, à savoir modifier les pas de quantification selon l'allure des distributions des densités de probabilités en fonction des sous-bandes et de la classe d'orientation (augmenter le pas de quantification lorsque ces distributions sont étalées, le réduire lorsqu'elles sont resserrées) ;
- tenir compte de la prépondérance de l'information d'une sous-bande par rapport à une autre à savoir que pour les régions classées comme mono-orientées, par exemple, l'information des hautes fréquences selon la direction de régularité est moins fondamentale que l'information des hautes fréquences selon la direction complémentaire horizontale ou verticale.

### 5.2.3 Pondération des pas de quantification en fonction de la classe structurelle

Afin d'exploiter la segmentation de l'image en des blocs  $16 \times 16$  et leur classification, les pas de quantification de base  $\Delta_s^l(o)$ , qui dépendent de l'orientation de filtrage  $o$ ,  $s \in \{OE, EO, OO\}$  et  $l \in 0, 1, 2$ , sont pondérés par les poids  $W_{cs}^l$  avec  $c$  représentant l'une des quatre classes : uniforme ( $u$ ), mono-orientée ( $mo$ ), multi-orientée ( $mu$ ) et texture isotrope ( $t$ ). L'objectif de cette pondération est d'adapter localement le pas de quantification des coefficients de la transformée en ondelettes orientées en exploitant uniquement les caractéristiques liées à la classification structurelle ou à la nature de la transformée elle-même. À titre d'exemple, utiliser une pondération où tous les poids  $W_{cs}^l$  d'un niveau  $l$  seraient inférieurs à 1 ne saurait exploiter l'une ou l'autre des propriétés ci-dessus mais tirer parti de la sensibilité au contraste en fonction de la fréquence spatiale, ce qui peut déjà être mis en application par une pondération des coefficients de la transformée en ondelettes à travers les pondérations données dans le tableau 5.1.

Pour une sous-bande  $s$  à un niveau  $l$ , le pas de la loi de quantification des éléments de cette sous-bande classés comme  $c$  est donné par la relation :

$$\Delta_{cs}^l = \lambda \cdot \frac{\Delta_s^l}{W_{cs}^l \cdot W_{CSF_s}^l}. \quad (5.7)$$

La détermination de la pondération idéale qui maximisera la qualité des images reconstruites, ou minimisera les distorsions introduites à débit fixé n'est pas triviale.

Ces deux notions, souvent indifférenciées, sont estimées à travers des métriques dites de qualité. Ces métriques sont généralement classées en trois catégories : les métriques avec référence complète, où l'image originale est disponible pour être comparée à l'image dont on souhaite estimer la qualité, les métriques avec référence réduite, où une partie de l'information relative à l'image originale est disponible, et enfin les métriques sans référence. Si la dernière catégorie est nécessairement une métrique de qualité, les deux premières sont généralement plutôt des métriques de similarité, dite de qualité car l'image originale est souvent considérée comme une référence de qualité maximale. En dépit de ces imperfections, la métrique de similarité certainement la plus utilisée est le PSNR.

En conséquence de quoi nous nous limiterons à étudier qualitativement l'impact visuel des modifications de la pondérations des pas de quantification en fonction des classes structurelles ainsi que des sous-bandes.

Pour démontrer l'intérêt de notre représentation structurelle des images, il convient d'exhiber des cas d'application où à débit constant la qualité des images reconstruites est meilleure, ou bien des images reconstruites de qualité identique présentant un coût de codage moindre. La preuve de l'identité en terme de qualité de deux images est ardue à démontrer sauf pour des images identiques. Si elle peut être admise d'un point de vue statistique à partir de mesures subjectives de qualité obtenue auprès d'un certain nombre d'observateurs dans des conditions normalisées, elle peut être trompeuse ponctuellement pour établir une observation fiable. Par conséquent nous cherchons à déterminer expérimentalement une pondération qui permet, au pire à débit constant, une amélioration sensible de la qualité.

La méthodologie employée pour établir une telle pondération comporte deux étapes : tout d'abord réduire le débit de codage sans altérer sensiblement la qualité d'une image, puis profiter du débit économisé pour améliorer la qualité de l'image. Pour réduire le coût de codage sans introduire de dégradations gênantes, nous étudierons l'augmentation des pas de quantification dans les classes mono-orientées, multi-orientées et texturées. L'économie réalisée sera alors réallouée pour quantifier plus finement les régions classées uniformes.

Cette méthodologie est appliquée sur quelques images de test pour un facteur général de quantification tel que les dégradations introduites sont perceptibles mais non gênantes. À cette fin, nous avons développé une application que nous avons appelé Psy-KoX (voir capture d'écran sur figure 5.8) pour pouvoir observer les variations de qualité perçue en modifiant le facteur général de quantification et les différentes pondérations structurelles.

Avant d'appliquer une quelconque pondération structurelle, étudions les images reconstruites d'une part avec une transformée en ondelettes, et d'autre part avec une transformée en ondelettes orientées. Le tableau 5.3 rend compte pour un facteur de quantification  $\lambda$  propre à chaque image, du coût de codage de chaque méthode ainsi que quelques mesures de qualité. Ce coût de codage est exprimé à travers l'entropie des éléments d'ondelettes, ainsi que du débit du train binaire obtenu après un codage SPIHT[SP96]. Ce codeur a été choisi pour des raisons de simplicité d'implantation.

La valeur du facteur général de quantification  $\lambda$  varie jusqu'à un rapport 2 entre

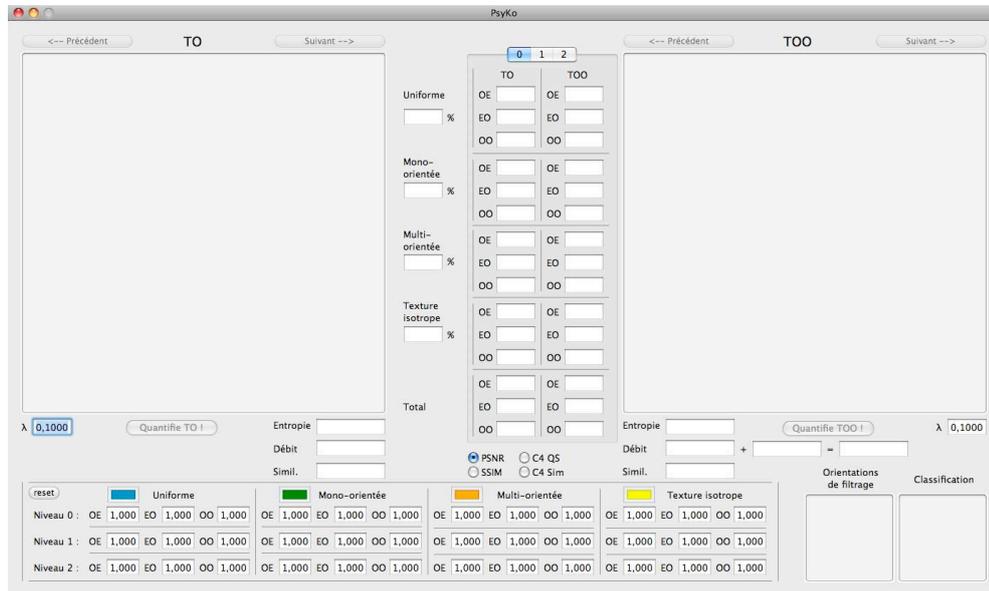


FIGURE 5.8 – Capture d'écran de l'application PsyKoX.

les images **baboon** et **boats**. À qualité équivalente donc, l'image **boats** est quantifiée deux fois plus finement que l'image **baboon**. Ceci peut s'expliquer en partie par la différence d'importance des effets de masquage. Pour l'image **boats**, pour un facteur  $\lambda = 0,075$ , les dégradations apparaissent en priorité sur les mâts et les cordages. Pour ces régions, l'effet de masquage de l'erreur de quantification d'un stimulus orienté est principalement dû aux contenus des sous-bandes de même orientation. Par contre pour l'image **baboon**, pour un facteur  $\lambda = 0,075$ , les dégradations ne sont pas perceptibles. Pour les moustaches comme pour le pelage, le signal masquant pour un stimulus orienté dans ces zones n'est pas restreint à un contenu fréquentiel de même orientation. Les effets de masquage apportés par ce contenu plus riche a tendance à être plus important, et donc diminuer la visibilité des erreurs de quantification.

Pour chaque image certaines différences notables entre les images reconstruites peuvent être remarquées :

**baboon** les fentes du museau du babouin, orientées selon  $45^\circ$  et  $135^\circ$ , sont de meilleure qualité avec la TOO (voir figure 5.9-a et b). Les moustaches sont aussi plus nettes en particulier les diagonales (figure 5.9-c et d). La texture de la joue gauche du babouin apparaît orientée après TOO, mais isotrope après TO (figure 5.9-e et f) ;

**barbara** pas de différence significative si ce ne sont les parties des bras orientées diagonalement (figure 5.9-g et h) ;

**boats** pas de différence significative à part les mats fortement contrastés orientés selon la diagonale (figure 5.9-i et j) ;

**monarch et train** pas de différence significative.

image	$\lambda$	TO					TOO						
		entropie	débit	PSNR	SSIM	C4	entropie	débit			PSNR	SSIM	C4
								coef.	géo.	total			
baboon	0,150	0,591	0,604	25,95	0,783	0,851	0,569	0,595	0,020	0,615	24,62	0,756	0,831
barbara	0,100	0,565	0,507	31,63	0,901	0,880	0,579	0,518	0,010	0,528	30,24	0,894	0,860
boats	0,075	0,450	0,440	34,97	0,929	0,914	0,466	0,456	0,010	0,466	33,51	0,925	0,900
farm	0,100	0,418	0,432	31,03	0,824	0,894	0,409	0,433	0,010	0,443	30,06	0,828	0,873
monarch	0,150	0,299	0,293	30,82	0,920	0,880	0,302	0,299	0,010	0,309	29,32	0,914	0,877
train	0,125	1,105	1,114	27,01	0,867	0,913	1,100	1,116	0,010	1,126	24,48	0,827	0,886

TABLE 5.3 – Débit et distorsion de différentes images quantifiées après une transformée en ondelettes (TO) ou une transformée en ondelettes orientées (TOO) sans pondération structurelle pour des facteurs généraux de quantification  $\lambda$  impliquant une qualité perçue après codage avec la transformée en ondelettes comparable pour les différentes images.

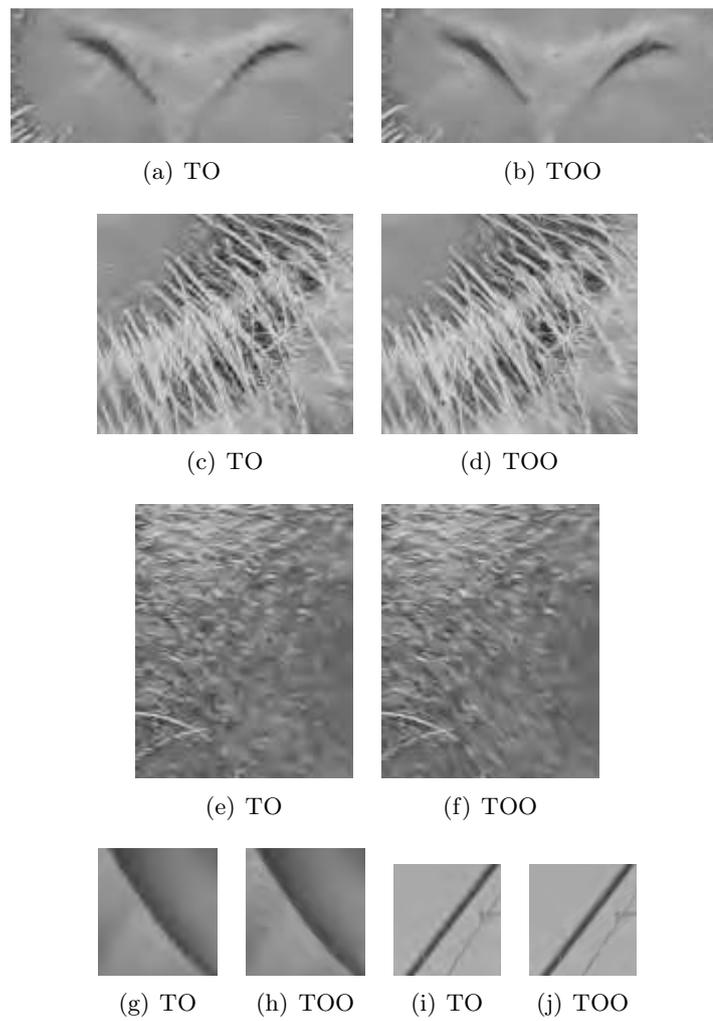


FIGURE 5.9 – Amélioration notable de la qualité pour les contours orientés selon  $45^\circ$  ou  $135^\circ$ .

De ces quelques résultats, trois observations peuvent être tirées de la reconstruction d'image après quantification avec un facteur général  $\lambda$  qui confirment les conclusions du chapitre 2 :

- la TOO seule n'apporte un avantage notable en terme de qualité de l'image reconstruite par rapport à la TO que pour les contours ou textures orientés selon  $45^\circ$  ou  $135^\circ$ . Les seules orientations de filtrage ne faisant pas intervenir d'interpolation et qui ne sont pas utilisées par la TO ;
- l'entropie est en moyenne plus élevée avec la TOO. Ceci peut s'expliquer par l'effet d'étalement des filtres d'analyse et de synthèse pour des orientations de filtrage faisant intervenir l'interpolation linéaire ;
- l'écart entre l'entropie et le coût réel de codage avec SPIHT est plus important pour la TOO que pour la TO. Ceci était prévisible car la cohérence des coefficients d'ondelettes entre les différents niveaux de résolution a pu être déstabilisée par le filtrage orienté.

Afin de réduire le coût de codage des coefficients de la transformée en ondelettes orientées, sans altérer la qualité de l'image reconstruite, nous nous intéressons d'abord à la négligence des coefficients de hautes fréquences selon la direction de plus forte régularité pour les classes mono-orientées et multi-orientées. Au premier niveau de décomposition, cette information peut être totalement supprimée sans introduire des dégradations visuelles. Au deux autres niveaux cette opération radicale ne peut par contre être répétée sans introduire de dégradations. Toutefois si l'on compare ces dégradations avec une image reconstruite avec la TO pour un débit quasi identique, on constate l'effet lisseur directionnel de l'opération. L'image apparaît plus floue, mais les orientations principales en sortent renforcées (voir figure 5.10). Hormis pour des raisons de différence de sensibilité aux fréquences spatiales, cette différence de comportement selon le niveau de décomposition s'explique par la capacité des ondelettes orientées à concentrer l'énergie dans la sous-bande EO (HFOrtho) au détriment de la sous-bande OE (HFReg). Cette capacité est plus forte pour le premier niveau de décomposition que pour les deux suivants. Ceci peut être en partie dû au fait que l'estimation de la géométrie est faite à pleine résolution. Il conviendrait donc de comparer les résultats avec une estimation de cette géométrie effectuée au premier niveau ou au second niveau pour confirmer ou infirmer cette hypothèse.

Le second levier pour réduire le débit est de tirer parti de l'effet de masquage des coefficients d'ondelettes élevés en augmentant le pas de quantification des sous-bandes EO des régions mono-orientées et multi-orientées. Ceci peut être comparé au principe énoncé précédemment de l'application d'une fonction puissance qui a pour but de quantifier plus grossièrement les coefficients forts et plus finement les coefficients faibles. La différence ici est que les coefficients faibles dans les zones mono-orientées et multi-orientées sont quantifiés tout aussi grossièrement.

Le dernier levier, qui concerne les zones texturées isotropes, est similaire au précédent à ceci près qu'il concerne les trois sous-bandes des hautes fréquences de chaque niveau de décomposition.

À partir de ces considérations, la méthode employée pour définir l'ensemble des pondérations satisfaisantes, résumé sur le tableau 5.4, a été la suivante. Les images ont été

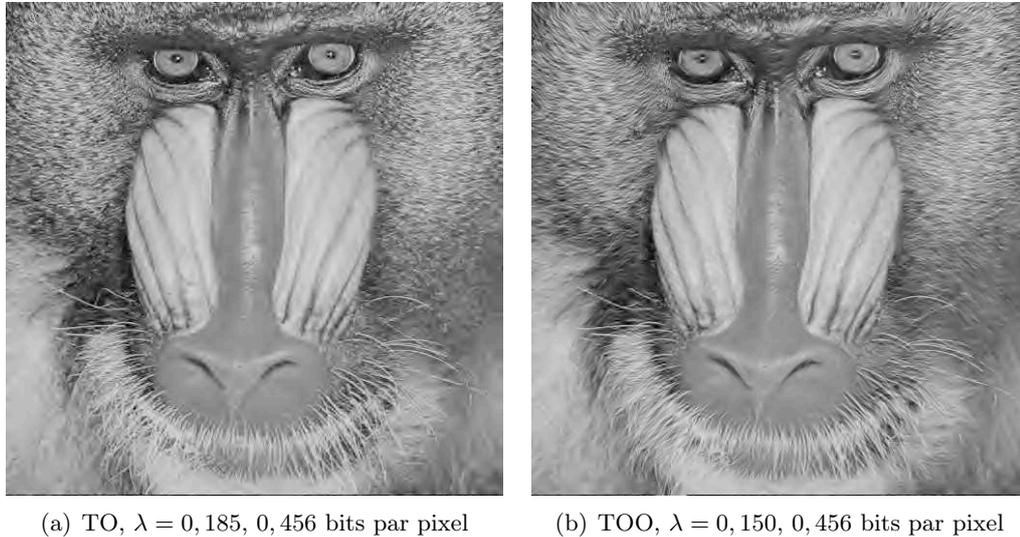


FIGURE 5.10 – Reconstruction après suppression complète de l'information de hautes fréquences selon la direction de plus grande régularité.

observées à une distance d'approximativement deux fois leur hauteur. Il s'agit d'une distance réduite d'observation. Cette distance a été choisie afin de pouvoir donner de l'importance aux hautes fréquences dans nos comparaisons de qualité car la concentration d'énergie dans une sous-bande avec les ondelettes orientées est plus efficace pour ces fréquences.

Nous avons considéré l'image **boats** car la classification des régions de cette image mettait en évidence des zones mono-orientées et multi-orientées intéressantes. En partant de la plus haute résolution, nous avons considéré à chaque fois l'augmentation des poids relatifs aux hautes fréquences dans la direction de régularité (OE et OO) pour les classes mono-orientées et multi-orientées, puis l'augmentation du poids relatif aux hautes fréquences dans la direction complémentaire (EO). Pour le premier niveau, les dégradations qui peuvent apparaître se perçoivent en premier au niveau des cordages fins. Pour les deux autres niveaux, outre les mâts et cordage, l'image est modifiée au niveau de la coque du bateau. Toutefois ces modifications, pour une pondération qui n'est pas excessivement faible, contribue à lisser certaines régions sans introduire de dégradations. Les pondérations obtenues relatives aux deux classes mono-orientée et multi-orientée peuvent être justifiées ainsi : comme illustré précédemment avec l'image **baboon**, il est normal que les poids pour les deuxième et troisième niveaux soient inférieurs aux poids du premier niveau. Le fait que le poids du deuxième niveau soit plus élevé que celui du troisième peut s'expliquer par la sensibilité au contraste qui atteint un maximum dans les moyennes fréquences qui correspond au deuxième niveau pour de telles distance de visualisation[NRK03].

Une fois ces pondérations fixées pour les classes mono-orientées et multi-orientées, nous avons considéré l'image **train** qui contient deux régions différentes texturées iso-

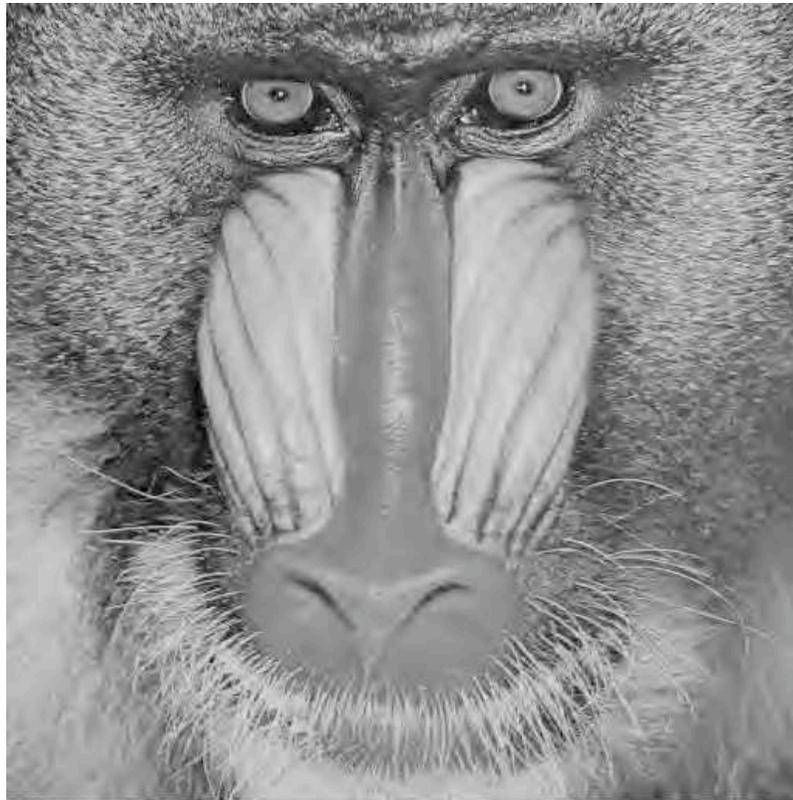
		Uniforme	Mono-orienté	Multi-orienté	Texture isotrope
	OE	2,000	0,001	0,001	0,900
Niveau 0	EO	2,000	0,750	0,750	0,900
	OO	2,000	0,001	0,001	0,900
	OE	4,000	0,800	0,750	0,800
Niveau 1	EO	4,000	1,000	0,900	0,800
	OO	4,000	0,800	0,750	0,800
	OE	4,000	0,750	0,750	0,700
Niveau 2	EO	4,000	0,900	0,800	0,700
	OO	4,000	0,750	0,750	0,700

TABLE 5.4 – Pondérations structurelles retenues pour les expérimentations.

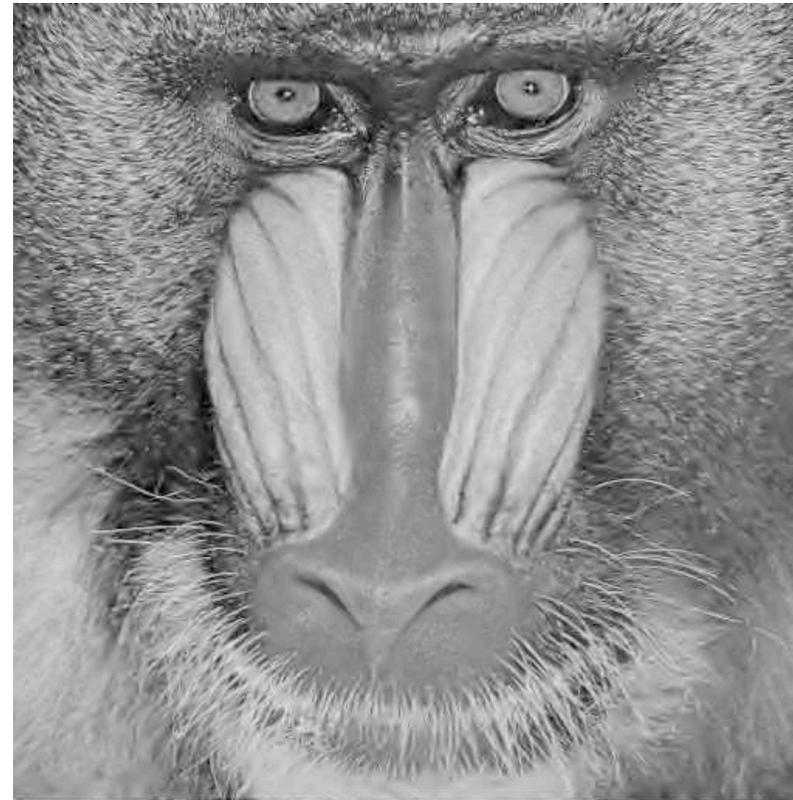
tropes : le sol et les feuilles de l'arbre. La modification de la pondération pour cette classe structurelle engendre un effet de flou. Nous avons donc réglé la pondération de façon à réduire le débit sans introduire trop ce type de dégradation. Une fois cette pondération fixée, nous avons cherché à améliorer la qualité des zones uniformes en augmentant les pondérations liées à cette classe. Il convenait alors d'augmenter ces poids en tenant compte à la fois du rapport amélioration de la qualité sur augmentation du débit, mais aussi tenir compte de la qualité des zones uniformes par rapport à la qualité du reste de l'image.

Ce système de pondérations a ensuite été appliqué à quelques images (**baboon**, **barbara**, **boats**, **farm**, **monarch** et **train**). Le facteur général de quantification pour la transformée en ondelettes standard est finalement modifié pour produire un débit avec le codeur SPIHT similaire à celui obtenu par la transformée en ondelettes orientées avec pondération structurelle. Les images résultantes sont données figures 5.11 à 5.16. Les résultats numériques sont présentés dans le tableau 5.5. Les notes de qualité ont été calculées par le PSNR, SSIM[WBSS04] et C4[CLCB08].

De façon générale, logiquement, on note des améliorations sensibles de la qualité dans les zones uniformes. Pour l'image **baboon**, le museau (voir figure 5.17) classé uniforme, est doté d'une texture qui avait disparu avec le codage par transformée en ondelettes classique. Sur la joue gauche du babouin, on remarque que pour la transformée en ondelettes standard, la texture est isotrope, alors qu'avec la transformée en ondelettes orientées, elle a conservé une orientation privilégiée (voir figure 5.18). Pour l'image **barbara**, des améliorations légères sont à noter au niveau des bras, et s'expliquent par le filtrage orienté selon  $45^\circ$ . Pour l'image **boats**, les mâts obliques n'apparaissent pas dégradés grâce au filtrage orienté. Les cordages sont mieux conservés lorsqu'ils ont été classés uniformes (cordages horizontaux dans la partie droite, voir figure 5.19) que lorsqu'ils ont été classés comme mono-orientés (cordage vertical dans la partie gauche). Ceci peut être expliqué par un manque d'efficacité du filtrage orienté lorsqu'il nécessite une interpolation et pour des contours faiblement contrastés. L'image **farm**, très texturée, permet d'utiliser le débit économisé dans ces régions pour améliorer la qualité des zones uniformes (voir figure 5.20). Pour l'image **monarch**, aucune amélioration notable n'est



(a) TO 2,  $\lambda = 0,1675$ , débit= 0,524 bits par pixel



(b) TOO 2,  $\lambda = 0,1500$ , débit= 0,528 bits par pixel

FIGURE 5.11 – Image baboon reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.



(a) TO 2,  $\lambda = 0,0950$ , débit= 0,537 bits par pixel



(b) TOO 2,  $\lambda = 0,1000$ , débit= 0,533 bits par pixel

FIGURE 5.12 – Image *barbara* reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.



(a) TO 2,  $\lambda = 0,0675$ , débit= 0,482 bits par pixel



(b) TOO 2,  $\lambda = 0,0750$ , débit= 0,486 bits par pixel

FIGURE 5.13 – Image **boats** reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.



(a) TO 2,  $\lambda = 0,0865$ , débit= 0,516 bits par pixel

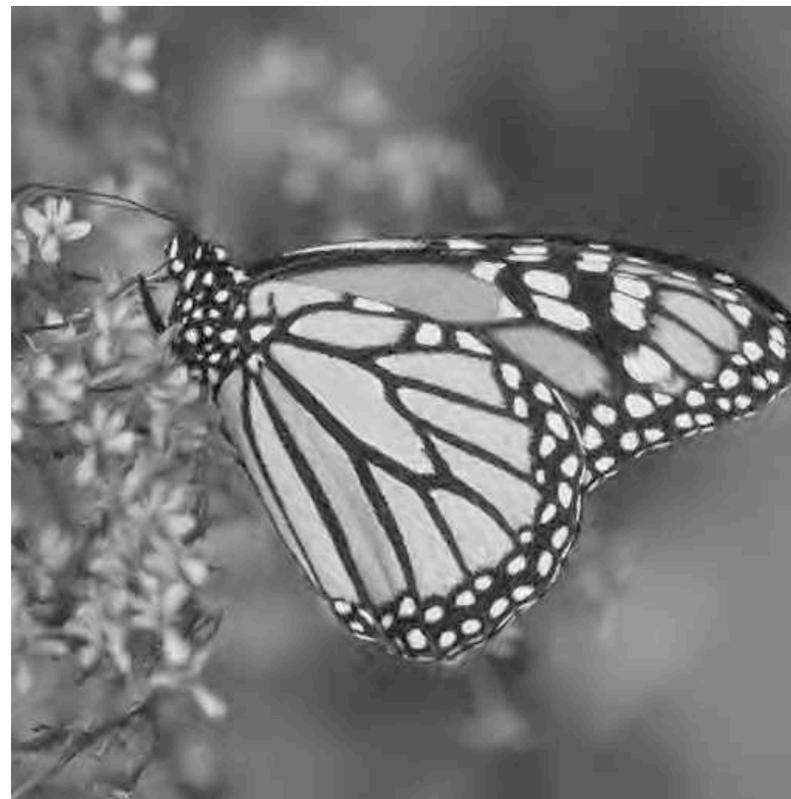


(b) TOO 2,  $\lambda = 0,1000$ , débit= 0,520 bits par pixel

FIGURE 5.14 – Image **farm** reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.



(a) TO 2,  $\lambda = 0,1550$ , débit= 0,277 bits par pixel



(b) TOO 2,  $\lambda = 0,1500$ , débit= 0,275 bits par pixel

FIGURE 5.15 – Image monarch reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.



(a) TO 2,  $\lambda = 0,1550$ , débit= 0,887 bits par pixel



(b) TOO 2,  $\lambda = 0,1250$ , débit= 0,881 bits par pixel

FIGURE 5.16 – Image **train** reconstruite après TO 2 à gauche, et TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT.

image	$\lambda$		entropie en bits par pixel				débit en bits par pixel			
	original	modifié	TO 1	TOO 1	TOO 2	TO 2	TO 1	TOO 1	TOO 2	TO 2
baboon	0,1500	0,1675	0,591	0,569	0,456	0,506	0,604	0,615	0,528	0,524
barbara	0,1000	0,0950	0,565	0,579	0,547	0,594	0,507	0,528	0,533	0,538
boats	0,0750	0,0675	0,450	0,466	0,451	0,498	0,440	0,466	0,486	0,482
farm	0,1000	0,0865	0,418	0,409	0,457	0,505	0,432	0,443	0,520	0,516
monarch	0,1500	0,1550	0,299	0,302	0,246	0,289	0,293	0,309	0,275	0,277
train	0,1250	0,1550	1,105	1,100	0,840	0,873	1,114	1,126	0,881	0,887

image	PSNR en $dB$				SSIM				C4			
	TO 1	TOO 1	TOO 2	TO 2	TO 1	TOO 1	TOO 2	TO 2	TO 1	TOO 1	TOO 2	TO 2
baboon	25,95	25,91	25,91	25,32	0,783	0,780	0,780	0,758	0,851	0,831	0,852	0,842
barbara	31,63	31,59	31,59	31,94	0,901	0,899	0,899	0,906	0,880	0,860	0,875	0,884
boats	34,97	34,69	34,69	35,59	0,929	0,926	0,926	0,935	0,914	0,900	0,906	0,902
farm	31,03	30,06	30,97	31,82	0,824	0,828	0,821	0,849	0,894	0,873	0,890	0,899
monarch	33,44	30,84	30,84	30,82	0,943	0,919	0,919	0,920	0,880	0,877	0,881	0,880
train	22,01	26,99	26,99	25,47	0,867	0,865	0,865	0,828	0,913	0,886	0,908	0,893

TABLE 5.5 – Tableaux récapitulatifs des coûts de codage et notes de qualité pour le codage d'images avec la transformée en ondelettes avec un facteur de quantification  $\lambda$  original (TO 1), la transformée en ondelettes orientées sans pondérations structurelles (TOO 1) et avec (TOO 2) (avec le facteur  $\lambda$  original) et la transformée en ondelettes standard avec le facteur de quantification  $\lambda$  modifié (TO 2) pour que le débit TO 2 corresponde à celui de TOO 2.

constatée. En général les améliorations sont perçues dans les zones uniformes, or ici les zones uniformes sont initialement très floues car la mise au point est faite sur le papillon. Par conséquent l'augmentation du débit dans ces zones n'engendre pas d'améliorations visibles. Enfin pour l'image `train`, la différence de qualité est aussi remarquable pour les régions uniformes.

Cette pondération introduit tout de même parfois des artefacts non souhaités. D'une part car pour certaines images, le surcoût introduit pour les régions uniformes a impliqué une augmentation du facteur général de quantification pour maintenir le débit constant. D'autre part une pondération trop faible pour les zones mono-orientées ou multi-orientées peut entraîner un effet de lissage directionnel parfois trop prononcé. Dans un tel cas il pourrait être intéressant d'appliquer un post-traitement pour rehausser les contours. Un défaut de classification peut aussi entraîner un traitement inadapté pour une région de l'image. Par exemple certaines régions de l'image `train` n'ont pas été classées comme uniformes (une partie du sol et une partie des arbres), et n'ont par conséquent pas pu profiter de l'augmentation de qualité résultant de l'augmentation de débit accordée à cette classe structurelle. Parfois aussi, la direction de filtrage est trop différente de l'orientation du contour ou alors une orientation secondaire a été trop négligée : un contour secondaire peut être présent lorsqu'il est considéré comme unique direction privilégiée dans une région, et disparaître brutalement à proximité d'une orientation plus importante.

Nous avons donc établi une pondération qui globalement permet d'améliorer la qualité dans les zones uniformes sans diminuer la qualité dans les autres zones, et ceci à débit constant. Si nous observons maintenant les notes de qualité données par PSNR, SSIM ou C4, nous constatons des tendances similaires pour les trois mesures. Celles-ci donnent un léger avantage à notre représentation pour les images `baboon` et `train` qui contiennent une forte proportion de l'image avec une activité orientée ou texturée isotrope très contrastée. Ceci permet d'économiser plus de débit qui est par conséquent ré-alloué aux zones uniformes. Il conviendrait de procéder à des tests subjectifs de qualité pour savoir si un ensemble d'observateurs confirment ou informent les mesures objectives.

## Conclusion

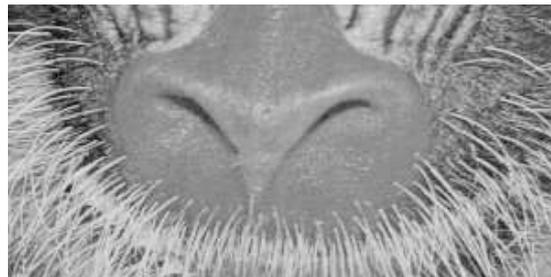
Dans ce chapitre nous avons comparé le coût de codage d'une image avec transformée en ondelettes par rapport à une transformée en ondelettes orientées implantées par redressement et interpolation jusqu'au quart de pixel.

Nous avons montré que les pas des lois de quantification dépendent, à la fois de la sous-bande et du niveau considéré, mais aussi des orientations de filtrage utilisées. Nous avons donc déterminé ces pas dans le cas d'une transformée en ondelettes orientées où la même orientation de filtrage est appliquée sur les trois premiers niveaux de décomposition pour une même région spatiale.

À partir de l'observation des lois de probabilités des éléments d'ondelettes et d'ondelettes orientées en fonction de la classe structurelle, et les caractéristiques des sous-



(a) TO 2



(b) Original



(c) TOO 2

FIGURE 5.17 – Partie de l'image **baboon** reconstruite après TO 2 à gauche, TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT, et image originale au centre.

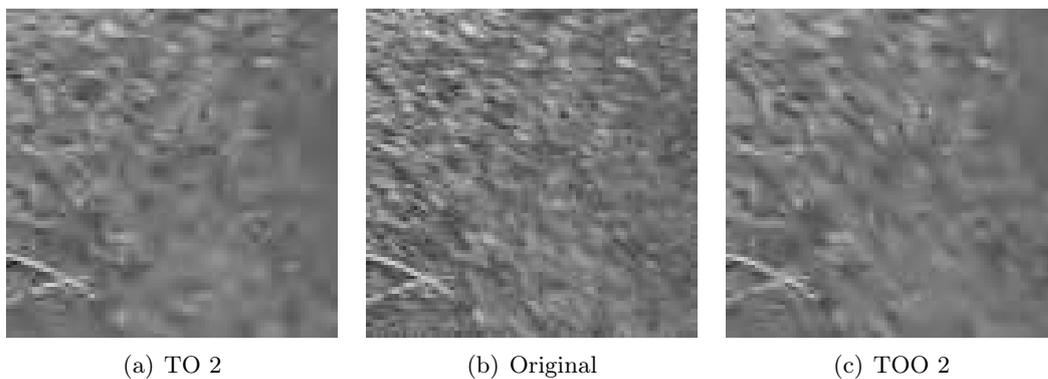


FIGURE 5.18 – Partie de l'image **baboon** reconstruite après TO 2 à gauche, TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT, et image originale au centre.

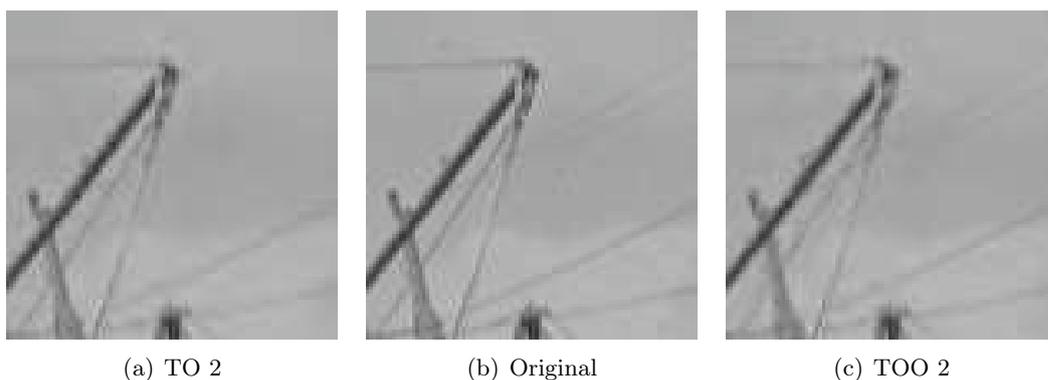


FIGURE 5.19 – Partie de l'image **boats** reconstruite après TO 2 à gauche, TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT, et image originale au centre.

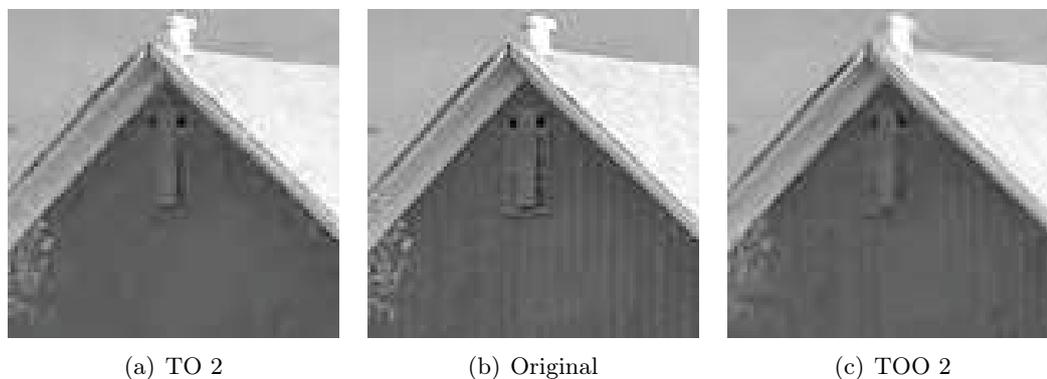


FIGURE 5.20 – Partie de l'image **farm** reconstruite après TO 2 à gauche, TOO 2 à droite en utilisant la pondération structurelle du tableau 5.4, pour un même débit avec le codeur SPIHT, et image originale au centre.

bandes de la transformée en ondelettes orientées, nous avons émis des hypothèses sur une pondération structurelle efficace en fonction de la classe de chaque région : les éléments des zones uniformes doivent être quantifiés plus finement, les éléments des zones texturées isotropes doivent être quantifiés plus grossièrement, ainsi que les éléments des sous-bandes EO (HFOrtho) des deux autres classes, et les éléments des sous-bandes OE et OO correspondant aux hautes fréquences selon la direction de régularité peuvent être plus ou moins négligés par rapport aux éléments de la sous-bande EO.

Une méthodologie expérimentale, par défaut d'une métrique objective de qualité suffisamment performante et précise, a permis de déterminer un système de pondérations structurelles qui améliore pour les images testées leur qualité à partir des hypothèses établies. Les améliorations sont particulièrement notables dans les zones uniformes grâce à la quantification plus fine permise par la réduction du débit dans les autres régions sans perte de qualité, et aussi dans les régions filtrées selon des orientations de filtres diagonales.



# Conclusion

Dans cette partie nous avons comparé le coût de codage d'une transformée en ondelettes orientées avec une transformée en ondelettes standard. Les représentations réduites des caractéristiques géométriques de l'image ont été codées de deux façons différentes. Pour l'une, les chaînes d'éléments de contours, nous avons proposé une technique de codage exploitant l'orientation du gradient de l'image décodée à la résolution précédente pour améliorer les techniques de l'état de l'art. Pour la deuxième, deux arbres quaternaires des orientations de filtrage et des classes de structures, nous avons optimisé le coût de codage avec un codage arithmétique adaptatif en utilisant le contexte spatial déjà codé. Nous avons constaté que le coût de codage des arbres quaternaires est inférieur au coût de codage des contours extraits pour une représentation réduite identique.

En tenant compte du coût de codage des arbres quaternaires, nous avons comparé le codage d'images avec la transformée en ondelettes et avec la transformée en ondelettes orientées prenant en compte des orientations de filtrage jusqu'au quart de pixel pour les trois premiers niveaux de décomposition. Pour un même facteur général de quantification pour les deux transformées tel que les dégradations des images reconstruites soient perceptibles et légèrement gênantes, on constate des nettes améliorations de qualité pour les contours et textures orientées selon les diagonales, par contre le coût de codage est en général légèrement plus élevé, même sans considérer le coût de la géométrie.

À partir de l'observation des histogrammes des éléments d'ondelettes et d'ondelettes orientées en fonction de la classe structurelle, et les caractéristiques des sous-bandes de la transformée en ondelettes orientées, nous avons émis des hypothèses sur une pondération structurelle efficace en fonction de la classe de chaque région : les éléments des zones uniformes doivent être quantifiés plus finement, les éléments des zones texturées isotropes doivent être quantifiés plus grossièrement, ainsi que les éléments des sous-bandes EO (HFOrtho) des deux autres classes, et les éléments des sous-bandes OE et OO correspondant aux hautes fréquences selon la direction de régularité peuvent être plus ou moins négligés par rapport aux éléments de la sous-bande EO.

Une méthodologie expérimentale, par défaut de métrique objective de qualité suffisamment efficace, a permis de déterminer un système de pondérations structurelles qui améliore pour les images testées leur qualité en modulant les hypothèse établies. Les améliorations sont particulièrement notables dans les zones uniformes grâce à la quantification plus fine permise par la réduction du débit dans les autres régions sans perte de qualité, et aussi dans les régions filtrées selon des orientations de filtrages diagonales.



# Conclusion générale et perspectives

## Conclusion

Les limites de la transformée en ondelettes 2D pour la représentation et le codage des images sont connues. De nouvelles représentations pour remédier à ces limites ont été déjà proposées. Nous avons choisi d'en considérer une famille : les transformées en ondelettes orientées basées sur des techniques de redressement.

L'objectif de notre travail était double. En premier, il fallait définir une représentation basée sur une de ces transformées en ondelettes orientées qui tienne compte à la fois des orientations locales des caractéristiques géométriques de l'image, mais aussi du caractère anisotrope de ces caractéristiques. Deuxièmement, il fallait concevoir une quantification adaptative et un codage prenant en compte ces caractéristiques locales très variables.

Dans une première partie, nous avons défini une représentation où localement la base est adaptative : la première orientation de décomposition ondelette est donnée par la direction estimée de régularité locale, la seconde orientation est donnée par la direction horizontale ou verticale la plus éloignée de la première orientation.

L'étude approfondie de trois implantations de techniques de redressement orienté a permis de les comparer et d'évaluer leur capacité selon la direction de régularité à concentrer l'énergie des coefficients d'ondelettes dans la sous-bande des hautes fréquences spatiales.

Lorsque le nombre d'orientations de filtrage est restreint à quatre ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ ), ces trois transformées sont équivalentes. La diminution d'énergie constatée pour des contours proches d'une diagonale est importante par rapport à l'utilisation des seules orientations horizontale et verticale.

Ces transformées se différencient pour des orientations différentes. Le gain est cependant limité par rapport à une décomposition à quatre orientations possibles. Il est significatif seulement pour des contours ou textures vraiment très contrastés. Parmi les trois techniques, celle basée sur une interpolation des échantillons permet le plus souvent une diminution de l'énergie. La variation de l'énergie est continue en fonction de l'orientation de filtrage. Par contre lorsqu'elle est efficace, la méthode correspondant à une transformée en ondelettes sur une grille discrète ré-échantillonnée est plus performante. Toutefois la variation d'énergie en fonction de l'orientation de filtrage est alors très grande et imprévisible.

Nous avons proposé et testé deux méthodes pour l'estimation des orientations locales

qui permettent en parallèle pour l'une, séquentiellement pour l'autre, de classifier en plus les régions de l'image en fonction de leurs activités directionnelles. Ces estimations ne correspondant pas nécessairement à l'orientation qui minimise réellement l'énergie, il est important qu'une faible variation de l'orientation de filtrage estimée n'entraîne pas une forte pénalisation. La technique de redressement basée sur une interpolation linéaire des valeurs des coefficients est celle qui correspond le mieux à ce critère et aussi en performance.

Ces deux formes de représentation de la géométrie de l'image sont codées par un codeur arithmétique adaptatif. La première sous la forme de chaînes d'éléments de contours, la deuxième sous la forme de deux arbres quaternaires représentant la classe de structure et l'orientation de filtrage à appliquer. Le codage des chaînes, issu de l'état de l'art, a été amélioré en moyenne de 3% sur une base de 10 images, avec même une diminution du débit de 8% en tenant compte de l'aspect multi-résolution de notre représentation. Toutefois, comparée au coût de codage des arbres quaternaires correspondant, cette réduction n'est pas suffisante. En effet, en moyenne sur les 10 images de la base de test, le coût de codage des arbres quaternaires pour une taille de blocs  $8 \times 8$  est 3,3 fois plus faible que le coût de codage des contours extraits aboutissant à ces deux arbres.

Les blocs de l'image sont classés selon une des quatre classes structurelles suivantes : uniforme, mono-orientée, multi-orientée, texturée isotrope. L'analyse des lois de probabilités des éléments d'ondelettes orientées sur les trois premiers niveaux de décomposition en fonction de la classe structurelle a mis en évidence la différence de dynamique de ces éléments, d'une part entre les différentes classes de structure, d'autre part entre la sous-bande des hautes fréquences dans la direction de régularité (HFReg) et la sous-bande des hautes fréquences dans la direction complémentaire horizontale ou verticale (HFOrtho) pour les classes mono-orientées et multi-orientées. Une méthodologie expérimentale s'appuyant sur des tests visuels de perception de dégradations a permis de déterminer un système de pondérations structurelles qui améliore la qualité visuelle des images codées de notre base. Les données des zones uniformes sont quantifiées globalement plus finement car les effets de masquage visuel dans ces zones sont quasi inexistantes. Les données des zones texturées isotropes sont quantifiées plus grossièrement. Pour les classes mono-orientées et multi-orientées, les données dans les sous-bandes des hautes fréquences selon la direction de régularité peuvent être négligées par rapport aux hautes fréquences dans la direction complémentaire pour le premier niveau de décomposition. Pour les deux autres niveaux, cette propriété n'est plus vérifiée, les dégradations dues à la quantification dans les hautes fréquences selon la direction de régularité ne sont pas totalement masquées par les signaux des hautes fréquences dans la direction horizontale ou verticale complémentaire, en particulier pour les régions multi-orientées.

De façon générale, des améliorations de qualité significatives sont observées pour les contours orientés selon  $45^\circ$  ou  $135^\circ$ . On peut se poser alors la question de la pertinence des autres orientations de filtrage, surtout que celles-ci ont tendance à étaler la valeur des coefficients ondelettes du fait de l'interpolation linéaire utilisée. La prise en compte des effets de masquage permet de réduire le coût de codage. Ces coûts ont été mesurés avec l'algorithme de codage SPIHT. Nous avons observé que la différence entre l'entropie et le débit après codage est significativement plus importante avec la transformée en

ondelettes orientées qu'avec la transformée en ondelettes standard. Les corrélations entre les différents niveaux de résolution exploitées dans le codage SPIHT sont mises à mal par la transformée en ondelettes orientées.

## Perspectives

Une première perspective de ce travail de thèse consiste en une évaluation par des tests subjectifs de la qualité des images codées/décodées avec la pondération structurelle que nous avons expérimentalement déterminée par rapport aux images codées/décodées avec la transformée standard en ondelettes. De plus, réduire le pas de quantification dans les régions uniformes où les valeurs des coefficients d'ondelettes sont faibles et augmenter ce pas dans les zones texturées où ces valeurs sont en général élevées est à rapprocher du principe d'application d'une fonction puissance sur les coefficients d'ondelettes avant quantification uniforme. Par contre, dans les zones mono-orientées et multi-orientées, les éléments des sous-bandes des hautes fréquences dans la direction de régularité qui sont en moyenne plus faibles que dans les autres sous-bandes sont quantifiés plus grossièrement avec notre pondération qu'ils ne le seraient après la fonction puissance du modèle de masquage dit de « self-masking ». Une comparaison avec ce modèle semble donc particulièrement intéressante.

Le codage des coefficients d'ondelettes orientées avec l'algorithme SPIHT est moins efficace qu'avec les coefficients d'ondelettes standards. Le codeur EBCOT offre l'avantage de travailler niveau par niveau. Toutefois certains contextes de ce codeur dépendent des caractéristiques de la transformée en ondelettes standard. Il faudrait donc les adapter pour que ce codeur soit pleinement applicable à la transformée en ondelettes orientées.

Nous avons exploité l'anisotropie de notre représentation avec la pondération structurelle des pas des lois de quantification des coefficients d'ondelettes orientées. Il est aussi possible d'utiliser des filtres d'ondelettes différents pour le filtrage selon la direction de régularité et pour la direction horizontale ou verticale complémentaire.

De part la connaissance que nous avons de l'orientation du contenu de l'image, nous disposons d'un a priori sur la structure des signaux dans les différentes sous-bandes. Cette information pourraient être mise à profit dans un contexte de quantification vectorielle.



# Annexes



## Annexe A

# Filtres d'ondelettes et paramètres de redressement associés

### A.1 Filtres 5/3

Les filtres d'analyse de l'ondelettes 5/3 sont pour le filtre passe-bas et le filtre passe-haut de réponses impulsionnelles respectivement :

$$l_a = \begin{pmatrix} -1 & 2 & 6 & 2 & -1 \end{pmatrix} \cdot \frac{\sqrt{2}}{8} \quad (\text{A.1})$$

$$h_a = \begin{pmatrix} -1 & 2 & -1 \end{pmatrix} \cdot \frac{1}{2\sqrt{2}} \quad (\text{A.2})$$

Le schéma de redressement correspondant,  $\mathcal{L}_{5/3}$ , représenté figure A.1, possède un seul et unique étage, la relation (2.12) devient dans ce cas :

$$\mathcal{L}_{5/3} = \mathcal{S}_\gamma \circ \mathcal{U}_\beta \circ \mathcal{P}_\alpha \circ \text{Split} \quad (\text{A.3})$$

avec :

$$\alpha = -\frac{1}{2} \quad (\text{A.4})$$

$$\beta = \frac{1}{4} \quad (\text{A.5})$$

$$\gamma = \sqrt{2} \quad (\text{A.6})$$

Le résultat peut être vérifié de deux méthodes différentes. La première méthode, analytique, consiste à expliciter les échantillons  $\{l_x\}_{x \in \mathbb{Z}}$  et  $\{h_x\}_{x \in \mathbb{Z}}$  en fonction des échantillons du signal d'entrée  $\{s_x\}_{x \in \mathbb{Z}}$ . La deuxième méthode, expérimentale, consiste à construire deux réponses en sortie quand l'entrée est un signal de type symbole de Kronecker. Compte tenu du sous-échantillonnage opéré par le schéma de redressement, deux transformations sont nécessaires pour obtenir un résultat complet. Une transformation d'un symbole de Kronecker placé sur un échantillon pair, et une transformation d'un

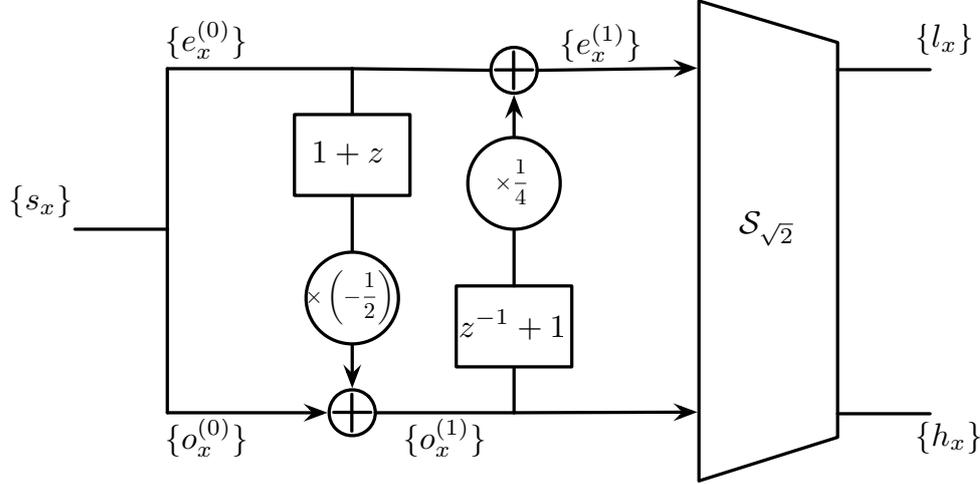


FIGURE A.1 – Schéma de redressement pour les filtres d'ondelettes 5/3.

symbole de Kronecker placé sur un échantillon impair. Les filtres d'analyse déterminés par la méthode analytique vérifient bien les relations (A.1) et (A.2) :

$$l_a = \begin{pmatrix} \alpha\beta & \beta & 1+2\alpha\beta & \beta & \alpha\beta \end{pmatrix} \cdot \gamma \quad (\text{A.7})$$

$$h_a = \begin{pmatrix} \alpha & \underset{\circ}{1} & \alpha \end{pmatrix} \cdot \frac{1}{\gamma} \quad (\text{A.8})$$

Les filtres de prédiction et de mise à jour étant linéaires, la relation (2.13) du schéma de redressement inverse, conjuguée aux relations (2.7) et (2.8) définissent le schéma de reconstruction exacte du signal d'entrée :

$$\mathcal{L}_{5/3}^{-1} = \text{Merge} \circ \mathcal{S}_{\frac{1}{\gamma}}^{-1} \circ \mathcal{P}_{-\alpha} \circ \mathcal{U}_{-\beta} \quad (\text{A.9})$$

De la même façon que les filtres d'analyse ont pu être explicités à partir d'un schéma de redressement, les filtres de synthèse peuvent être formulés ainsi :

$$l_s = \begin{pmatrix} -\alpha & \underset{\circ}{1} & -\alpha \end{pmatrix} \cdot \frac{1}{\gamma} \quad (\text{A.10})$$

$$h_s = \begin{pmatrix} \alpha\beta & -\beta & 1+2\alpha\beta & -\beta & \alpha\beta \end{pmatrix} \cdot \gamma \quad (\text{A.11})$$

## A.2 Filtres 9/7

Les filtres d'analyse de l'ondelettes 9/7 sont approximativement pour le filtre passe-bas et le filtre passe-haut de réponses impulsionnelles respectivement :

$$l_a = \begin{pmatrix} l_{a4} & l_{a3} & l_{a2} & l_{a1} & l_{a0} & l_{a1} & l_{a2} & l_{a3} & l_{a4} \end{pmatrix} \quad (\text{A.12})$$

$$h_a = \begin{pmatrix} h_{a3} & h_{a2} & h_{a1} & h_{a0} & h_{a1} & h_{a2} & h_{a3} \end{pmatrix} \quad (\text{A.13})$$

avec :

$$\begin{array}{rcl} l_{a0} & \approx & 0,8526986788 \\ l_{a1} & \approx & 0,3774028555 \\ l_{a2} & \approx & -0,1106244041 \\ l_{a3} & \approx & -0,02384946501 \\ l_{a4} & \approx & 0,0378284555 \\ \hline h_{a0} & \approx & 0,788485617 \\ h_{a1} & \approx & -0,4180922729 \\ h_{a2} & \approx & -0,04068941765 \\ h_{a3} & \approx & 0,06453888268 \end{array}$$

Un schéma possible de redressement correspondant,  $\mathcal{L}_{9/7}$ , représenté figure A.2, possède deux étages, la relation (2.12) devient dans ce cas :

$$\mathcal{L}_{9/7} = \mathcal{S}_\zeta \circ \mathcal{U}_\delta \circ \mathcal{P}_\gamma \circ \mathcal{U}_\beta \circ \mathcal{P}_\alpha \circ \text{Split} \quad (\text{A.14})$$

avec :

$$\alpha \approx -1.586134342 \quad (\text{A.15})$$

$$\beta \approx -0.05298011854 \quad (\text{A.16})$$

$$\gamma \approx 0.8829110762 \quad (\text{A.17})$$

$$\delta \approx 0.4435068522 \quad (\text{A.18})$$

$$\zeta \approx 1.149604398 \quad (\text{A.19})$$

Explicités en fonction des coefficients du schéma de redressement, les filtres d'analyse sont donc de réponses impulsionnelles :

$$l_a = \begin{pmatrix} l'_{a4} & l'_{a3} & l'_{a2} & l'_{a1} & l'_{a0} & l'_{a1} & l'_{a2} & l'_{a3} & l'_{a4} \end{pmatrix} \cdot \zeta \quad (\text{A.20})$$

$$h_a = \begin{pmatrix} h'_{a3} & h'_{a2} & h'_{a1} & h'_{a0} & h'_{a1} & h'_{a2} & h'_{a3} \end{pmatrix} \cdot \frac{1}{\zeta} \quad (\text{A.21})$$

$$\begin{array}{rcl} l'_{a0} & = & (1 + 2\alpha\beta) + 2\delta(\alpha + \gamma + 3\alpha\beta\gamma) \\ l'_{a1} & = & \beta + \delta(1 + 3\beta\gamma) \\ l'_{a2} & = & \alpha\beta + \delta(\alpha + \gamma + 4\alpha\beta\gamma) \\ l'_{a3} & = & \beta\gamma\delta \\ l'_{a4} & = & \alpha\beta\gamma\delta \\ \hline h'_{a0} & = & 1 + 2\beta\gamma \\ h'_{a1} & = & \alpha + \gamma + 3\alpha\beta\gamma \\ h'_{a2} & = & \beta\gamma \\ h'_{a3} & = & \alpha\beta\gamma \end{array}$$

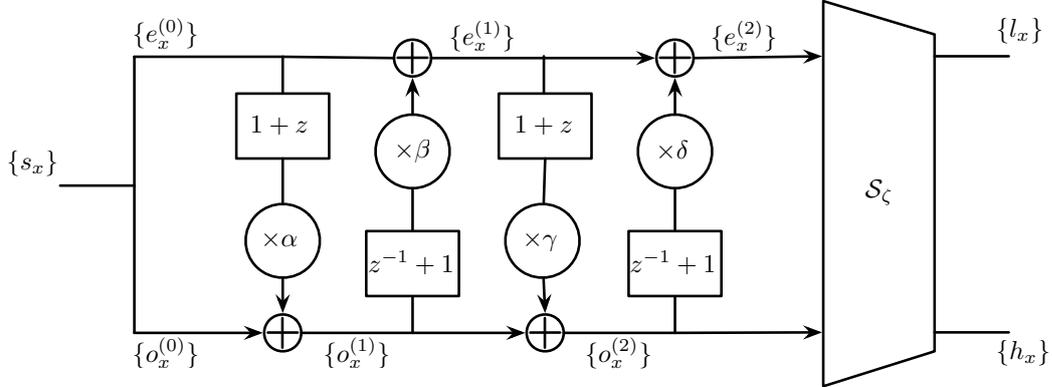


FIGURE A.2 – Schéma de redressement pour les filtres d'ondelettes 9/7.

et les filtres de synthèse :

$$l_s = \begin{pmatrix} l'_{s3} & l'_{s2} & l'_{s1} & l'_{s0} & l'_{s1} & l'_{s2} & l'_{s3} \end{pmatrix} \cdot \frac{1}{\zeta} \quad (\text{A.22})$$

$$h_s = \begin{pmatrix} h'_{s4} & h'_{s3} & h'_{s2} & h'_{s1} & h'_{s0} & h'_{s1} & h'_{s2} & h'_{s3} & h'_{s4} \end{pmatrix} \cdot \zeta \quad (\text{A.23})$$

$$\begin{aligned} l'_{s0} &= 1 + 2\beta\gamma \\ l'_{s1} &= -(\alpha + \gamma + 3\alpha\beta\gamma) \\ l'_{s2} &= \beta\gamma \\ l'_{s3} &= -\alpha\beta\gamma \\ \hline h'_{s0} &= 1 + 2(\alpha\beta + \alpha\delta + \gamma\delta) + 6\alpha\beta\gamma\delta \\ h'_{s1} &= -(\beta + \delta + 3\beta\gamma\delta) \\ h'_{s2} &= \alpha\beta + \alpha\delta + \gamma\delta + 4\alpha\beta\gamma\delta \\ h'_{s3} &= -\beta\gamma\delta \\ h'_{s4} &= \alpha\beta\gamma\delta \end{aligned}$$

## Annexe B

# Paramètres de $\theta$ -redressement orienté et filtres d'ondelettes orientées associés

### B.1 Filtres 5/3 pour $0 \leq \theta \leq 45^\circ$

Filtres d'analyse, en posant  $t = \tan \theta$  :

$$l_a = \begin{pmatrix} l_{2,2} & & & & & \\ l_{2,1} & l_{1,1} & l_{0,1} & & & \\ l_{2,0} & l_{1,0} & l_{0,0} & l_{1,0} & l_{2,0} & \\ & & l_{0,1} & l_{1,1} & l_{2,1} & \\ & & & & & l_{2,2} \end{pmatrix} \cdot \gamma \quad (\text{B.1})$$

avec :

$$\begin{aligned} l_{0,0} &= 1 + 2\alpha\beta[(1-t)^2 + t^2] \\ l_{0,1} &= 2\alpha\beta(1-t)t \\ \hline l_{1,0} &= \beta(1-t) \\ l_{1,1} &= \beta t \\ \hline l_{2,0} &= \alpha\beta(1-t)^2 \\ l_{2,1} &= 2\alpha\beta(1-t)t \\ l_{2,2} &= \alpha\beta t^2 \end{aligned}$$

$$h_a = \begin{pmatrix} \alpha t & & & \\ \alpha(1-t) & 1 & \alpha(1-t) & \\ & & & \alpha t \end{pmatrix} \cdot \frac{1}{\gamma} \quad (\text{B.2})$$



avec :

$$\begin{array}{rcl}
h_{0,0} & = & 1 + 2\beta\gamma(u^2 + v^2) \\
h_{0,1} & = & 2\beta\gamma uv \\
\hline
h_{1,-1} & = & 3\alpha\beta\gamma u^2 v \\
h_{1,0} & = & (\alpha + \gamma)u + 3\alpha\beta\gamma u^3 + 6\alpha\beta\gamma uv^2 \\
h_{1,1} & = & (\alpha + \gamma)v + 3\alpha\beta\gamma v^3 + 6\alpha\beta\gamma u^2 v \\
h_{1,2} & = & 3\alpha\beta\gamma uv^2 \\
\hline
h_{2,0} & = & \beta\gamma u^2 \\
h_{2,1} & = & 2\beta\gamma uv \\
h_{2,2} & = & \beta\gamma v^2 \\
\hline
h_{3,0} & = & \alpha\beta\gamma u^3 \\
h_{3,1} & = & 3\alpha\beta\gamma u^2 v \\
h_{3,2} & = & 3\alpha\beta\gamma uv^2 \\
h_{3,3} & = & \alpha\beta\gamma v^3
\end{array}$$



## Annexe C

# Base d'images

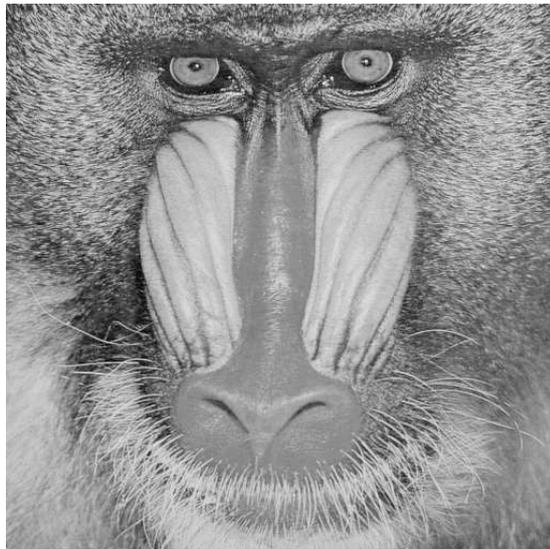


FIGURE C.1 – baboon

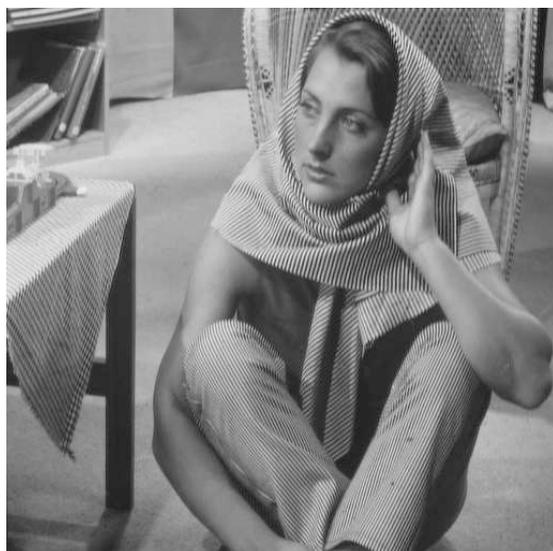


FIGURE C.2 – barbara



FIGURE C.3 – boats



FIGURE C.4 – farm



FIGURE C.5 – fruits



FIGURE C.6 – goldhill

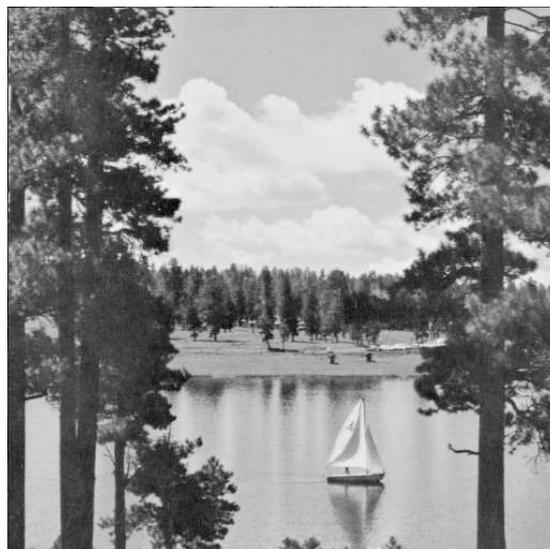


FIGURE C.7 – lake



FIGURE C.8 – lena



FIGURE C.9 – monarch



FIGURE C.10 – parrots



FIGURE C.11 – peppers



FIGURE C.12 – plane

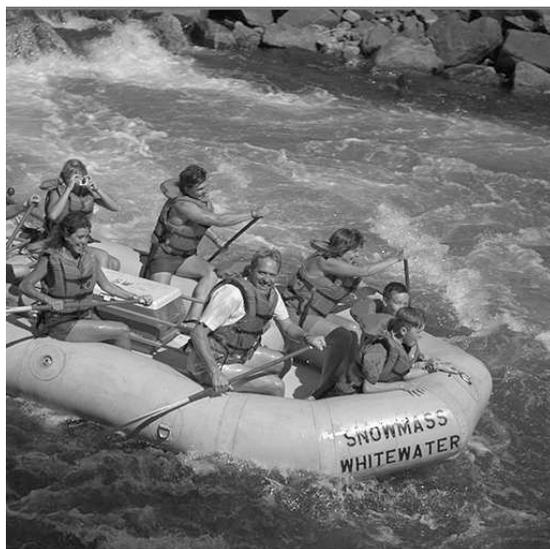


FIGURE C.13 – raft



FIGURE C.14 – train



FIGURE C.15 – yosemite

## Annexe D

### Exemples de cartes de contours extraits à trois niveaux de résolution

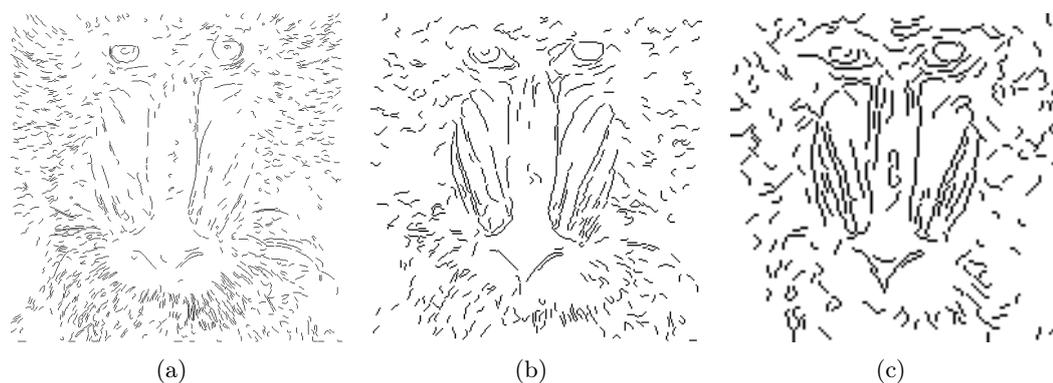


FIGURE D.1 – baboon

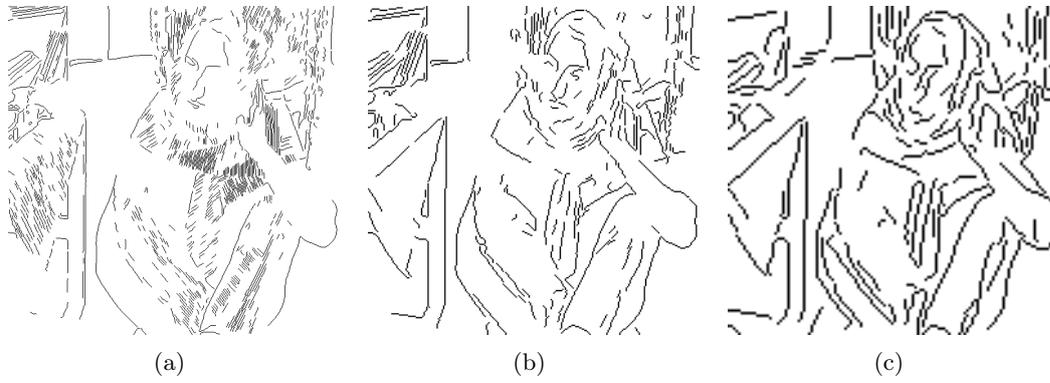


FIGURE D.2 – barbara

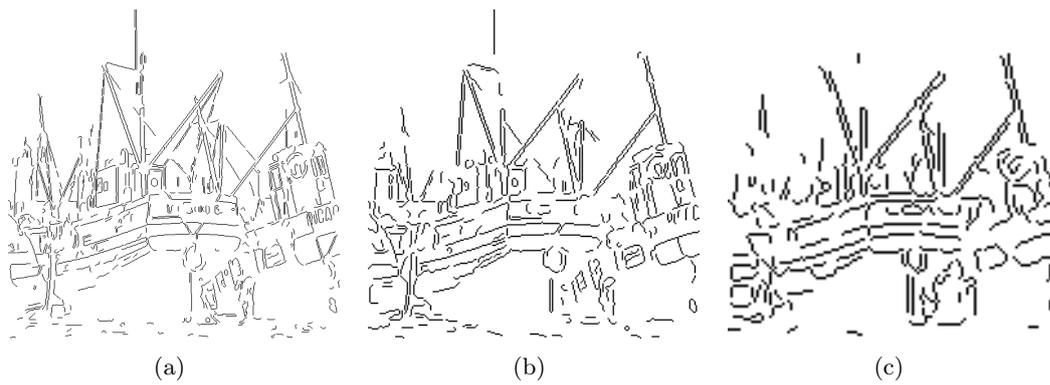


FIGURE D.3 – boats

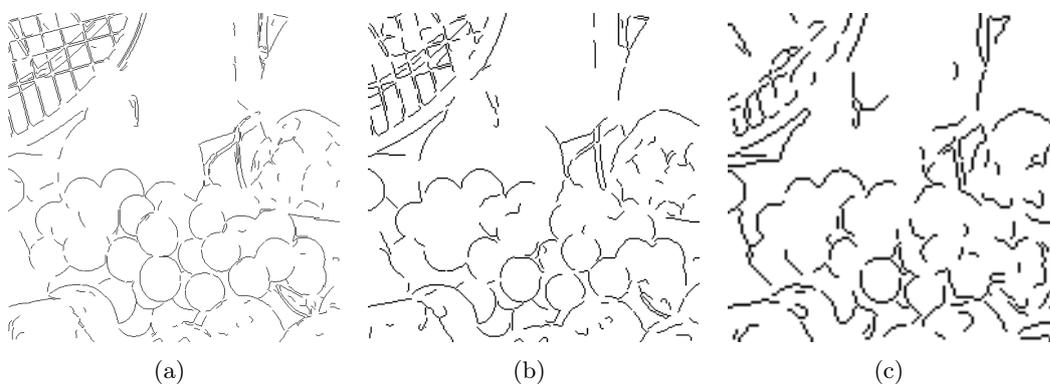


FIGURE D.4 – fruits

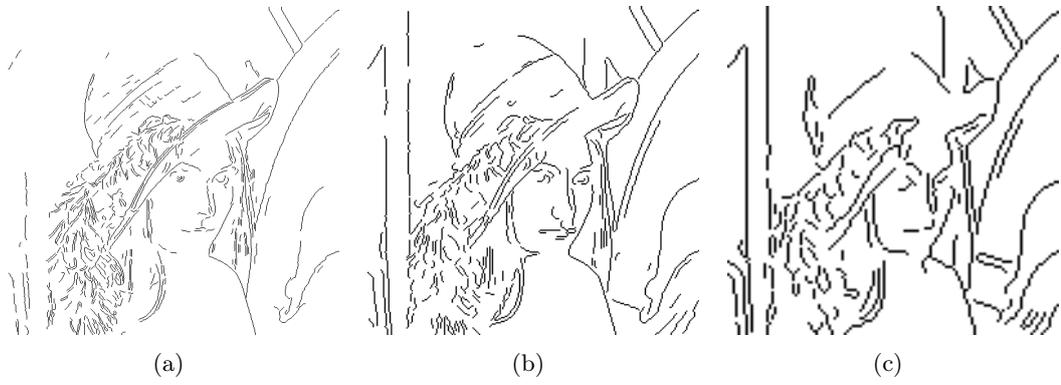


FIGURE D.5 – lena

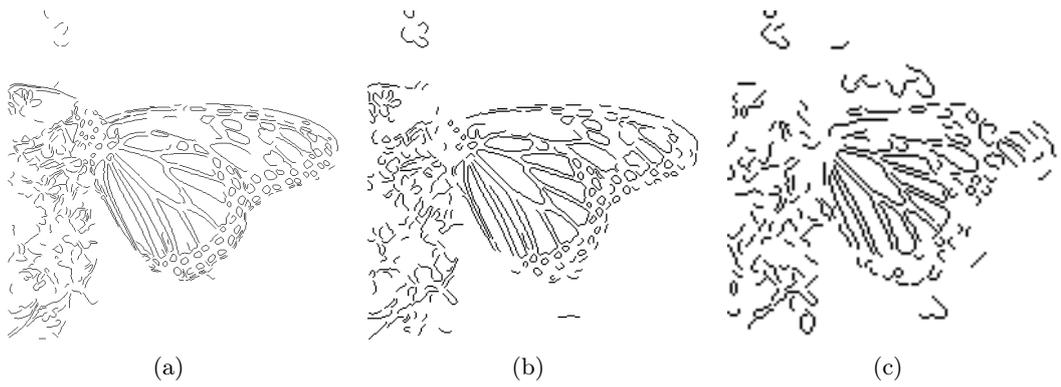


FIGURE D.6 – monarch

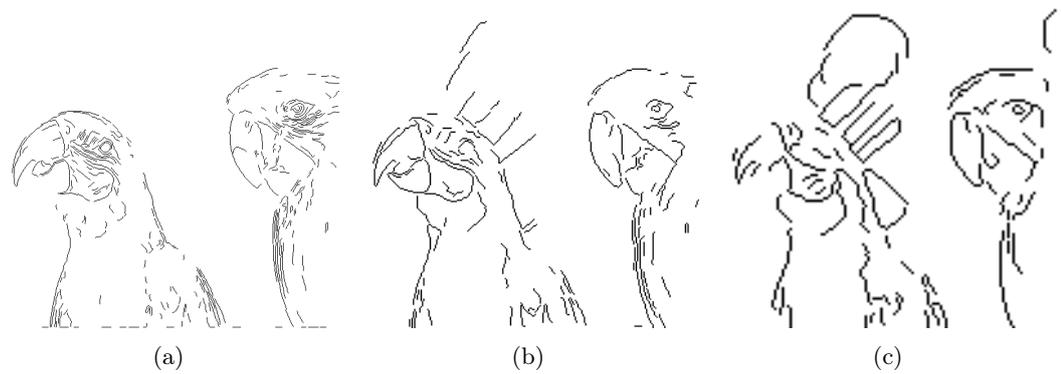


FIGURE D.7 – parrots

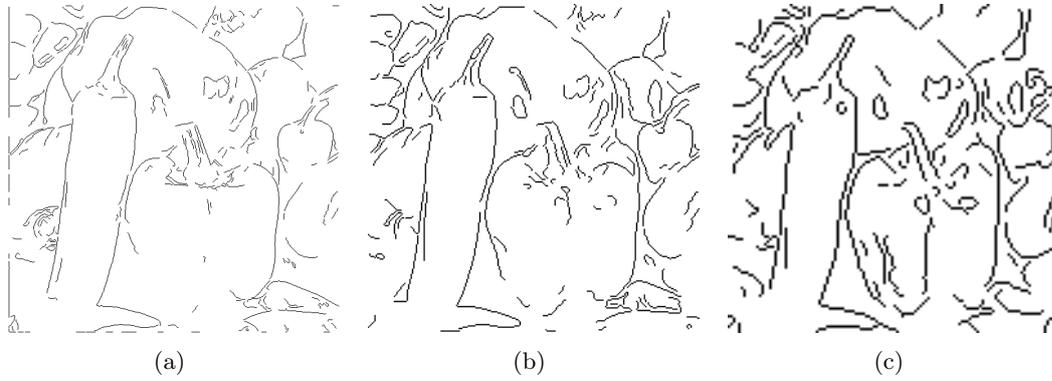


FIGURE D.8 – peppers

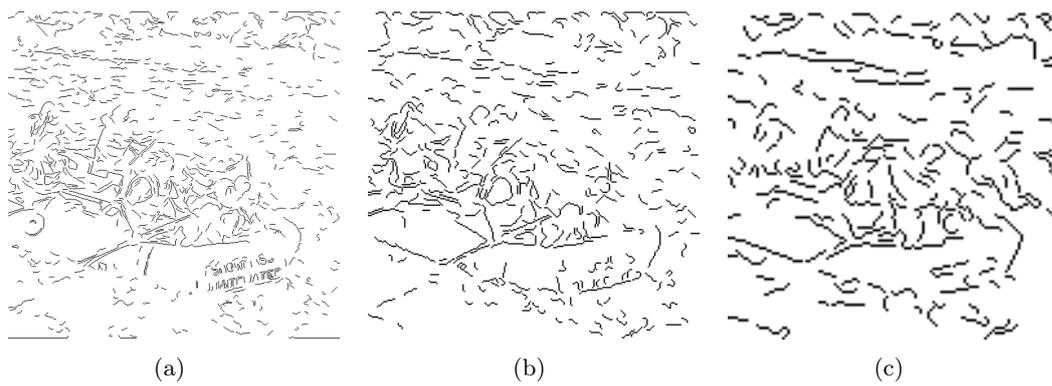


FIGURE D.9 – raft

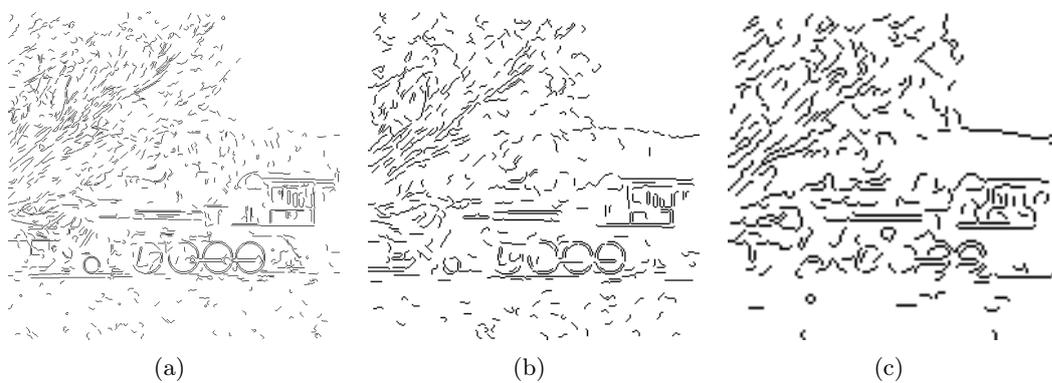


FIGURE D.10 – train

## Annexe E

# Exemples de classification structurelle et d'orientations de filtrage estimées

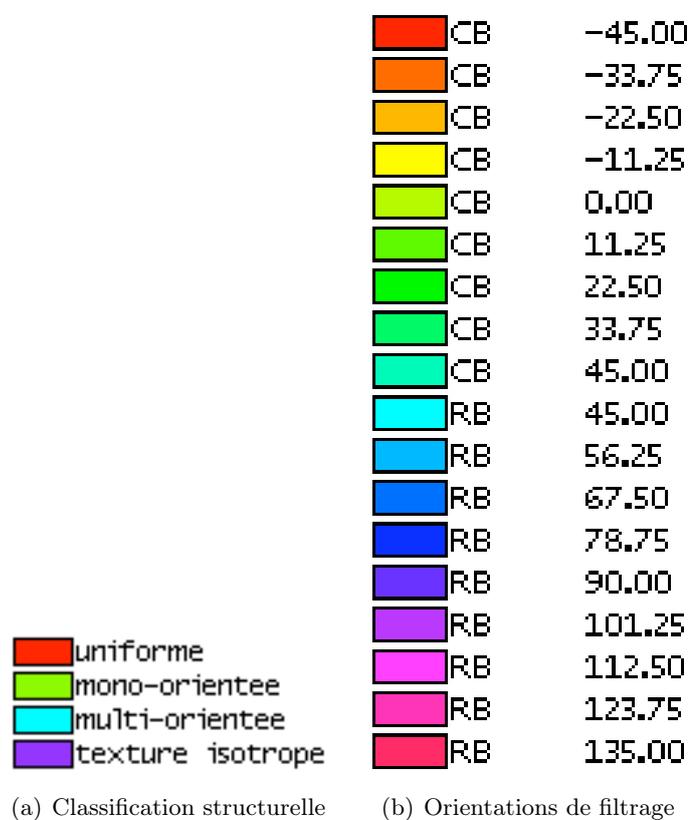


FIGURE E.1 – Légende des couleurs.

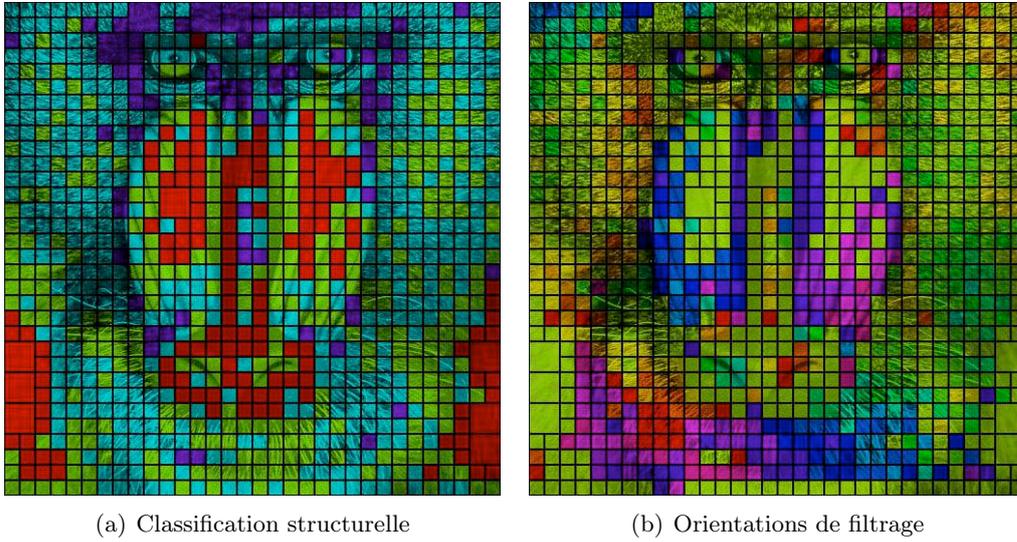
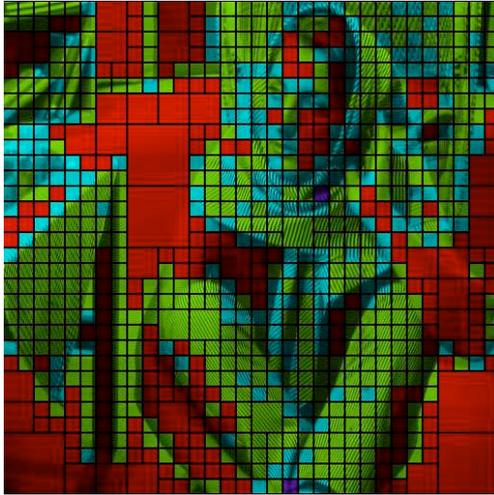
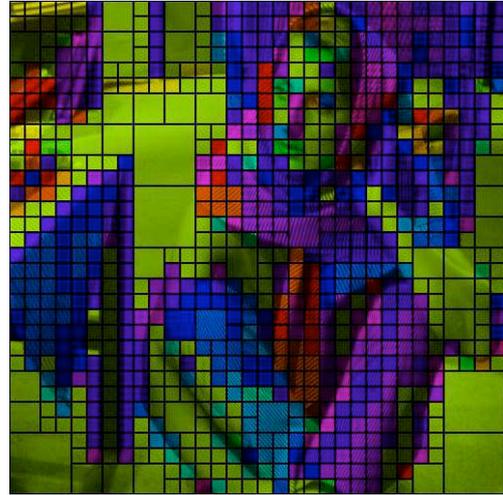


FIGURE E.2 – baboon

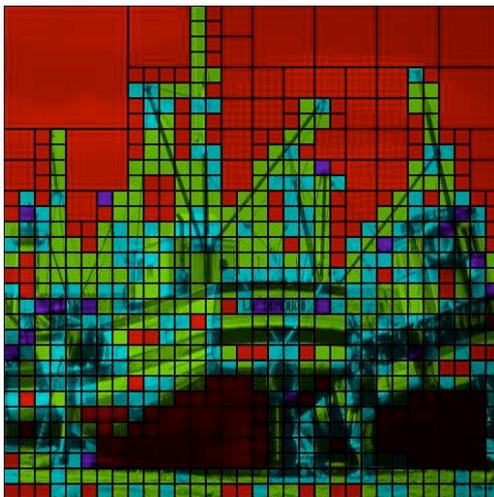


(a) Classification structurelle

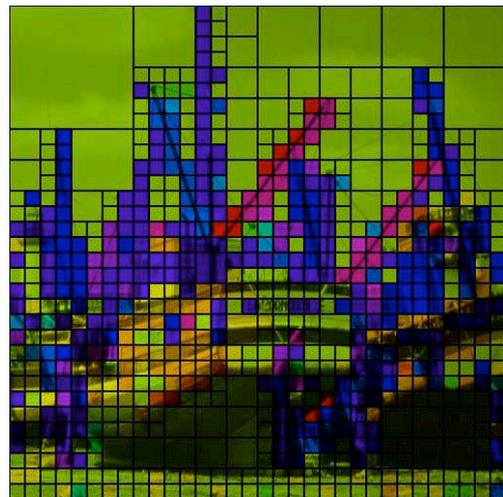


(b) Orientations de filtrage

FIGURE E.3 – barbara



(a) Classification structurelle



(b) Orientations de filtrage

FIGURE E.4 – boats

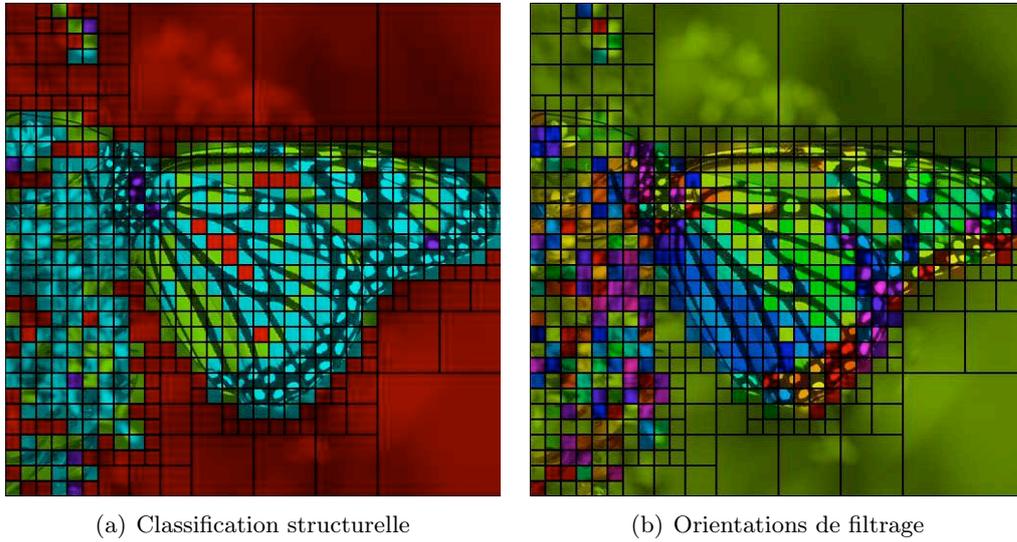


FIGURE E.5 – monarch

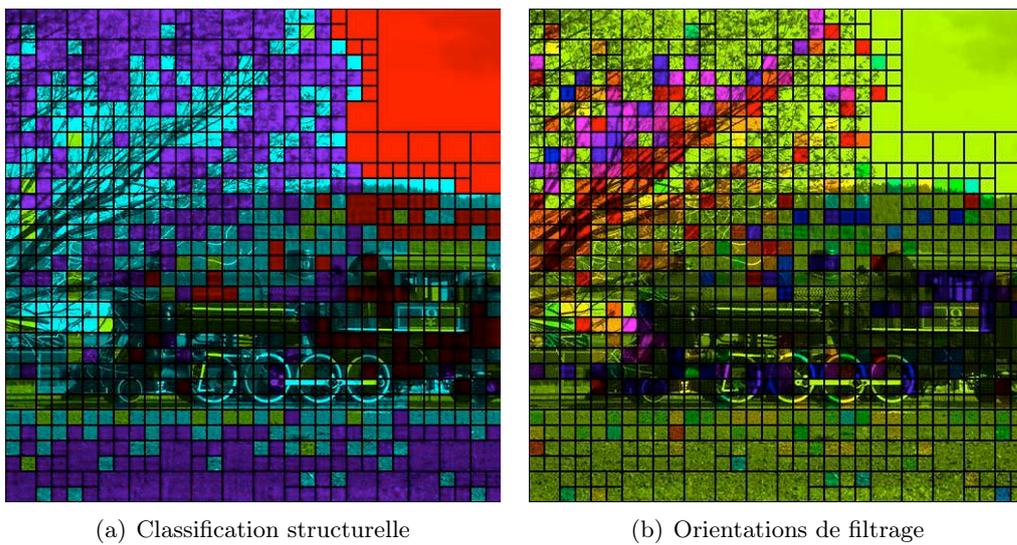


FIGURE E.6 – train

# Bibliographie

- [Ant03] Marc ANTONINI : Compression des images et des vidéos numériques — dix années de recherches au cnrs. Habilitation à Diriger des Recherches, 2003.
- [BCK02] Eric BODDEN, Malte CLASEN et Joachim KNEIS : Arithmetic Coding in revealed. *In Proseminar Datenkompression 2001*. RWTH Aachen University, 2002.
- [Ber87] Jean-François BERTRAND : *Codage d'images par segmentation et séparation en deux composantes ; Optimisation des performances d'un système de codage d'image monochromes*. Thèse de doctorat, INSA, mars 1987.
- [BS92] R. H. BAMBERGER et M. J. T. SMITH : A filter bank for the directional decomposition of images : theory and design. *IEEE Transactions on Signal Processing*, 40(4):882–893, avril 1992.
- [CA04] Philippe CARRE et Eric ANDRES : Discrete analytical ridgelet transform. *Signal Processing*, 84(11):2165–2173, november 2004.
- [Can98] Emmanuel J. CANDÈS : *Ridgelets : Theory and Applications*. Thèse de doctorat, Stanford, août 1998.
- [CD99] Emmanuel J. CANDÈS et David L. DONOHO : Ridgelets : a key to higher-dimensional intermittency ? *Philosophical Transactions Royal Society London A*, (357):2495–2509, 1999.
- [CD02] Emmanuel J. CANDÈS et David L. DONOHO : New tight frames of curvelets and optimal representations of objects with smooth singularities. Rapport technique, Stanford University, 2002.
- [CG05] Vivien CHAPPELIER et Christine GUILLEMOT : Oriented wavelet transform for image compression and denoising. *In IEEE Transactions On Image Processing*, 2005.
- [CG07] Chuo-Ling CHANG et Bernd GIROD : Direction-adaptive discrete wavelet transform for image compression. *IEEE Transactions on Image Processing*, 16(5), mai 2007.
- [Cha05] Vivien CHAPPELIER : *Codage progressif d'images par ondelettes orientées*. Thèse de doctorat, Université de Rennes 1, 2005.
- [CLCB08] Mathieu CARNEC, Patrick LE CALLET et Dominique BARBA : Objective quality assessment of color images based on a generic perceptual reduced

- reference. *Signal Processing : Image Communication*, 23(4):239–256, avril 2008.
- [CS95] Yuk-Hee CHAN et Wan-Chi SIU : Highly efficient coding schemes for contour line drawings. *In International Conference on Image Processing*, 1995.
- [DM83] Dan E. DUDGEON et Russell M. MERSEREAU : *Multidimensional Digital Signal Processing*. Prentice Hall, 1983.
- [DS98] Ingrid DAUBECHIES et Wim SWELDENS : Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):247–269, 1998.
- [DV03] Minh N. DO et Martin VETTERLI : The finite ridgelet transform for image representation. *IEEE Transactions On Image Processing*, 12(1), january 2003.
- [DV05] Minh N. DO et Martin VETTERLI : The contourlet transform : An efficient directional multiresolution image representation. *IEEE Transactions On Image Processing*, 14(12):2091–2106, décembre 2005.
- [DWW<sup>+</sup>07] Wenpeng DING, Feng WU, Xiaolin WU, Shipeng LI et Houqiang LI : Adaptive directional lifting-based wavelet transform for image coding. *IEEE Transactions on Image Processing*, 16(2):416–427, février 2007.
- [Gou02] Annabelle GOUZE : *Schéma de lifting quinconce pour la compression d'images*. Thèse de doctorat, Université de Nice - Sophia Antipolis, décembre 2002.
- [Gué01] Jeanpierre GUÉDON : L'histoire de la transformée mojette (la forme et la projection). Habilitation à Diriger des Recherches, 2001.
- [HPPP06] Henk HEIJMANS, Gemma PIELLA et Béatrice PESQUET-POPESCU : Adaptive wavelets for image compression using update lifting : Quantisation and error analysis. *International Journal of Wavelets, Multiresolution and Information Processing*, 4(1):41–65, 2006.
- [JRB07a] Guillaume JEANNIC, Vincent RICORDEL et Dominique BARBA : The edge driven oriented wavelet transform : An anisotropic multidirectional representation with oriented lifting scheme. *In Visual Communications and Image Processing*, San Jose, USA, janvier 2007.
- [JRB07b] Guillaume JEANNIC, Vincent RICORDEL et Dominique BARBA : A multi-resolution approach for the coding of edges of still images using adaptive arithmetic coding. *In PCS*, Lisbonne, Portugal, novembre 2007.
- [LD91] Cheng-Chang LU et James G. DUNHAM : Highly efficient coding schemes for contour lines based on chain code representations. *IEEE Transactions On Communications*, 39(10):1511–1514, octobre 1991.
- [LD03] Yue LU et Minh N. DO : Crisp-contourlets : a critically sampled directional multiresolution image representation. *In Proc. of SPIE conference on Wavelet Applications in Signal and Image Processing X*, San Diego, USA, august 2003.

- [LG08] Benjamin LE GUEN : *Adaptation du contenu spatio-temporel des images pour un codage par ondelettes*. Thèse de doctorat, Université de Rennes 1, février 2008.
- [LP02] Erwan LE PENNEC : *Bandelettes et représentation géométrique des images*. Thèse de doctorat, École Polytechnique, décembre 2002.
- [LPM03] Erwan LE PENNEC et Stéphane MALLAT : Bandelettes et représentation géométrique des images. *In GRETSI*, Paris, september 2003.
- [LPM05] Erwan LE PENNEC et Stéphane MALLAT : Sparse geometric image representations with bandelets. *IEEE Transactions On Image Processing*, avril 2005.
- [Mal99] Stéphane MALLAT : *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- [MLB<sup>+</sup>02] Michael W. MARCELLIN, Margaret A. LEPLEY, Ali BILGIN, Thomas J. FLOHR, Troy T. CHINEN et James H. KASNER : An overview of quantization in JPEG 2000. *Signal Processing : Image Communications ISSN 0923-5965*, 17(1):73–84, 2002.
- [NRK03] Marcus J. NADENAU, Julien REICHEL et Murat KUNT : Wavelet-based color image compression : exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing*, 12(1):58–70, 2003.
- [PL97] Stéphane PATEUX et Claude LABIT : Efficient way of coding arbitrary geometric partition for region-based video sequence coding. *In Visual Communications and Image Processing*, volume 3024, pages 742–753, janvier 1997.
- [PM05] Gabriel PEYRÉ et Stéphane MALLAT : Surface compression with geometric bandelets. *ACM Transactions on Graphics*, 24(3), juillet 2005.
- [PPPH02] Gemma PIELLA, Béatrice PESQUET-POPESCU et Henk HEIJMANS : Adaptive update lifting with a decision rule based on derivative filters. *IEEE Signal Processing Letters*, pages 329–332, octobre 2002.
- [PPPH05] Gemma PIELLA, Béatrice PESQUET-POPESCU et Henk HEIJMANS : Gradient-driven update lifting for adaptive wavelets. *Signal Processing : Image Communication*, 20(9-10):813–831, août 2005.
- [Rev91] Jean-Pierre REVEILLÈS : *Géométrie discrète, calcul en nombres entiers et algorithmique*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, France, décembre 1991.
- [SBK05] Ivan W. SELESNICK, Richard G. BARANIUK et Nick G. KINGSBURY : The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, 22(6):123–151, 2005.
- [SCD00] 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.

- [SP96] A. SAID et W. A. PEARLMAN : A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6:243–250, juin 1996.
- [Tau98] David TAUBMAN : Ebcot : Embedded block coding with optimized truncation. *IEEE transactions on image processing ISSN 1057-7149*, 9(7):1158–1170, 1998.
- [TM01] David TAUBMAN et Michael W. MARCELLIN : JPEG 2000 image compression fundamentals, standards and practice. *Kluwer International Series in Engineering and Computer Science, Secs 642, ISBN-10 : 079237519X, ISBN-13 : 978-0792375197*, 2001.
- [VBLVD06] 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.
- [Vil05] Emmanuel VILLÉGER : *Constance de largeur et désocclusion dans les images digitales*. Thèse de doctorat, Université de Nice - Sophia Antipolis, décembre 2005.
- [WBSS04] Z. WANG, A. C. BOVIK, H. R. SHEIKH et E. P. SIMONCELLI : Image quality assessment : from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, avril 2004.
- [WZS06] Demin WANG, Liang ZHANG et Filippo SPERENZA : Curved wavelet transform for image coding. *IEEE Transactions On Image Processing*, 15(8):2413–2421, août 2006.
- [ZDL00] Wenjun ZENG, Scott DALY et Shawmin LEI : Visual optimization tools in JPEG 2000. *In International Conference on Image Processing*, septembre 2000.
- [ZDL02] Wenjun ZENG, Scott DALY et Shawmin LEI : An overview of the visual optimization tools in jpeg 2000. *Signal rocessing : Image Coommunication*, 17(1):85–104, janvier 2002.



## Résumé

Dans un contexte de compression d'images fixes, de nouvelles représentations implantées par des techniques de redressement (lifting) orienté ont été imaginées. Pour ce travail de thèse, trois variantes de ces transformées sont évaluées comparativement selon un critère de minimisation de l'énergie des hautes fréquences reconstruites et selon la direction de régularité locale. Cette direction de régularité est estimée par des approches basées extraction de contours ou minimisation d'une fonction de coût. Les deux approches proposées permettent de plus une classification de l'activité directionnelle des blocs de l'image. Ces représentations des caractéristiques géométriques de l'image utilisées pour guider le filtrage en ondelettes orientées, est codé pour être transmis au décodeur. Les coûts de ces différentes représentations sont comparés. Finalement une quantification adaptée est proposée pour tenir compte à la fois des différentes classes structurelles et de l'anisotropie de la représentation.

## Abstract

In the context of still image compression, some new representation implemented with oriented lifting schemes has been conceived. For this thesis, three implementations of these transforms have been benchmarked upon a criteria considering the minimization of the high frequencies along the local direction of regularity. This orientation is estimated via edge extraction or minimization of a cost function. Both approaches lead to the classification of the directional activity of the image blocks. Those representations of the geometrical features that drive the oriented filtering process are coded in order to be transmitted to the decoder. The coding cost of these representations is compared. Finally an adaptive quantization is proposed to take into account the structural classes and the anisotropy of the representation.