

THESE

Pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ DE POITIERS
UFR des sciences fondamentales et appliquées
Pôle poitevin de recherche pour l'ingénieur en mécanique, matériaux et énergétique - PPRIMME
(Diplôme National - Arrêté du 7 août 2006)

École doctorale : Sciences et ingénierie pour l'information, mathématiques - S2IM
Secteur de recherche : Automatique et traitement du signal

Présentée par :
Raphaël Soulard

Ondelettes analytiques et monogènes pour la représentation des images couleur

Directeur(s) de Thèse :
Christine Fernandez-Maloigne, Philippe Carré

Soutenue le 19 novembre 2012 devant le jury

Jury :

Président	Laure Blanc-Féraud	Directrice de recherche CNRS
Rapporteur	Frédéric Truchetet	Professeur - Université de Dijon
Rapporteur	Jean-Christophe Pesquet	Professeur - Université de Marne la vallée
Membre	Christine Fernandez-Maloigne	Professeur - Université de Poitiers
Membre	Philippe Carré	Professeur - Université de Poitiers
Membre	Laurent Duval	Ingénieur de recherche - IFP - Energies Nouvelles

Pour citer cette thèse :

Raphaël Soulard. *Ondelettes analytiques et monogènes pour la représentation des images couleur* [En ligne]. Thèse Automatique et traitement du signal . Poitiers : Université de Poitiers, 2012. Disponible sur Internet
<<http://theses.univ-poitiers.fr>>

THÈSE

pour l'obtention du Grade de
DOCTEUR DE L'UNIVERSITÉ DE POITIERS
(Faculté des Sciences Fondamentales et Appliquées)
(Diplôme National - Arrêté du 7 août 2006)

École Doctorale : Sciences et Ingénierie pour l'Information
Secteur de Recherche : Automatique et Traitement du Signal

Présentée par :
Raphaël SOULARD

ONDELETTES ANALYTIQUES ET MONOGÈNES POUR LA REPRÉSENTATION DES IMAGES COULEUR

Directeur de Thèse : Philippe CARRÉ

Soutenue le 19 novembre 2012 devant la Commission d'Examen composée de :

Jean-Christophe PESQUET, Professeur	Rapporteur
Frédéric TRUCHETET, Professeur	Rapporteur
Laure BLANC-FÉRAUD, Directrice de recherche CNRS	Examinatrice
Laurent DUVAL, Ingénieur de recherche	Examineur
Christine FERNANDEZ-MALOIGNE, Professeur	Examinatrice
Philippe CARRÉ, Professeur	Examineur

À Elsa

PRÉAMBULE

C E mémoire de thèse présente le travail que j'ai mené ces trois dernières années au laboratoire XLIM-SIC de l'Université de Poitiers, avec le financement de la région Poitou-Charentes, en association avec le projet ANR VERSO-CAIMAN, et sous la direction de Philippe Carré.

Je tiens avant tout à remercier Philippe de m'avoir attiré vers le traitement du signal et des images, qui est une discipline passionnante et pleine d'avenir. Merci également de m'avoir encadré avec confiance et soutien.

Je remercie également l'ensemble des personnes que j'ai croisées au laboratoire, en particulier Christine Fernandez-Maloigne, l'équipe ICONES, le personnel administratif, les congénères doctorants. . .

Je remercie infiniment les membres de ma famille de leur curiosité envers ce travail et pour leurs encouragements, en particulier ma sœur Sarah, ma mère pour son amour sans limite et son admiration, et mon père qui me donne parfois en secret quelques conseils depuis là-haut.

Enfin, je remercie Elsa du fond du cœur de m'avoir encouragé, soutenu, supporté, attendu, aidé, consolé, rassuré, préparé des tisanes aux pouvoirs magiques, et bien des choses encore, tout en rédigeant son propre mémoire. Merci d'avoir partagé l'aventure avec moi.

Poitiers, le 3 octobre 2012.

TABLE DES MATIÈRES

Préambule	5
Table des matières	7
Glossaire	9
Introduction	11
1 Amplitude, phase et fréquence	15
1.1 Outils de base 1D	16
1.1.1 Domaine fréquentiel	16
1.1.2 Domaine temporel	17
1.2 Analyse multibande 1D	19
1.2.1 Représentations continues : Gabor et Morlet	19
1.2.2 Bancs de filtres orthogonaux à coefficients réels	21
1.2.3 Extensions complexes	23
1.3 Vers des signaux 2D	25
1.3.1 La transformée de Fourier 2D	26
1.3.2 Ondelettes de Gabor directionnelles	28
1.3.3 Ondelettes orthogonales séparables	29
1.3.4 Dual-tree 2D	30
1.4 Résumé et discussion	33
2 Approches quaternioniques	35
2.1 Image et quaternions	35
2.1.1 Algèbre des quaternions	35
2.1.2 Outils quaternioniques pour les images couleur	37
2.2 Le formalisme de T. Bülow	37
2.2.1 Transformée de Fourier quaternionique (QFT)	38
2.2.2 Le signal quaternionique analytique	41
2.2.3 Ondelettes quaternioniques	42
2.3 Utilisation de la QWT	46
2.3.1 Les angles φ et θ et le flot optique	46
2.3.2 L'angle ψ	48
2.4 Etude de la parcimonie de la QWT	49
2.4.1 Codage de l'amplitude	50
2.4.2 Codage de la phase quaternionique	50
2.4.3 Schéma complet de quantification	55
2.5 Mise à l'épreuve en classification de textures	58
2.5.1 Bases d'images de textures	59
2.5.2 Protocole expérimental	59
2.5.3 Extraction des descripteurs	60
2.5.4 Résultats expérimentaux	62
2.6 Conclusion sur les approches quaternioniques	67

3	Approches isotropes et extension couleur	69
3.1	Analyse locale des structures	70
3.1.1	Tenseur de structure basé gradient	70
3.1.2	Approche « signal » du tenseur	73
3.1.3	Filtres en quadrature orientables	76
3.2	Un signal analytique isotrope	79
3.2.1	Signal monogène	79
3.2.2	Dépasser les nombres complexes	83
3.2.3	Discussion	88
3.3	Ondelettes monogènes peu redondantes	88
3.3.1	Ondelettes isotropes	89
3.3.2	Ondelettes monogènes de M. Unser <i>et al.</i>	92
3.4	Extension couleur par les algèbres géométriques	94
3.4.1	Le signal monogène couleur de G. Demarcq	94
3.4.2	Décomposition en ondelettes	96
3.4.3	La reconstruction	98
3.5	Mise à l'épreuve en débruitage couleur	99
3.5.1	Étude des coefficients liés au bruit	99
3.5.2	Seuillage des coefficients	103
3.5.3	Résultats et analyse	103
3.6	Bilan et conclusion du chapitre	109
4	Une nouvelle représentation monogène couleur	111
4.1	Variations locales d'une image couleur	113
4.2	Un nouveau signal monogène couleur	115
4.2.1	Transformée de Riesz couleur	115
4.2.2	Phase de signaux vectoriels	117
4.2.3	Signal monogène couleur	121
4.3	Ondelettes monogènes couleur	124
4.3.1	Définition	124
4.3.2	Reconstruction	128
4.4	Mise en pratique	129
4.4.1	Représentation parcimonieuse des images couleur	129
4.4.2	Analyse invariante d'images couleur	133
4.5	Bilan	137
4.6	Une nouvelle piste de schéma numérique	139
4.6.1	Le domaine de Radon	140
4.6.2	Algorithme discret	141
4.6.3	Discussion	142
	Conclusion	143
	Publications associées	147
	A Résultats de classification de texture	149
	Liste des figures	153
	Bibliographie	157
	Index	165
	Résumé - Abstract	167

GLOSSAIRE

$\mathbb{R}, \mathbb{R}^+$	Ensembles des nombres réels et des réels positifs
$\mathbb{Z}$	Ensemble des nombres entiers
$\mathbb{C}$	Ensemble des nombres complexes
j	nombre imaginaire complexe
$\Re\{z\}, \Im\{z\}$	parties réelle et imaginaire d'un nombre complexe z
z^*	nombre complexe conjugué
$\mathbb{H}$	Ensemble des quaternions
j_1, j_2, j_3	nombres imaginaires quaternioniques
$\mathcal{G}_n$	Algèbre géométrique construite sur $\mathbb{R}^n$
$e_i, e_{ij}, e_{ijk} \dots$	éléments de base de l'algèbre géométrique (vecteurs, bivecteurs, trivecteurs <i>etc.</i>)
$\begin{bmatrix} a & b & c \end{bmatrix}^T$	notation vectorielle
$\mathbf{x}, \boldsymbol{\omega}, \mathbf{k}$	coordonnées 2D ($\mathbf{x} = [x_1 \ x_2]^t$ <i>etc.</i>)
$L^2(\mathbb{R})$	Espace de Hilbert des signaux complexes 1D d'énergie finie
$\langle h, g \rangle$	produit scalaire de h et g
$h * g$	convolution entre h et g
$s(t)$	signal continu
$s[k]$	séquence discrète
$v.p. \int s(t) dt$	Valeur principale de Cauchy de l'intégrale
$\mathcal{F}, \mathcal{F}^{-1}$	transformation de Fourier et transformée de Fourier inverse
$\hat{s}(\omega) = \mathcal{F}\{s(t)\}$	s « chapeau », transformée de Fourier de s
$\mathcal{H}$	Transformée de Hilbert
$\mathcal{H}_{\theta}^{\text{part}}$	Transformée de Hilbert partielle dans la direction θ
$\mathcal{R}$	Transformée de Riesz
$\mathcal{R}_0, \mathcal{R}_{\frac{\pi}{2}}$	composantes cartésiennes de la transformée de Riesz
∇	Opérateur de gradient
$\nabla_{\mathbb{C}}$	Opérateur de gradient complexe
Δ	Laplacien
$(-\Delta)^\alpha$	Laplacien fractionnel
AM/FM	Modulation d'amplitude et de fréquence
STFT	Transformée de Fourier fenêtrée (Short-time Fourier transform)
QFT	Transformée de Fourier quaternionique
AMR	Analyse multirésolution
DWT	Transformée en ondelettes séparable
QWT	Transformée en ondelettes quaternionique
MWT	Transformée en ondelettes monogènes
RGB	Espace colorimétrique rouge-vert-bleu (Red Green Blue)

INTRODUCTION

La couleur est par excellence la partie de l'art qui détient le don magique. Alors que le sujet, la forme, la ligne s'adressent d'abord à la pensée, la couleur n'a aucun sens pour l'intelligence, mais elle a tous les pouvoirs sur la sensibilité.

— Eugène DELACROIX, XIX^e siècle

REPRÉSENTER des éléments visuels avec des nombres, telle est la tâche qu'accomplissent aujourd'hui des centaines de millions d'ordinateurs et appareils photo numériques. La représentation numérique basique d'une image est celle de l'écran, qui consiste en une suite de *pixels*, codés par des nombres, et organisés dans un quadrillage rectangulaire très fin. Lorsque l'on veut analyser ou modifier une image automatiquement, ce « tableau de nombres » ne fournit pas toujours d'informations suffisamment explicites. La première étape des systèmes d'analyse et de traitement d'image consiste alors souvent à *transformer* ce tableau, de façon à obtenir des données plus pertinentes.

Depuis plus de vingt ans, les *transformations en ondelettes* permettent de séparer les informations liées aux détails de l'image en fonction de leur orientation et de leur taille. Certains objets fondamentaux comme des contours, des points particuliers, des textures, peuvent être plus simplement analysés et même modifiés, en travaillant dans le domaine des ondelettes plutôt que directement sur les pixels. On peut alors évoluer vers des procédures de plus haut-niveau comme la reconnaissance de formes complexes. Dans le cas d'un *traitement*, comme par exemple la restauration d'une image d'archive, on pourra modifier les données dans le domaine des ondelettes ; ensuite, une transformation *inverse* permettra de les reconvertir en pixels, pour reconstituer l'image restaurée.

Les ondelettes en traitement du signal et des images font référence aujourd'hui, particulièrement dans l'algorithme de compression JPEG-2000. Grâce au formalisme de l'analyse multi-résolution, les ondelettes forment un lien théorique entre les données *discrètes* disponibles et des objets *continus* sous-jacents. Elles se concrétisent par des algorithmes de transformation numériquement stables, assurant la cohérence entre les opérations appliquées dans le domaine des ondelettes d'une part, et les changements physiques résultants dans l'image d'autre part.

Cependant, et comme pour la plupart des outils mathématiques dédiés aux images en général, la définition de transformées en ondelettes se heurte à un problème de *dimension* des données. La plupart des applications utilisent une implantation dite *séparable*, qui applique une méthode 1D séparément sur chaque ligne et chaque colonne de l'image. Cette solution très simple à mettre en œuvre ne permet néanmoins pas de capter la *géométrie* des structures de l'image de façon satisfaisante, et privilégie les formes rectangulaires. C'est pourquoi de nombreuses variantes de transformées en ondelettes 2D, utilisant des schémas non séparables, ont également été proposées (la grande famille des « x-lets »).

Parmi ces multiples propositions, les *ondelettes quaternioniques* (2004) ont retenu notre attention, car c'est à partir d'innovations autour du concept classique de *signal analytique* qu'elles

parviennent à exploiter efficacement la géométrie des structures de l'image. La particularité de cette transformée est que ses coefficients ne sont plus de simples nombres réels mais des vecteurs composés d'une *amplitude* et d'une *phase*. Elle réunit deux travaux fondateurs datant de 1999 : l'algorithme « dual-tree », qui réalise des ondelettes 1D approximativement *analytiques* de façon numériquement stable, et le formalisme de signal quaternionique, qui généralise le *signal analytique* pour les images à travers un concept nouveau de « phase 2D ». L'approche sous-jacente dans ce domaine de recherche est l'étude des signaux 2D à travers les notions oscillatoires de l'analyse de Fourier : amplitude, phase et fréquence. Le but recherché est la qualité et la richesse de l'information portée par les coefficients. D'une part, l'*amplitude* est invariante par translation, propriété partagée avec la transformée de Fourier, mais pas avec la transformée en ondelettes classique. D'autre part, il s'avère que la *phase*, issue d'une caractérisation physique d'un phénomène vibratoire, traduit d'une certaine façon la géométrie locale des structures de l'image. Cette approche *signal*, bien définie en 1D pour l'étude des signaux électriques, radars, ou le son par exemple, suscite actuellement une recherche active visant à étendre ces notions pour les images.

À partir de 2001, un nouvel axe est venu compléter le formalisme quaternionique, dans le but d'améliorer cette analyse des images basée sur un concept de phase. Il repose sur une réelle généralisation de la transformée de Hilbert, à savoir la transformée de Riesz. Le formalisme *monogène* offre alors une définition de « phase 2D » alternative, qui permet de rendre l'analyse invariante par rotation, ce qui n'est pas le cas des outils quaternioniques. De plus, cette nouvelle phase traduit des caractéristiques géométriques plus claires que la phase quaternionique, notamment une mesure directe de l'orientation des structures. Ces propriétés sont très utiles en pratique car les objets contenus dans les images peuvent prendre des directions diverses. Basé sur ce concept prometteur, le premier algorithme de transformée en *ondelettes monogènes* offrant une reconstruction parfaite de l'image à partir des données d'analyse a été proposé en même temps que cette thèse commençait (2009). Le signal monogène constitue une amélioration théorique du signal quaternionique, et la transformée en ondelettes monogène offre de nouvelles perspectives en termes de représentation des images. Nous nous sommes donc fortement intéressé à ces deux généralisations 2D de la transformée en ondelettes analytique.

À partir du constat que l'algèbre des nombres complexes, classiquement utilisée en traitement du signal, est trop limitée pour modéliser les signaux multidimensionnels de manière satisfaisante, les travaux sur l'extension 2D du signal analytique (le signal quaternionique et le signal monogène) ont été proposés dans des algèbres plus générales : l'algèbre des quaternions, et l'algèbre de Clifford. C'est ainsi que les coefficients d'analyse ne sont plus de simples nombres réels mais des objets plus évolués, intégrant les notions d'amplitude et de phase dans une seule entité.

Nous pensons que ce domaine de recherche est propice à étendre les outils de traitement du signal aux signaux *vectoriels*, ce qui permettrait de définir une représentation en ondelettes adaptée aux images *couleur*. En effet, tous les outils que nous avons mentionnés jusqu'ici ne sont applicables que sur des images en *niveaux de gris*. Les pixels des images en niveaux de gris sont composés d'un simple nombre réel, traduisant l'intensité lumineuse, alors que les pixels d'une image couleur consistent en général en *trois* nombres, correspondant aux composantes primaires « rouge », « vert » et « bleu ». En pratique, lorsqu'il est nécessaire d'analyser ou de traiter des images couleur, la méthode la plus utilisée est *marginale*, c'est-à-dire qu'elle consiste à séparer l'image couleur en trois images scalaires, chacune liée à un canal couleur, puis à appliquer séparément un outil scalaire (« gris »). Ces traitements marginaux ont parfois des effets indésirables comme l'apparition de *fausses couleurs* ou la non-détection de certaines formes, ce qui est dû au fait que les couleurs de l'images ne sont pas prises en compte comme des entités, et que donc le traitement n'est pas adapté à cette notion physique de couleur. Mathématiquement, les images couleur posent à nouveau un problème de *dimension*, puisque l'on doit étendre les outils de façon à prendre en compte les signaux dont les valeurs sont *vectérielles*. Notons que ce problème est encore plus important avec des données multimédia plus riches, impliquant par

exemple des objets en trois dimensions, de la video, des images multi- ou hyper- spectrales, des images en stéréo-vision. . .

D'une manière générale en traitement d'image, les outils vectoriels sont rares. En ce qui concerne le signal analytique, la transformée de Fourier et les transformées en ondelettes analytiques, il n'existe pas à notre connaissance aujourd'hui d'outil vectoriel ou couleur parfaitement établi.

C'est ici que nous plaçons notre travail, dans une problématique de redéfinition des outils classiques de traitement du signal pour les images couleur. De tels outils pourraient à l'avenir mieux prendre en compte les informations conjointes entre les canaux couleurs, en manipulant de réelles entités géométriques couleur, tout en conservant la qualité de description géométrique des outils scalaires modernes. Ce travail nécessite une réflexion sur le sens que l'on peut donner à des objets comme une « oscillation couleur », un « contour couleur », une « phase couleur », ainsi qu'à la définition de l'orientation, et plus généralement des caractéristiques géométriques des objets en couleur. Cette démarche passe par une modélisation mathématique particulière des signaux couleur, de façon à identifier certaines de leurs caractéristiques physiques, oscillatoires, géométriques. Elle fera notamment appel aux outils issus de la géométrie différentielle. Nous attendons de cette approche *signal* qu'elle aboutisse à une méthode peu redondante et inversible, grâce aux notions d'échantillonnage et de théorie de l'information.

Le travail de recherche présenté dans ce mémoire de thèse a été guidé par des questions, auxquelles nous tenterons de répondre :

Dans une perspective de manipulation des signaux aux dimensions diverses, quels nouveaux outils peuvent émerger de l'étude d'algèbres générales comme les quaternions ou les algèbres géométriques ?

Dans quelle mesure les notions liées aux phénomènes vibratoires, formant le cœur des outils de traitement du signal et indispensables aux technologies du son, peuvent améliorer les méthodes d'analyse en apportant une dimension physique à l'étude des images ? Et comment une grandeur physique comme la couleur peut-elle être prise en compte efficacement dans une telle démarche ?

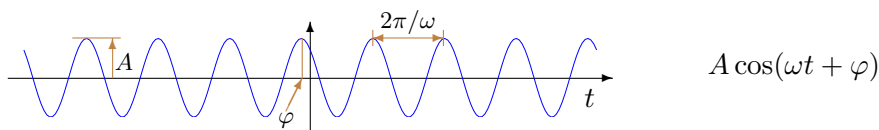
Comment les outils classiques d'analyse d'image peuvent-ils être intégrés dans un processus de synthèse, en d'autres termes, comment des mesures concrètes de caractéristiques physiques peuvent à la fois résumer les structures visibles et permettre d'en reconstruire l'essentiel *a posteriori* ? Et à quel point une méthode de représentation peut-elle concilier puissance d'analyse et compacité des données ?

Ce mémoire commence avec la présentation des outils classiques de traitement du signal, à travers les notions oscillatoires d'amplitude, de phase et de fréquence, pour aboutir à la description de l'algorithme dual-tree, qui ouvre la voie vers des schémas numériques d'ondelettes analytiques. Ensuite, les approches quaternioniques pour le traitement d'image seront traitées, et les ondelettes quaternioniques en particulier seront étudiées en détail, dans un travail d'abord théorique, puis pratique. Les approches isotropes pour l'extension du signal analytique, aboutissant aux ondelettes monogènes, seront étudiées au chapitre 3, ainsi qu'une première réflexion sur l'extension couleur. Enfin, le chapitre 4 présentera une nouvelle définition de représentation en ondelettes monogènes couleur.

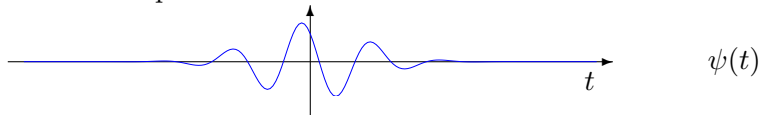
AMPLITUDE, PHASE ET FRÉQUENCE

LES outils élémentaires de représentation des signaux 1D et 2D sur lesquels se base notre travail sont présentés dans ce chapitre. Il s'agit de la transformée de Fourier, du signal analytique et des transformées en ondelettes.

Un phénomène oscillatoire élémentaire peut se modéliser par une fonction sinusoïdale s dépendant du temps t telle que celle-ci :



La transformée de Fourier permet de construire n'importe quelle forme de signal en superposant de telles sinusoïdes pour différentes valeurs d'amplitude A , de phase φ et de fréquence $\frac{\omega}{2\pi}$ (plus précisément, ω est la *pulsation*, exprimée en radians par secondes, et $\frac{\omega}{2\pi}$ est la *fréquence* exprimée en Hertz). L'analyse de Fourier est à l'origine de la discipline du traitement du signal et les caractéristiques oscillatoires sont les notions centrales de notre travail. L'analyse par ondelettes permet de construire n'importe quelle forme de signal en superposant des sinusoïdes *localisées* dans le temps comme celle-ci :



Ces « ondelettes » peuvent être paramétrées par une amplitude, une fréquence et parfois une phase, mais aussi par une *position* dans le temps, et tout signal pourra s'écrire comme une somme d'ondelettes.

Nous présentons dans ce chapitre les outils classiques de *représentation* des signaux en nous focalisant sur ceux qui intègrent les notions d'*amplitude* et de *phase*. Ces derniers utilisent en général l'algèbre des nombres complexes $\mathbb{C}$, par opposition aux représentations « à valeurs réelles », qui sont insuffisantes pour manipuler ces grandeurs physiques.

Le but de ce chapitre est de donner les bases pour l'étude des représentations en ondelettes « analytiques » peu redondantes, avec les difficultés sous-jacentes concernant l'extension de ces outils aux *images*. Ainsi, les premières méthodes d'extension des représentations analytiques

aux signaux 2D seront présentées et permettront d'ouvrir vers les approches quaternioniques et monogènes qui seront étudiées dans les chapitres suivants.

1.1 OUTILS DE BASE 1D

1.1.1 Domaine fréquentiel

La transformée de Fourier décompose toute fonction $s \in L^2(\mathbb{R})$ en une somme de signaux élémentaires oscillants tel que :

$$s(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{s}(\omega) e^{j\omega t} d\omega \quad \xleftrightarrow{\mathcal{F}} \quad \hat{s}(\omega) = \int_{\mathbb{R}} s(t) e^{-j\omega t} dt = A(\omega) e^{j\varphi(\omega)} \quad (1.1)$$

Les termes $e^{j\omega t} = \cos(\omega t) + j \sin(\omega t)$ sont des sinusoides complexes, correspondant à une « fréquence pure ». Les coefficients complexes $\hat{s} = Ae^{j\varphi}$ définis par une *amplitude* A et une *phase* φ composent la transformée de Fourier de s et caractérisent complètement ce signal. La plupart du temps, les signaux étudiés sont à valeurs réelles ($s(t) \in \mathbb{R}$), et la symétrie hermitienne de leur spectre $\hat{s}(\omega) = \hat{s}^*(-\omega)$ nous permet de les représenter comme une somme plus intuitive de sinusoides réelles, décomposée en une « partie paire » et une « partie impaire » :

$$s(t) = \frac{1}{\pi} \int_{\mathbb{R}^+} \Re\{\hat{s}(\omega)\} \cos(\omega t) d\omega - \frac{1}{\pi} \int_{\mathbb{R}^+} \Im\{\hat{s}(\omega)\} \sin(\omega t) d\omega \quad (1.2)$$

$$= \frac{1}{\pi} \int_{\mathbb{R}^+} A(\omega) \cos(\omega t + \varphi(\omega)) d\omega \quad (1.3)$$

On voit bien dans cette equation que le *module* et l'*argument* des coefficients de Fourier encodent directement l'*amplitude* et la *phase* pour chaque fréquence. L'amplitude est liée à une notion d'énergie, comme le montre l'égalité de Parseval-Plancherel $\int s(t)^2 = \frac{1}{2\pi} \int A(\omega)^2$, et donne accès à la puissance moyenne de la sinusoïde $A \cos(\omega t + \varphi)$ qui vaut $A^2/2$. Elle possède une propriété appréciée en analyse, l'invariance par translation, régie par le théorème du retard (ou du « décalage ») :

$$s(t - \tau) \quad \xleftrightarrow{\mathcal{F}} \quad e^{-j\omega\tau} \hat{s}(\omega) = A(\omega) e^{j(\varphi(\omega) - \omega\tau)} \quad (1.4)$$

Il est clair que l'analyse de Fourier est fondamentalement liée à l'algèbre des nombres complexes $\mathbb{C}$, alors même que les signaux considérés sont réels. Les exponentielles complexes sont en fait les vecteurs propres des systèmes linéaires invariant dans le temps (ou « à coefficients constants »), dont l'application à un signal s'effectue par une *convolution*. En d'autres termes, les sinusoides complexes sont les seules fonctions qui lorsqu'on les passe dans un système linéaire invariant dans le temps, en ressortent simplement « pondérées » par un facteur complexe. Cette pondération affecte l'amplitude et la phase. Il faut de plus reconnaître que cette algèbre très répandue permet d'unifier de façon élégante les notions physiques d'amplitude (réel positif) et de phase (angle), par le module et l'argument, dans *un seul* coefficient ; et de modéliser une double opération d'amplification par A et de déphasage par φ par une simple *multiplication* $\Re\{Ae^{j\varphi} e^{j\omega t}\} = A \cos(\omega t + \varphi)$.

Notons que la discrétisation de la transformée de Fourier est très bien établie grâce au formalisme des *distributions*. Une séquence $s[n]$ de N nombres caractérise un signal T -périodique continu $s(t)$ dont on a extrait N échantillons consécutifs à intervalles réguliers T/N . La définition de la transformée de Fourier discrète d'un signal $s[n]$ défini pour $n = 0..N-1$ est :

$$\hat{s}[k] = \sum_{n=0}^{N-1} s[n] e^{-j2\pi kn/N} \quad s[n] = \frac{1}{N} \sum_{k=0}^{N-1} \hat{s}[k] e^{j2\pi kn/N} \quad (1.5)$$

Cette représentation non-redondante est calculée classiquement par l'algorithme rapide « FFT ».

Les informations obtenues dans le domaine fréquentiel concernent des sinusoides infinies et ne sont donc pas localisées dans le temps. Nous allons voir qu'au contraire, le *signal analytique* fournit une amplitude et une phase *instantanées*.

1.1.2 Domaine temporel

La représentation à valeurs complexes de la transformée de Fourier formée d'une amplitude et d'une phase trouve son analogie temporelle dans le *signal analytique*, construit à partir de la transformée de Hilbert [55, 151].

Transformée de Hilbert

La transformée de Hilbert est apparue dans la communauté mathématique comme la relation entre la partie réelle et la partie imaginaire des fonctions holomorphes [78], et se définit comme une transformation d'une fonction s , par la valeur principale de Cauchy de l'intégrale singulière suivante :

$$\mathcal{H}s(t) = v.p. \int \frac{s(t-\tau)}{\pi\tau} d\tau \quad \xleftrightarrow{\mathcal{F}} \quad -j \operatorname{sgn}(\omega) \hat{s}(\omega) \quad (1.6)$$

En dépit de sa singularité, elle possède la plupart des propriétés classiques des opérateurs linéaires invariants par translation [78]. Unitaire dans L^2 , elle conserve l'énergie ($\int |s|^2 = \int |\mathcal{H}s|^2$) et son inverse/adjoint se réduit à $-\mathcal{H}$. Le théorème de Bedrosian [6], que nous utiliserons par la suite, donne une condition suffisante pour que $\mathcal{H}$ commute avec la multiplication :

$$\hat{s}_1(\omega)\hat{s}_2(\omega) = 0 \quad \Rightarrow \quad \mathcal{H}\{s_1 s_2\}(t) = s_1(t)\mathcal{H}s_2(t) \quad (1.7)$$

(Les deux signaux occupent une bande fréquentielle distincte). La réponse fréquentielle de $\mathcal{H}$ montre qu'elle déphase de $-\pi/2$ toutes les sinusoïdes qui composent le signal. Dans le domaine temporel, cela correspond à un décalage à droite de $\frac{\pi}{2\omega}$ (différent pour chaque sinusoïde), qui se traduit sur le signal par une modification locale de sa structure oscillante, comme illustré figure 1.1a. Le déphasage local de la transformée de Hilbert se distingue donc d'un simple décalage. L'amplitude n'est pas affectée à l'exception de la composante continue qui est annulée, $\mathcal{H}$ est donc un filtre « passe-tout », un « déphaseur pur ». On dit que s et $\mathcal{H}s$ forment une « paire de Hilbert », et sont « en quadrature » (terme qui renvoie à l'orthogonalité entre les fonctions, puisque $\langle s, \mathcal{H}s \rangle = 0$, autant qu'à « l'angle droit » - 90° qui sépare leurs phases respectives).

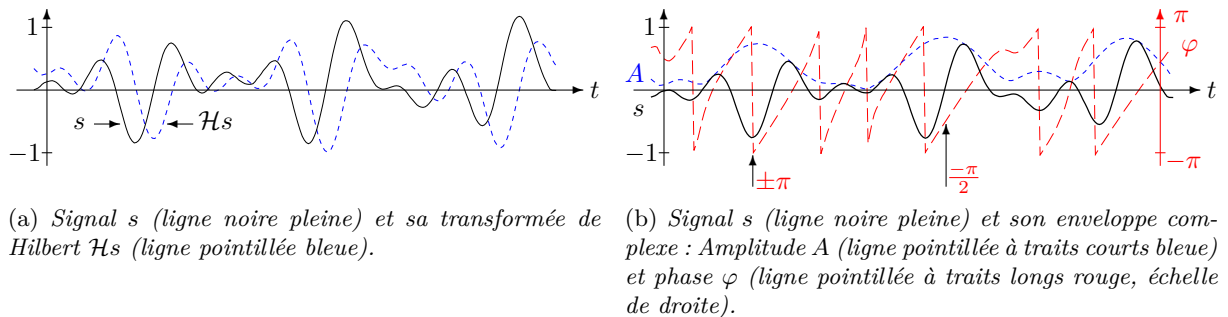


FIGURE 1.1 – *Signal analytique 1D*

Remarquons le lien étroit entre $\mathcal{H}$ et la *dérivation* :

$$\mathcal{H}s(t) \quad \xleftrightarrow{\mathcal{F}} \quad -\frac{j\omega}{|\omega|} \hat{s}(\omega) \quad \frac{ds(t)}{dt} \quad \xleftrightarrow{\mathcal{F}} \quad j\omega \hat{s}(\omega) \quad (1.8)$$

Plus particulièrement, dans le cas de la sinusoïde, la transformée de Hilbert et la dérivée sont égales au signe près : $\mathcal{H} \cos(t) = \sin(t) = -\frac{d \cos(t)}{dt}$. Ce lien sera abordé à nouveau au chapitre 3.

L'implantation discrète ne peut malheureusement pas se définir aussi bien que pour la transformée de Fourier. En effet, la transformée de Hilbert est associée à une réponse impulsionnelle infinie, ce qui empêche le calcul numérique exact de $\mathcal{H}s$. En pratique, on calculera $\mathcal{H}s$ en multipliant ses coefficients de FFT par $\pm j$. On considère en général que cette approximation est satisfaisante.

Notons que la transformée de Hilbert n'est pas adaptée pour manipuler des signaux de moyenne non-nulle, à cause de sa singularité autour de $\omega = 0$.

Phase φ	0	$\frac{\pi}{2}$	$\pm\pi$	$-\frac{\pi}{2}$
Forme locale				

TABLE 1.1 – Correspondance entre la phase et la forme locale [61, 15].

Signal analytique

Le *signal analytique* est un signal à valeurs complexes s_A construit à partir d'un signal réel s afin d'extraire son « enveloppe ». Il est défini par :

$$s_A(t) = s(t) + j\mathcal{H}s(t) = A(t)e^{j\varphi(t)} \xrightarrow{\mathcal{F}} (1 + \text{sgn}(\omega))\hat{s}(\omega) \quad (1.9)$$

On peut remarquer que cette extension de s correspond simplement à annuler les « fréquences négatives » de son spectre, ce qui a pour effet de convertir toutes les sinusoïdes réelles $A(\omega) \cos(\omega t + \varphi(\omega))$ qui le composent en sinusoïdes complexes $A(\omega)e^{j\varphi(\omega)}$ correspondantes. On a donc une représentation de redondance $2\times$.

L'enveloppe s_A est composée d'une amplitude $A(t)$ et d'une phase $\varphi(t)$ *instantanées*, la phase pouvant être dérivée pour obtenir la fréquence instantanée du signal $\nu(t) = \frac{1}{2\pi} \frac{d}{dt} \varphi(t)$. Cette extraction de caractéristiques locales du signal réalise une *démodulation* en amplitude et en fréquence (AM/FM), ce qui a bien sûr une application en télécommunications, mais aussi pour la détection de la hauteur des sons par exemple.

On peut voir figure 1.1b que l'amplitude A enveloppe le signal de façon lisse. Cette donnée analogue à celle de Fourier s'interprète comme l'amplitude d'une sinusoïde, mais cette fois-ci localisée dans le temps, indépendamment de sa fréquence et de sa phase. On peut la voir comme traduisant la *présence* d'une oscillation locale. La phase instantanée $\varphi(t)$, dont les valeurs sont dans $[-\pi; \pi]$, traduit localement la position du signal par rapport à son oscillation courante. Par exemple, les deux points particuliers de la figure 1.1b représentent respectivement un « creux » et une « montée », ce qui correspond, au point central de ces structures, à des valeurs de phase de $\pm\pi$ et $-\pi/2$. On peut alors interpréter la phase comme un descripteur continu de forme géométrique locale [81, 61], dont les cas particuliers sont illustrés dans la table 1.1. Le signal analytique peut donc être vu comme une représentation locale en termes de *présence* et de *forme*. Cette interprétation structurelle ouvre de nouvelles perspectives dans l'extension de cet outil pour l'image.

La modélisation du signal à partir de son enveloppe et de sa phase

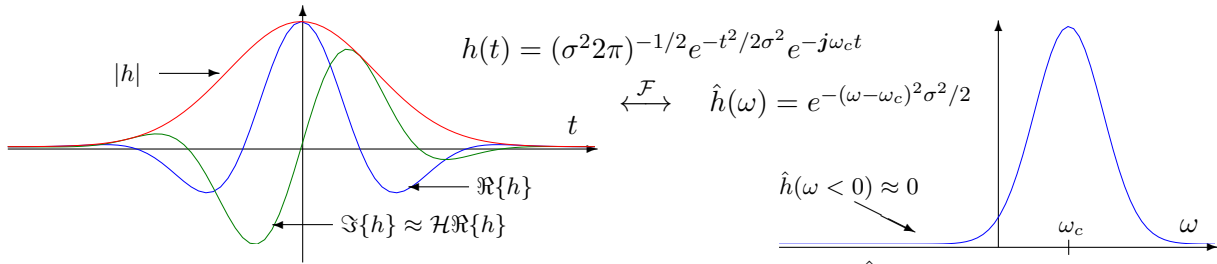
$$s(t) = A(t) \cos(\varphi(t)) \quad (1.10)$$

révèle le modèle sous-jacent à cette représentation : une sinusoïde modulée en amplitude et en phase/fréquence. Nous avons pointé précédemment que la transformée de Hilbert est mal adaptée pour les signaux à moyenne non-nulle, nous allons maintenant voir qu'un signal à *bande étroite* est requis pour les deux raisons suivantes :

D'abord, le concept de fréquence instantanée est incompatible avec l'éventuelle présence de *plusieurs* composantes fréquentielles dans s . Par exemple, un signal contenant deux fréquences d'amplitudes constantes sera assimilé à une seule fréquence modulée en amplitude, par l'identité trigonométrique suivante :

$$\cos(\omega_1 t + \varphi_1) + \cos(\omega_2 t + \varphi_2) = 2 \cos\left(\frac{\omega_1 + \omega_2}{2} t + \frac{\varphi_1 + \varphi_2}{2}\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t + \frac{\varphi_1 - \varphi_2}{2}\right) \quad (1.11)$$

Dans le cas où les deux fréquences concernées sont proches ($\omega_1 \approx \omega_2$), cela est bien connu en acoustique sous le nom de phénomène de « battement ». Le modèle du signal analytique est alors cohérent avec la perception auditive, puisque l'on entend une seule fréquence dont l'intensité varie. Par contre, dans le cas où les deux fréquences sont bien distinctes, la modélisation des deux

FIGURE 1.2 – Filtre en quadrature de Gabor h et sa réponse fréquentielle $\hat{h}$.

oscillations par une onde modulée n'est pas satisfaisante, ce qui suggère clairement l'étude de signaux à bande étroite, qui ne peuvent contenir simultanément que des fréquences suffisamment proches.

Ensuite, il semble raisonnable que la fréquence soit relativement constante au voisinage d'un cycle, ce qui implique une certaine *régularité* de la phase et donc du signal. D'un point de vue « Fourier », la régularité signifie l'absence de hautes fréquences, corroborant à nouveau la contrainte de bande étroite.

Cette restriction suggère donc d'analyser une certaine *bande* fréquentielle du signal. Afin d'avoir une approche complète, prenant en compte *toutes sortes* de signaux, il faut utiliser la représentation analytique dans une analyse *multibande*.

1.2 ANALYSE MULTIBANDE 1D

Au coeur des techniques de traitement du signal, les représentations multibandes sont le compromis nécessaire entre le domaine fréquentiel - dont les informations ne sont pas localisées dans le temps - et le domaine spatial. Notons que ce genre d'outil est en accord avec la perception humaine du son, ce qui explique son succès par exemple en codage de la parole pour les télécommunications. Les transformées en ondelettes facilitent ainsi l'analyse des signaux en localisant l'information à la fois en temps et en fréquence, et permettent de les approximer efficacement par des reconstructions partielles simples.

Après la présentation de l'outil classique qu'est la représentation de Gabor, nous verrons les bancs de filtres orthogonaux et leurs extensions à valeurs complexes. De plus amples détails pourront être trouvés dans [102, 150, 49].

1.2.1 Représentations continues : Gabor et Morlet

Atomes de Gabor

A cause de la contrainte de bande étroite sur la représentation analytique, l'analyse de l'enveloppe temporelle des signaux se fait en général par un *filtre en quadrature*, qui sélectionne une bande fréquentielle tout en calculant la transformée de Hilbert associée. Un tel filtre est défini à partir d'un filtre passe-bande réel $h_1(t)$ et par son extension analytique $h(t) = h_1(t) + j\mathcal{H}h_1(t)$ (voir figure 1.2). La linéarité de $\mathcal{H}$ implique que le signal filtré par h est lui aussi analytique. En pratique, puisque $\mathcal{H}$ est associé à une réponse impulsionnelle à support infini, une paire de Hilbert ne peut pas être constituée de deux filtres à support compacts. Il est donc courant d'utiliser des approximations de paire de Hilbert, comme le célèbre atome de Gabor défini par le fenêtrage d'une exponentielle complexe, illustré figure 1.2 dans la cas d'une fenêtre gaussienne. L'approximation se traduit par une énergie non-nulle dans la partie négative du spectre, due à sa forme gaussienne. Cette approximation peut aussi être expliquée par le théorème de Bedrosian (equation (1.7)). Ce dernier indique que si les supports des spectres de la fenêtre $e^{-t^2/2\sigma^2}$ et de la sinusoïde $\cos(\omega_c t)$ sont *distincts*, alors on peut effectuer la commutation suivante :

$$e^{-t^2/2\sigma^2} \mathcal{H} \cos(\omega_c t) = \mathcal{H} \left\{ e^{-t^2/2\sigma^2} \cos(\omega_c t) \right\} \quad (1.12)$$

et réécrire $h(t)$ comme le signal analytique (exact) associé à l'atome $e^{-t^2/2\sigma^2} \cos(\omega_c t)$. En réalité, les supports ne sont pas exactement distincts, mais le sont approximativement si la fréquence centrale ω_c est suffisamment grande par rapport à la largeur de la gaussienne σ . Le choix d'une fenêtre gaussienne n'est pas anodin car il correspond au meilleur compromis entre la localisation temporelle et la localisation fréquentielle.

En faisant varier la position de la fenêtre et la fréquence de l'exponentielle complexe, on obtient une famille d'atomes temps-fréquence. Cela correspond à une transformée de Fourier fenêtrée (STFT, pour *Short Time Fourier Transform*), définie pour un signal $s(t)$ de $L^2(\mathbb{R})$ comme ceci :

$$w(\omega, x) = \int_{\mathbb{R}} s(t) e^{-j\omega t} g(t - x) dt \quad (1.13)$$

où g est la fenêtre utilisée. Le coefficient $w(\omega, x)$ mesure l'amplitude et la phase du signal autour de la position x et de la fréquence ω , constituant une fonction de deux variables. L'identité de Parseval $\int |s|^2 = \frac{1}{2\pi} \int |w|^2$ confère au carré de l'amplitude $|w|^2$ une valeur d'énergie qui définit le « spectrogramme » de s . La mesure associée à chaque coefficient se fait sur un certain *voisinage* en temps ainsi qu'en fréquence, ce qui introduit une certaine régularité de w ¹. Une implantation discrète classique de la STFT consiste à effectuer N transformées de Fourier pour un signal de taille N , ce qui produit N^2 coefficients.

La reconstruction de s à partir de ses coefficients est définie *en continu* par :

$$s(t) = \frac{1}{2\pi} \int_{\mathbb{R}^2} w(\omega, x) e^{j\omega t} g(t - x) d\omega dx \quad (1.14)$$

modélisant donc le signal comme une somme d'oscillations locales à des fréquences différentes, mais ne fonctionne pas toujours lorsque l'on discrétise la transformée (on ne peut pas calculer un ensemble continu de coefficients).

La particularité des atomes de Gabor à osciller dans une fenêtre de durée fixée fait que les atomes haute-fréquence présentent de plus nombreuses oscillations que les atomes basse-fréquence. Cette variabilité de *forme* a suscité la recherche vers une analyse temps-fréquence dont les atomes se ressembleraient plus, notamment pour l'analyse de sismogrammes [62]. Il s'agit des ondelettes de Morlet.

Ondelettes de Morlet

Une évolution de cet outil est la représentation en ondelettes de Morlet/Gabor [62], qui reprend l'atome de Gabor :

$$\psi(t) = \frac{1}{(\sigma^2\pi)^{1/4}} e^{-t^2/2\sigma^2} e^{j\omega_c t} \quad \xleftrightarrow{\mathcal{F}} \quad \hat{\psi}(\omega) = (\sigma^2\pi)^{1/4} e^{-(\omega - \omega_c)^2 \sigma^2 / 2} \quad (1.15)$$

pour générer une famille d'*ondelettes*. Au lieu de simplement décaler la fenêtre, on va plutôt *dilater* et *translater* l'atome tout entier :

$$\psi_{s,x}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-x}{s}\right) \quad \xleftrightarrow{\mathcal{F}} \quad \sqrt{s} e^{-j\omega x} \hat{\psi}(s\omega) \quad \text{avec } s, x \in \mathbb{R} \quad (1.16)$$

Cela permet d'une part d'avoir des atomes qui portent tous la même énergie, puisque $\int |\psi_{s,x}|^2 = \int |\psi|^2$, mais plus généralement les fonctions analysantes seront toutes *semblables*, contrairement aux atomes de la STFT qui contiennent un nombre d'oscillations différent.

La notion d'*échelle* s vient remplacer la notion de fréquence, et si ψ a une fréquence centrale ω_c , celle de ψ_s sera $s \times \omega_c$. La décomposition est équivalente à des filtrages *passes-bande* :

$$w_s(x) = \int_{\mathbb{R}} s(t) \psi_{s,x}(t) dt = (s * \psi_{s,0})(x) \quad (1.17)$$

1. Elle est égale à son lissage par un *noyau reproducteur* (voir aussi *fonction d'ambiguïté*), et ne produit pas d'interférences, contrairement à la distribution de Wigner-Ville.

Le module au carré des coefficients $|w|^2$ s'interprète comme un « scalogramme », analogue au spectrogramme mais avec un découpage « log-fréquentiel » (par octaves), impliquant des bandes plus grandes en hautes-fréquences, à l'image de notre perception du son.

La STFT de Gabor et les ondelettes de Morlet, que nous engloberont par la suite par le seul nom « Gabor », possèdent une bonne qualité d'analyse grâce à une amplitude invariante par translation et une résolution temps-fréquence très fine. Elle est utile pour l'analyse des signaux audio par exemple, qui peuvent souvent être modélisés par une somme d'oscillations modulées en amplitude et en fréquence. L'analyse des *formants* de la parole (composantes sinusoïdale apparaissant pour une courte durée et dont la fréquence varie lentement), peut se faire dans le domaine de Gabor notamment via la détection de maxima de w (notion de « ridges »).

La condition d'*admissibilité* sur ψ (de moyenne nulle et continûment dérivable) assure la reconstruction du signal à partir de l'ensemble *continu* des coefficients, et de nombreuses ondelettes à valeurs réelles ont été proposées (Chapeau Mexicain, Shannon, Meyer, Haar, Battle-Lemarié, Daubechies...), avec différentes propriétés d'orthogonalité, de taille de support, de symétrie et de nombre de moments nuls.

Dans le cadre de cette thèse, seules nous intéressent les ondelettes *analytiques* (ou approximativement), car elles permettent d'extraire l'amplitude et la phase locale du signal analysé. Malheureusement, ces représentations définies dans le domaine *continu* posent un problème de discrétisation, et ne permettent pas en pratique la reconstruction stable du signal à partir des coefficients d'analyse. C'est pourquoi l'algorithme de S. Mallat [101] que nous présentons maintenant a marqué l'essor des représentations en ondelettes.

1.2.2 Bancs de filtres orthogonaux à coefficients réels

Les décompositions temps-fréquence peuvent être vues comme des projections du signal dans des bases d'ondelettes de $L^2(\mathbb{R})$. L'article de S. Mallat [101] montre que l'on peut produire tous les coefficients de projection efficacement avec un banc de filtres à échantillonnage critique et reconstruction parfaite. Ce résultat est issu du recoupement entre l'*analyse multirésolution* [106], et le *codage en sous-bandes*, apparu dans les années 1970 en communications [32].

Une *analyse multirésolution* (AMR) est une séquence de sous-espaces emboîtés $\{V_j\}_{j \in \mathbb{Z}}$ de $L^2(\mathbb{R})$ tels que $V_j \subset V_{j+1}$. Cette théorie permet l'étude des approximations d'un signal dans lesquelles les détails sont progressivement supprimés par lissages successifs. A chaque valeur de j correspond une notion d'*échelle*, les signaux les plus « détaillés » étant pour des valeurs de j grandes². Pour une AMR donnée, il existe une unique *fonction d'échelle* $\phi(t)$ telle que ses translations et dilatations

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (j, k) \in \mathbb{Z}^2 \quad (1.18)$$

forment des bases orthonormales des V_j . On approxime donc un signal s à l'échelle j en le projetant dans V_j :

$$s_j(t) = \sum_k \underbrace{\langle s, \phi_{j,k} \rangle}_{v_j[k]} \phi_{j,k}(t) \in V_j \quad (1.19)$$

Les coefficients de cette projection sur une base *discrète* forment un signal discret $v_j[k]$ qui approxime la fonction continue s . Cette relation permet de faire apparaître une convolution discrète dans une expression faisant intervenir les coefficients à deux échelles voisines. Ainsi, les coefficients d'approximation de s à l'échelle j peuvent être obtenus en filtrant les coefficients de l'échelle $j+1$ et en ne gardant qu'un échantillon sur deux :

$$v_j[n] = \langle s, \phi_{j,n} \rangle = \sum_k \tilde{h}[2n - k] \langle s, \phi_{j+1,k} \rangle = (\tilde{h} * v_{j+1})[2n] \quad (1.20)$$

2. Nous reprenons ici pour les valeurs de j la définition de l'article de Mallat [101].

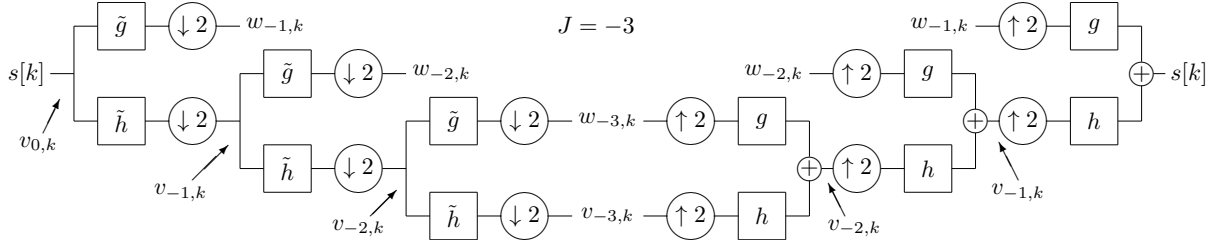


FIGURE 1.3 – Banc de filtres à échantillonnage critique 1D. Dans cet exemple la profondeur de décomposition choisie est de trois niveaux.

Le filtre h , appelé *filtre d'échelle*, est un filtre passe-bas déterminé par ϕ :

$$h[n] = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \phi\left(\frac{t}{2}\right) \phi(t-n) dt \quad \tilde{h}[n] = h[-n] \quad (1.21)$$

L'orthogonalité de la base formée par les translations de la fonction d'échelle ($\langle \phi_{j,k}, \phi_{j,k+n} \rangle = \delta[n]$) impose les contraintes suivantes sur ce filtre :

$$|\hat{h}(0)| = 1 \quad \text{et} \quad |\hat{h}(\omega)| \neq 0 \quad \text{pour} \quad \omega \in [0; \pi/2] \quad \text{« filtre passe-bas »} \quad (1.22)$$

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 1 \quad \text{« filtre conjugué »} \quad (1.23)$$

Les coefficients dits *d'ondelette* sont construits comme « signal de détail », correspondant aux détails perdus lorsque l'on passe à une résolution inférieure. On les obtient en projetant le signal dans les espaces complémentaires W_j , qui sont tels que $W_j \oplus V_j = V_{j+1}$. Les bases concernées sont des bases d'ondelettes, construites par translation/dilatation d'une *fonction d'ondelette* ψ . A nouveau, une convolution discrète apparaît dans l'expression des coefficients entre deux échelles. Pour satisfaire l'orthogonalité de la base d'ondelettes, le *filtre d'ondelette* g qui apparaît doit satisfaire :

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 1 \quad \text{« filtre conjugué »} \quad (1.24)$$

$$\hat{h}(\omega) \hat{g}^*(\omega) + \hat{h}(\omega + \pi) \hat{g}^*(\omega + \pi) = 0 \quad (1.25)$$

Ces deux conditions sont vérifiées si g est défini en fonction du filtre d'échelle de la façon suivante :

$$g[n] = (-1)^{1-n} h[1-n] \quad \tilde{g}[n] = g[-n] \quad (1.26)$$

La fonction d'ondelette est alors définie dans le domaine de Fourier par : $\hat{\psi}(\omega) = \hat{g}(\omega/2) \hat{\phi}(\omega/2)$.

L'utilisation conjointe des détails et approximations permet de décomposer les signaux sur un nombre fini d'échelles. On dispose en pratique d'un signal discret $s[n]$, dont on fait l'hypothèse qu'il encode l'approximation à l'échelle 0 d'un signal $s(t)$ continu, soit : $s[n] = v_{0,n}$. La décomposition à l'échelle suivante donne une approximation plus grossière $v_{-1,k}[n] = (v_{0,k} * \tilde{h})[2n]$ complétée par un signal de détail $w_{-1,k}[n] = (v_{0,k} * \tilde{g})[2n]$. L'itération sur le signal d'approximation $v_{j,k}$ permet d'obtenir tous les coefficients de projection sur la base d'ondelettes. On arrête la décomposition à une profondeur fixée J (négative), telle que les détails aux échelles $j < J$ sont alors tous regroupés dans le seul signal d'approximation restant $v_{J,k}$. La décomposition complète de s sur $-J$ échelles s'écrit finalement :

$$s(t) = \sum_{k=0}^{N-1} v_{J,k} \phi_{J,k} + \sum_{j=J}^0 \sum_{k=0}^{N-1} w_{j,k} \psi_{j,k} \quad (1.27)$$

Le processus complet de décomposition et de reconstruction à partir des coefficients peut s'effectuer efficacement de façon récursive, à travers une structure de banc de filtres à échantillonnage critique, comme illustré figure 1.3. Une telle utilisation correspond au cas particulier de banc de filtres *orthogonal*, qui était déjà étudié dans le domaine du codage en sous-bandes

[150]. Dans ce cas, les filtres de décomposition $\tilde{h}$ et $\tilde{g}$ sont définis par retournement temporel des filtres de reconstruction h et g (comme nous l'avons spécifié aux équations (1.21) et (1.26)), et respectent les conditions de *paraunitarité*, correspondant aux équations (1.23), (1.24) et (1.25). Au final, une décomposition en ondelettes orthogonale est entièrement déterminée par le filtre d'échelle h . En termes de décomposition et reconstruction d'un signal, on peut relâcher la contrainte d'orthogonalité, permettre $\tilde{h}$ et $\tilde{g}$ d'être différents de h et g , et se concentrer simplement sur la capacité du banc de filtres à reconstruire parfaitement le signal de départ. Les conditions plus générales de *reconstruction parfaite* sur les 4 filtres sont les suivantes :

$$\begin{cases} \hat{h}(\omega)\hat{\tilde{h}}(\omega) + \hat{g}(\omega)\hat{\tilde{g}}(\omega) = 2 \\ \hat{h}(\omega + \pi)\hat{\tilde{h}}(\omega) + \hat{g}(\omega + \pi)\hat{\tilde{g}}(\omega) = 0 \end{cases} \quad (1.28)$$

Dans ce cas la décomposition est dite « bi-orthogonales », et il est possible d'avoir des filtres symétriques à support compact et coefficients réels [33].

Cet outil est puissant pour représenter les signaux réguliers par morceaux, c'est-à-dire présentant des singularités éparées. Cette nouvelle famille de décompositions en temps et en fréquence est non-redondante et se calcule par un algorithme rapide de complexité linéaire. Il connaît un réel succès dans des domaines très variés, dont la compression d'images, aboutissant à l'intégration d'une transformée en ondelettes dans le format JPEG-2000. L'orthogonalité de la décomposition en fait un outil de choix pour les analyses statistiques car le bruit blanc y est décorrélié. L'intégration rigoureuse du discret dans la théorie permet la reconstruction parfaite, et donc le surpassement des problèmes de discrétisation posés par les ondelettes continues.

En revanche, le concept de phase est totalement perdu car les coefficients sont considérés réels dans la plupart des cas traités. Ceci est associé d'une part à une forte variance par translation, qui n'existe pas dans les représentation analytiques et dans la transformée de Fourier, et d'autre part à l'introduction d'oscillations indésirables (effet « pseudo-Gibbs ») par traitements non-linéaires dans le domaine des ondelettes. Toutes ces contreparties ont amené une communauté grandissante à travailler sur des bancs de filtres complexes, voire analytiques, dont nous allons parler maintenant, afin de trouver un compromis entre les représentations analytiques continues et les bancs de filtres orthogonaux.

1.2.3 Extensions complexes

Suite à la constatation du défaut de variance et du manque d'interprétation amplitude/phase des ondelettes, et vu la qualité d'analyse des ondelettes continues analytiques, il semble naturel de s'intéresser à la définition de bancs de filtres *complexes*, éventuellement légèrement redondants, et si possible analytiques, au moins approximativement.

Dès 1993, une version complexe des ondelettes de Daubechies permet de concilier orthogonalité, symétrie et support compact [92, 99, 157, 56]. La redondance de $2\times$ réside dans les valeurs *complexes* des coefficients (deux nombres réels). En 1994, il a été proposé d'appliquer la transformée de Hilbert sur le signal *avant* une décomposition en ondelettes classique [1, 16] (permis par la linéarité de $\mathcal{H}$), mais cela a l'inconvénient de supprimer sa composante continue. En 2003, une solution proposée dans [149] consiste à utiliser un filtrage complexe supprimant les fréquences négatives en sortie d'un banc de filtres classique. Avec un sous-échantillonnage de 4, on obtient une redondance de $1\times$ ($N/2$ coefficients complexes pour N échantillons de départ). A l'exception de cette non-redondance, cette méthode n'a pas montré à notre connaissance de propriété significative comme l'invariance par translation où l'interprétation géométrique de la phase.

Une alternative intéressante et surtout plus flexible pour des ondelettes analytiques est d'utiliser deux décompositions orthogonales réelles en parallèle, l'une étant approximativement « la transformée de Hilbert de l'autre » (redondance $2\times$). Nous présentons maintenant cette méthode à travers l'algorithme « dual-tree » qui a été proposé en 1998 par N. Kingsbury [79, 124].

Cet outil constitue également la base de la transformée en ondelettes quaternioniques, que nous étudierons au chapitre suivant.

Le « dual-tree »

L'extension du formalisme des ondelettes aux valeurs complexes ne permet pas de définir une décomposition *analytique*. D'une part, l'équation (1.23) empêche les filtres d'avoir une réponse *faible* pour les fréquences négatives, d'autre part, le découpage fréquentiel associé est trop contraint pour avoir une *sélectivité* satisfaisante. Afin d'avoir plus de degrés de libertés dans la construction, la transformée en ondelettes complexes « dual-tree » (DT-CWT) [124] consiste à appliquer deux transformées en ondelettes orthogonales en parallèle, dont la définition des filtres (à support fini) s'approche le plus possible d'une relation de Hilbert entre les deux analyses (voir figure 1.4). Il s'agit donc de définir un premier banc de filtre à partir des filtres passe-bas et

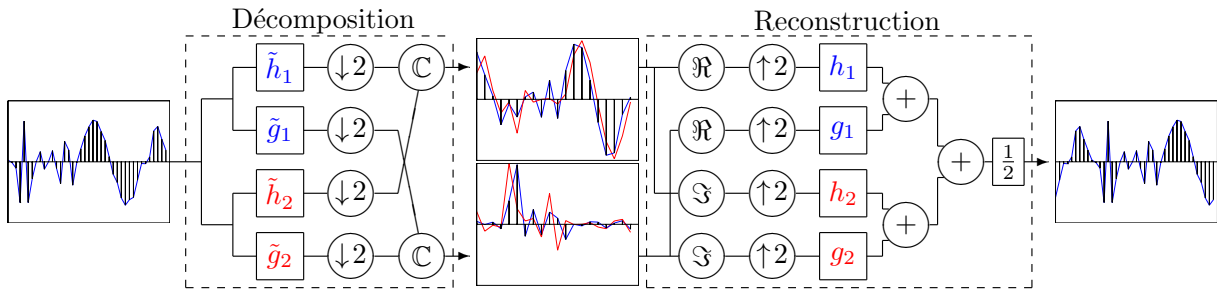


FIGURE 1.4 – Banc de filtres « dual-tree » 1D, sur un seul niveau de décomposition.

passe-haut $h_1[n]$ et $g_1[n]$, associé à l'ondelette $\psi_1(t)$; et un second à partir de $h_2[n]$ et $g_2[n]$, associé à l'ondelette $\psi_2(t)$, tels que $\psi_2(t) \approx \mathcal{H}\psi_1(t)$. Il suffit ensuite de combiner les sorties des deux bancs de filtres comme parties réelle et imaginaire d'une sortie complexe.

Il est montré dans [155] que la contrainte d'analyticité dans le cas orthogonal est équivalente à :

$$h_2[n] = h_1[n - 0.5] \quad \xleftrightarrow{\mathcal{F}} \quad \hat{h}_2(\omega) = e^{-j\omega/2} \hat{h}_1(\omega) \quad (1.29)$$

ce qui n'est bien sûr pas réaliste puisqu'il implique des indices non entiers, et rappelle l'impossibilité de définir une paire de Hilbert à support compact. Plusieurs solutions relevant du *design* de bancs de filtres sont alors possibles [124] pour approcher cette contrainte, et ainsi approcher l'invariance par translation de l'amplitude.

La solution « odd-even » proposée par Kingsbury [79] - qui avait alors pour seul but l'invariance par translation - consiste à utiliser deux passe-bas symétriques, l'un de taille paire et l'autre de taille impaire (notons que la symétrie implique la bi-orthogonalité). La relation (1.29) est alors partiellement obtenue, car on obtient la différence de phase voulue $\arg\{\hat{h}_2(\omega)\} = e^{-j\omega/2} \arg\{\hat{h}_1(\omega)\}$. La définition de tels filtres doit être telle que leurs gains $|\hat{h}_1|$ et $|\hat{h}_2|$ soient le plus proche possible.

La solution « q-shift », que nous retiendrons dans nos expérimentations, consiste à définir un des deux filtres passe-bas par la version retournée de l'autre : $h_2[n] = h_1[N - 1 - n]$ (où N est le nombre *pair* de coefficients du filtre). On obtient nécessairement le même gain fréquentiel en amplitude, et la définition des filtres doit être telle que la différence de phase soit proche de $\omega/2$. Il vient que cette contrainte implique une symétrie approximative de h_1 autour de $n = 0.5(N - 1) - 0.25$, point qui se trouve à un *quart* d'unité (d'où le nom « q-shift ») avant l'indice naturel de symétrie. Différentes solutions dans le cas orthogonal sont données dans [80], la cas où les filtres sont de longueur 10 est donné dans la table 1.2.

Des ajustements algorithmiques sont également proposés. En ce qui concerne la première échelle, les auteurs se sont aperçus que l'approximation de la condition (1.29) est significativement meilleure en utilisant des filtres différents pour le premier niveau de décomposition [124]. Il est également suggéré d'inverser les bancs de filtres à chaque niveau, pour compenser les légères différences de gain fréquentiel entre les deux décompositions.

n	-4	-3	-2	-1	0
$h[n]$	0.03516384	0	-0.08832942	0.23389032	0.76027237
n	1	2	3	4	5
$h[n]$	0.58751830	0	-0.11430184	0	0

TABLE 1.2 – Coefficients « q -shift » pour le filtre passe-bas du premier banc de filtres dans l'algorithme dual-tree, en fixant la taille à 10 coefficients. Ce cas intéressant n'a que 6 coefficients non-nuls.

Notons que la reconstruction est effectuée en moyennant les reconstructions des deux bancs de filtres, ce qui est parfaitement analogue au schéma de la transformée de Fourier, conformément à l'équation (1.2). Les deux banc de filtres correspondent à la partie paire et à la partie impaire des atomes de reconstruction, dont la somme $w_1\psi_1(t) + w_2\psi_2(t)$ permet de reconstituer un atome « pondéré et déphasé » selon l'amplitude $\sqrt{w_1^2 + w_2^2}$ et la phase $\arg\{w_1 + jw_2\}$. Cette mécanique permet de localiser les singularités précisément, contrairement aux décompositions réelles qui impliquent un *ensemble* d'ondelettes dont les oscillations doivent se compenser les unes avec les autres dans un voisinage. En conséquence, un traitement non linéaire comme un seuillage produira beaucoup moins d'artefacts avec des ondelettes complexes. Notons que la définition de la reconstruction est unique, ce qui n'est pas toujours le cas avec des représentations redondantes, qui obligent parfois un choix arbitraire de méthode de recombinaison des coefficients.

On notera également les paires de Hilbert de bases d'ondelettes de type spline qui ont été proposées dans [26]. L'approximation sous-jacente réside dans la discrétisation de la transformée de Hilbert - à l'origine continue - qui correspond à une periodisation de sa réponse fréquentielle $-j \operatorname{sgn}(\omega)$, où ω est alors compris modulo π . Comme la réponse impulsionnelle est infinie, le banc de filtres est réalisé dans le domaine de Fourier, à l'aide de FFT's. Les *phaselets* [58] généralisent également le dual-tree à travers la construction d'ondelettes complexes approximativement in-variantes par translation.

Les ondelettes complexes offrent un compromis intéressant entre l'analyse très fine de Gabor et le codage parcimonieux des bancs de filtres orthogonaux. Elles permettent d'analyser les signaux temporellement et fréquentiellement en termes d'amplitude et de phase, avec une certaine invariance par translation, obtenue par une redondance légère de $2\times$ tout en permettant une reconstruction parfaite et unique. Les applications des ondelettes complexes sont donc exactement les mêmes que pour les ondelettes classiques, avec toutefois un bémol pour la compression sans-pertes, qui ne profitera pas forcément de la redondance³. Cette évolution constitue une claire amélioration des ondelettes classiques, et valide l'idée d'intégrer la notion de signal analytique dans le domaine des ondelettes. L'extension de ces concepts pour des signaux 2D comme les images suscite depuis le début des années 1980 et encore à ce jour une recherche active, décrite dans la section suivante.

1.3 VERS DES SIGNAUX 2D

Depuis une vingtaine d'années, de nombreuses décompositions multi-échelle « géométriques » à reconstruction parfaite ont été proposées pour l'image, afin de mieux prendre en compte les informations structurelles 2D (transformées en « -let » et autres [72]). Ces approches n'utilisent pas le concept de phase, ce dernier étant plus une grandeur physique que géométrique. Étant donné le lien important que nous venons de mettre en avant entre la prise en compte de la phase et la qualité de représentation des ondelettes complexes en 1D, notre travail de recherche tend vers la définition d'une transformée en ondelettes *analytique* pour les images, qui posséderait d'une part les avantages du dual-tree 1D (invariance de l'amplitude, phase liée à la forme du

3. En revanche, la compression *avec* pertes, souvent plus efficace avec des transformées redondantes pour de forts taux de compression, est tout-à-fait envisageable.

signal, pas d'artefacts oscillants), et d'autre part une prise en compte efficace de la géométrie 2D.

Nous présentons dans cette section les travaux qui ont donné naissance aux représentations quaternioniques et monogènes - qui sont au cœur de cette thèse. Il s'agit d'une part de l'analyse directionnelle de Gabor, qui a initié l'utilisation d'une phase en 2D, et d'autre part de la version 2D du dual-tree, qui réalise une décomposition type Gabor dans un contexte de banc de filtres orthogonaux à faible redondance. Ces outils correspondent implicitement à une extension *directionnelle* de la notion de signal analytique, que l'on peut alors utiliser dans plusieurs directions spatiales fixées.

A partir de maintenant, les coordonnées 2D seront notées en caractères gras et correspondront à des vecteurs, comme $\mathbf{x} = [x_1 \ x_2]^\top$ et $\boldsymbol{\omega} = [\omega_1 \ \omega_2]^\top$.

1.3.1 La transformée de Fourier 2D

La version multi-dimensionnelle de la transformée de Fourier est bien définie mathématiquement. Elle décompose les signaux n D dans une base d'*ondes planes* complexes $e^{j\boldsymbol{\omega}^\top \mathbf{x}}$, paramétrées par leur fréquence $\xi = \|\boldsymbol{\omega}\|$ et leur orientation $\theta = \arg\{\omega_1 + j\omega_2\}$. Une onde plane oscille dans la direction θ et reste constante dans les directions perpendiculaires $\theta \pm \pi/2$. La transformée de Fourier 2D du signal $s(\mathbf{x})$ et son inverse s'écrivent :

$$\hat{s}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} s(\mathbf{x}) e^{-j(\mathbf{x}^\top \boldsymbol{\omega})} d\mathbf{x} = A(\boldsymbol{\omega}) e^{j\varphi(\boldsymbol{\omega})} \quad (1.30)$$

$$s(\mathbf{x}) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} \hat{s}(\boldsymbol{\omega}) e^{j(\mathbf{x}^\top \boldsymbol{\omega})} d\boldsymbol{\omega} \quad (1.31)$$

On a dans ce modèle une notion de phase $\varphi(\boldsymbol{\omega})$, qui s'interprète donc comme dans le cas 1D mais selon la direction spatiale θ . Comme en 1D, la restriction aux signaux réels $s(\mathbf{x}) \in \mathbb{R}$ implique une symétrie hermitienne du spectre $s(-\mathbf{x}) = s^*(\mathbf{x})$ qui permet d'écrire une somme d'oscillations *réelles* à partir d'une partie paire et d'une partie impaire :

$$s(\mathbf{x}) = \frac{1}{2\pi^2} \int_{\mathbb{R}^+ \times \mathbb{R}} \Re\{\hat{s}(\boldsymbol{\omega})\} \cos(\mathbf{x}^\top \boldsymbol{\omega}) d\boldsymbol{\omega} - \frac{1}{2\pi^2} \int_{\mathbb{R}^+ \times \mathbb{R}} \Im\{\hat{s}(\boldsymbol{\omega})\} \sin(\mathbf{x}^\top \boldsymbol{\omega}) d\boldsymbol{\omega} \quad (1.32)$$

$$= \frac{1}{2\pi^2} \int_{\mathbb{R}^+ \times \mathbb{R}} A(\boldsymbol{\omega}) \cos(\mathbf{x}^\top \boldsymbol{\omega} + \varphi(\boldsymbol{\omega})) d\boldsymbol{\omega} \quad (1.33)$$

La sommation se fait dans le demi-plan $\mathbb{R}^+ \times \mathbb{R}$ mais pourrait tout aussi bien se faire dans un autre demi plan comme $\mathbb{R} \times \mathbb{R}^+$ par exemple. On remarque à cette occasion que la notion de fréquence négative devient ambiguë en dimension supérieure à 1. D'ailleurs, on parlera plutôt désormais de fréquence *spatiale*, exprimée en nombre de cycles par unité de *longueur* plutôt que par unité de temps. Comme en 1D, l'amplitude est invariante par translation, grâce au théorème du décalage :

$$s(\mathbf{x} - \mathbf{x}_0) \xleftrightarrow{\mathcal{F}} e^{-j\varphi_0(\boldsymbol{\omega})} \hat{s}(\boldsymbol{\omega}) = A(\boldsymbol{\omega}) e^{j(\varphi(\boldsymbol{\omega}) - \varphi_0(\boldsymbol{\omega}))} \quad (1.34)$$

avec $\varphi_0(\boldsymbol{\omega}) = \boldsymbol{\omega}^\top \mathbf{x}_0$

Alors que l'amplitude des coefficients $A(\boldsymbol{\omega})$ donne une information d'énergie fréquentielle dépendante de l'orientation, la phase $\varphi(\boldsymbol{\omega})$ contient la structure globale de l'image, en termes de contours. En effet, la configuration des différentes valeurs de phase dans une somme d'oscillations en traduit les singularités, comme le montre le test très simple fait par exemple dans [114], qui consiste à reconstruire une image à partir de ses coefficients de Fourier après avoir neutralisé l'amplitude ($A(\boldsymbol{\omega}) = 1$) ou la phase ($\varphi(\boldsymbol{\omega}) = 0$). Ce test illustré figure 1.5 confirme que la géométrie de l'image est contenue dans la phase.

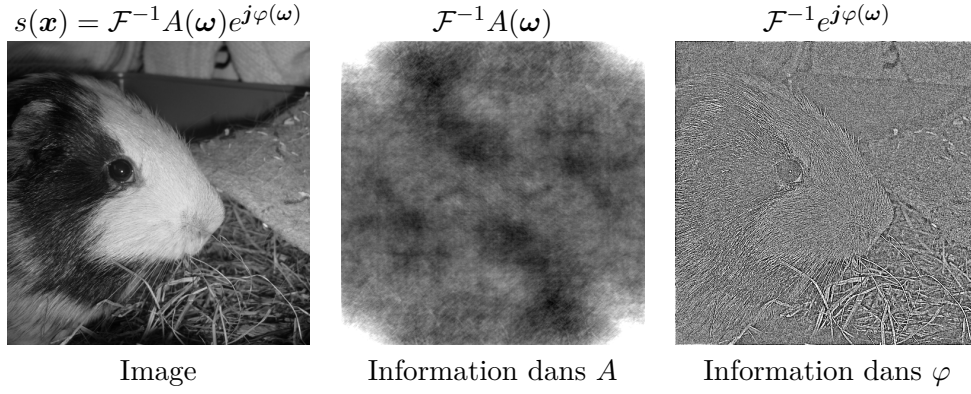


FIGURE 1.5 – Importance de la phase dans Fourier (le contraste des deux images de droite a été renforcé).

Cette généralisation 2D de la transformée de Fourier suggère donc le réemploi de la phase 1D, conjointement à la nouvelle donnée d'*orientation*. C'est naturellement que l'approche multi-résolution s'est tout d'abord faite de façon « directionnelle », c'est-à-dire indépendamment dans plusieurs directions fixées. Avant d'aborder concrètement ces représentations nous allons introduire la *transformée de Hilbert partielle*, qui permet de faire le lien avec le signal analytique 1D.

Transformée de Hilbert partielle

La transformée de Hilbert partielle est définie de façon analogue à une dérivée partielle, et concerne une direction particulière θ :

$$\mathcal{H}_\theta^{\text{part}} s(\mathbf{x}) \xleftrightarrow{\mathcal{F}} -j \operatorname{sgn}(\omega_1 \cos(\theta) + \omega_2 \sin(\theta)) \hat{s}(\boldsymbol{\omega}) \quad (1.35)$$

Sa réponse fréquentielle est identique à la transformée de Hilbert sur la ligne d'orientation θ , et sépare le spectre en deux parties, analogues aux fréquences « positives » et « négatives » du cas 1D. On retiendra en particulier pour la suite les $\mathcal{H}^{\text{part}}$ le long des axes x et y , ainsi que la composée des deux aussi appelée « Hilbert totale » :

$$\mathcal{H}_0^{\text{part}} s(\mathbf{x}) \xleftrightarrow{\mathcal{F}} -j \operatorname{sgn}(\omega_1) \hat{s}(\boldsymbol{\omega}) \quad \begin{array}{c} \uparrow \omega_2 \\ \begin{array}{|c|c|} \hline j & -j \\ \hline j & -j \\ \hline \end{array} \rightarrow \omega_1 \end{array} \quad (1.36)$$

$$\mathcal{H}_{\pi/2}^{\text{part}} s(\mathbf{x}) \xleftrightarrow{\mathcal{F}} -j \operatorname{sgn}(\omega_2) \hat{s}(\boldsymbol{\omega}) \quad \begin{array}{c} \uparrow \omega_2 \\ \begin{array}{|c|c|} \hline -j & -j \\ \hline j & j \\ \hline \end{array} \rightarrow \omega_1 \end{array} \quad (1.37)$$

$$\mathcal{H}_0^{\text{part}} \mathcal{H}_{\pi/2}^{\text{part}} s(\mathbf{x}) \xleftrightarrow{\mathcal{F}} -\operatorname{sgn}(\omega_1) \operatorname{sgn}(\omega_2) \hat{s}(\boldsymbol{\omega}) \quad \begin{array}{c} \uparrow \omega_2 \\ \begin{array}{|c|c|} \hline 1 & -1 \\ \hline -1 & 1 \\ \hline \end{array} \rightarrow \omega_1 \end{array} \quad (1.38)$$

Les schémas de droite indiquent la réponse fréquentielle associée. D'un point de vue numérique, ces trois opérateurs s'implantent en combinant les transformées de Hilbert des lignes et des colonnes de l'image.

Cette transformée relie les parties réelle/imaginaire des atomes de Fourier 2D de la façon suivante :

$$\Im \{e^{j\boldsymbol{\omega}^\top \mathbf{x}}\} = \mathcal{H}_\theta^{\text{part}} \Re \{e^{j\boldsymbol{\omega}^\top \mathbf{x}}\} \quad \text{si } \boldsymbol{\omega} = \xi [\cos(\theta) \quad \sin(\theta)]^\top \quad (1.39)$$

La transformée de Hilbert partielle doit être paramétrée dans la même direction que l'onde plane pour vérifier cette équation. Par rapport au cas 1D, on n'a plus une unique transformée qui relie l'ensemble des atomes de Fourier, mais une *famille* de transformées orientées, qui transforment les ondes planes de type *cosinus* en ondes planes de type *sinus*.

1.3.2 Ondelettes de Gabor directionnelles

Les décompositions directionnelles pour l'analyse d'images ont été initiées probablement par G. Granlund en 1978 sous forme d'un fenêtrage gaussien dans le domaine de Fourier 2D [60], et poursuivies avec les filtres en quadrature⁴ 2D de H. Knutsson [81] ainsi que les ondelettes de Gabor directionnelles proposée par J. Daugman [34]. Ces dernières sont en accord avec une certaine modélisation des « cellules simples » du cortex visuel primaire, qui détectent des motifs d'intensité lumineuse en fonctions de leurs position, orientation et échelle [70]. C'est pourquoi l'utilisation des ondelettes de Gabor pour l'image est souvent liée à la prise en compte du système visuel humain.

Les ondelettes de Gabor sont construites par fenêtrage gaussien isotrope d'une onde plane complexe de fréquence ξ dans la direction θ :

$$\psi^\theta(\mathbf{x}) = \frac{e^{-\|\mathbf{x}\|^2/2}}{2\pi} e^{-j(\mathbf{x}^\top \boldsymbol{\omega}_0)} \quad \text{avec} \quad \boldsymbol{\omega}_0 = \xi[\cos(\theta) ; \sin(\theta)]^\top \quad (1.40)$$

La décomposition dépend du nombre d'orientations K que l'on se fixe, et que l'on peut répartir uniformément dans $[0; \pi[$:

$$\theta \in \Theta = \left\{ \frac{k\pi}{K} ; 0 \leq k < K \right\} \quad (1.41)$$

On peut alors décomposer tout signal 2D réel $s(\mathbf{x})$ par produit scalaire avec les atomes suivants :

$$\left\{ \psi_{j,u}^\theta(\mathbf{x}) = 2^{-j} \psi^\theta \left(2^{-j}(\mathbf{x} - \mathbf{u}) \right) \right\}_{\theta \in \Theta, j \in \mathbb{Z}, \mathbf{u} \in \mathbb{R}^2} \quad (1.42)$$

La redondance de cette représentation correspond au nombre d'échelles fois le nombre d'orientations, et comme en 1D, la discrétisation rend la reconstruction inexacte en pratique.

La décomposition se calcule en filtrant le signal par une famille de *filtres de Gabor* dont la réponse impulsionnelle est $\psi_{j,0}^\theta(\mathbf{x})$. Ces filtres sont approximativement *en quadrature* [61], ou « analytiques » au sens de la transformée de Hilbert partielle, et donnent donc accès à une amplitude et une phase locale, pour chaque valeur de θ .

Cette décomposition fait référence encore aujourd'hui en analyse d'image, notamment en segmentation de texture et en détection de contours, pour des applications où la taille des données et le temps de calcul ne sont pas trop contraints, et où un certain lien avec le système visuel humain est désiré.

L'orientation des filtres ne coïncide pas nécessairement avec celle des l'orientation réelle des éléments de l'image, ce qui peut introduire un biais dans l'interprétation de la phase. Une étape supplémentaire d'adaptation locale de l'orientation des filtres sera donc parfois nécessaire. Ce genre de représentation fournit un grand nombre d'informations, et manquera parfois de synthèse dans une perspective d'analyse des structures géométriques. En revanche, celà en fait un outil de choix pour une caractérisation riche des textures. Le problème majeur réside dans les difficultés de discrétisation de la transformée, qui empêchent de mettre en place un schéma numérique satisfaisant.

Dans la littérature, des filtres de Gabor ainsi que la notion de transformée de Hilbert partielle ont été utilisés dans une problématique de *démodulation* des images [11]. Il s'agit de voir une image comme une somme d'ondes planes modulées en amplitude et en fréquence (modèle AM-FM). Pour l'analyse des textures, l'amplitude est utilisée comme une information de contraste et la fréquence comme une mesure de la « granularité » [12]. La phase de Gabor est quant à elle classiquement utilisée pour l'estimation de flot optique [51], par utilisation du théorème du décalage qui est approximativement vérifié (nous y reviendrons à la section 2.3.1).

On peut citer un travail substantiel autour du concept de phase, qui a été mené à travers la notion de *congruence de phase*, introduite en 1987 dans [108] comme un détecteur de contours.

4. En dimension 2, le terme « en quadrature » n'est pas clairement lié à une notion 2D de signal analytique, mais désigne la plupart du temps des filtres construits par la transformée de Hilbert *partielle* dans une direction fixée.

Un signal présente une forte congruence de phase en un point si les sinusôides qui le composent y sont *en phase*. Ces endroits du signal portent une énergie relativement forte, et exhibent une forme localement sinusôidale, dont la valeur de phase indique s'il s'agit plutôt d'une « bosse », d'un « creux » ou d'une « pente » (cf. table 1.1). Il vient que la mesure de congruence de phase est en fait équivalente à l'*amplitude instantanée* $A(t)$, que les auteurs appellent *énergie locale*. Pour étendre ce concept en 2D, les auteurs proposent d'appliquer cette méthode dans plusieurs directions, via une décomposition de Gabor. A nouveau, la difficulté est dans la recombinaison des informations obtenues dans les différentes orientations. Les valeurs de phase aux points de congruence permettent de classifier les contours dans un modèle continu entre la « ligne » et le « bord », qui est présenté comme judicieux vis-à-vis du système visuel humain [108], apparemment sensible à la congruence de phase. De plus, l'information de phase est invariante par changement d'intensité.

Finalement, la transformée de Gabor 2D permet une analyse riche et flexible mais nécessite une paramétrisation des orientations utilisées et souffre d'un manque de définition discrète, exactement comme en 1D. Notons que les transformées « géométriques » type Curvelets et Contourlets [72] offrent une analyse également très orientée avec redondance raisonnable et reconstruction parfaite. La notion de phase n'est pas recherchée dans ces approches.

1.3.3 Ondelettes orthogonales séparables

Nous présentons ici la célèbre extension séparable de l'algorithme de S. Mallat [101], construite par produit tensoriel de deux représentations orthogonales 1D, et réalisée par l'application du banc de filtres 1D sur les lignes et colonnes de l'image. Cet outil fait référence, en pratique, autour des ondelettes pour l'image, et constitue un point de départ pour notre travail, qui vise à lui trouver une meilleure alternative. Contrairement à la transformée de Gabor présentée plus haut, les coefficients ici sont réels.

A partir d'une décomposition orthogonales ou bi-orthogonale 1D associée à la fonction d'échelle $\phi(t)$, à l'ondelette $\psi(t)$, et aux filtres $h[n]$ et $g[n]$, on peut définir une décomposition 2D de façon séparable à travers une fonction d'échelle et *trois* fonctions d'ondelettes 2D construites par produit tensoriel :

$$\phi_{LL}(\mathbf{x}) = \phi(x_1)\phi(x_2) \quad \phi_{LL,j,\mathbf{k}} = 2^j \phi_{LL}(2^j \mathbf{x} - \mathbf{k}) \quad (1.43)$$

$$\psi_{LH}(\mathbf{x}) = \phi(x_1)\psi(x_2) \quad \psi_{LH,j,\mathbf{k}} = 2^j \psi_{LH}(2^j \mathbf{x} - \mathbf{k}) \quad (1.44)$$

$$\psi_{HL}(\mathbf{x}) = \psi(x_1)\phi(x_2) \quad \psi_{HL,j,\mathbf{k}} = 2^j \psi_{HL}(2^j \mathbf{x} - \mathbf{k}) \quad (1.45)$$

$$\psi_{HH}(\mathbf{x}) = \psi(x_1)\psi(x_2) \quad \psi_{HH,j,\mathbf{k}} = 2^j \psi_{HH}(2^j \mathbf{x} - \mathbf{k}) \quad (1.46)$$

La base de décomposition associée est :

$$\{ \phi_{LL,j,\mathbf{k}} , \psi_{LH,j,\mathbf{k}} , \psi_{HL,j,\mathbf{k}} , \psi_{HH,j,\mathbf{k}} \}_{j \in \mathbb{Z}, \mathbf{k} \in \mathbb{Z}^2} \quad (1.47)$$

et la reconstruction du signal satisfait :

$$s(\mathbf{x}) = \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{N-1} w_{LL,j}[\mathbf{k}] \phi_{LL,j,\mathbf{k}} + \sum_{j=J}^0 \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{N-1} (w_{LH,j}[\mathbf{k}] \psi_{LH,j,\mathbf{k}} + w_{HL,j}[\mathbf{k}] \psi_{HL,j,\mathbf{k}} + w_{HH,j}[\mathbf{k}] \psi_{HH,j,\mathbf{k}}) \quad (1.48)$$

Le premier avantage de cette définition est la simplicité mathématique, puisqu'on conserve les propriétés classiques d'orthogonalité, de régularité *etc.* D'un point de vue algorithmique, le calcul des coefficients peut se faire de façon efficace avec un banc de filtres 1D que l'on applique sur les lignes et les colonnes de l'image. Cet algorithme est très utilisé dans la littérature notamment sur le codage d'image en général, le débruitage et la classification de textures. Nous donnons une illustration de cette décomposition à la figure 1.6. Comme en 1D, la transformée en ondelettes séparable offre une représentation parcimonieuse de l'image car la plupart des coefficients sont

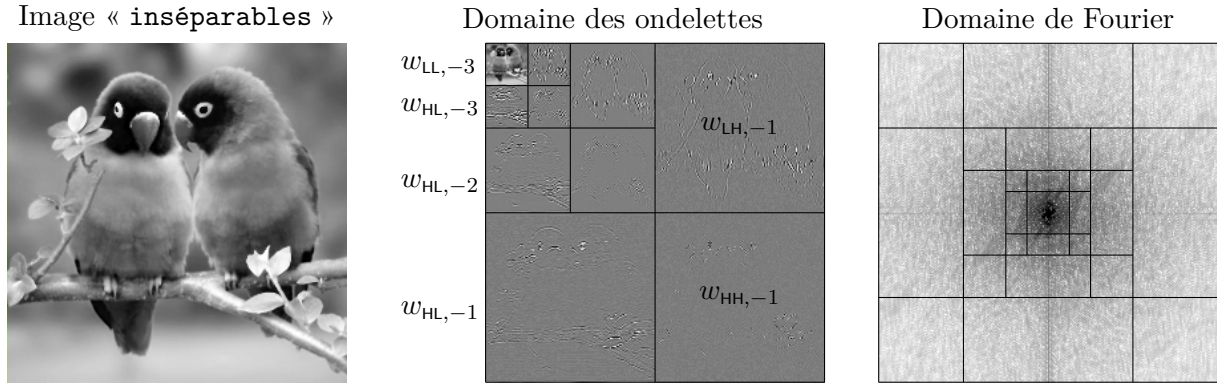


FIGURE 1.6 – Représentation en ondelettes séparables et découpage de spectre associé.

très faibles, et donc ne portent pas d'information. L'orthogonalité de la base permet de ne pas introduire de corrélation dans la représentation d'un bruit (et de le « blanchir »), ce qui est utile en restauration d'image, mais aussi d'envisager le *codage* des coefficients sous un angle statistique pour optimiser leur quantification.

En ce qui concerne les désavantages, au problème de *variance par translation* vient s'ajouter un problème de mauvaise directionnalité. En effet, la méthode séparable résulte en un découpage fréquentiel rectangulaire très arbitraire (voir figure 1.6), qui correspond à une classification des détails de l'image en trois orientations globales : horizontale, verticale, et les deux diagonales mélangées. Il vient que cette représentation est efficace pour représenter les structures géométriques « impulsives », mais mal adaptée pour la représentation des *contours*, objets fondamentaux en images, qui ont en général des orientations diverses.

Les ondelettes orthogonales *non-séparables* permettent de résoudre certains de ces problèmes [76, 30, 85, 4, 46], mais ne correspondent pas à des décompositions analytiques - nous ne les traiterons pas ici.

Nous avons vu en 1D que l'utilisation de bancs de filtres complexes peu redondants comme le dual-tree permettait d'obtenir une quasi-invariance par translation et une interprétation physique des coefficients en termes d'amplitude et de phase. C'est pourquoi nous étudions maintenant l'extension 2D du dual-tree, qui a ouvert la voie des bancs de filtres 2D analytiques peu redondants, et qui a soulevé le manque de définitions fondamentales en signal pour l'image.

1.3.4 Dual-tree 2D

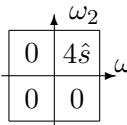
Face aux limites désormais classiques des ondelettes biorthogonales séparables (variance par translation, mauvaise directionnalité, peu adapté aux structures géométriques), et sachant l'efficacité du dual-tree due à l'intégration du signal analytique, il paraît naturel d'étudier le dual-tree 2D [124].

En dépit de la mauvaise réputation des schémas séparables, l'extension 2D du dual-tree est réalisée par un produit tensoriel. Reprenons les fonctions de base du schéma séparable ϕ_{LL} , ψ_{HL} , ψ_{LH} et ψ_{HH} ; et considérons désormais que les fonctions 1D ϕ et ψ sur lesquelles elles sont construites sont *analytiques*, et de la forme $\phi(t) = \phi_1(t) + j\phi_2(t)$ et $\psi(t) = \psi_1(t) + j\psi_2(t)$. On obtient des fonctions 2D complexes de la forme suivante (ici pour le cas de ψ_{HH}) :

$$\psi_{HH}(\mathbf{x}) = \psi(x_1)\psi(x_2) = \psi_1(x_1)\psi_1(x_2) - \psi_2(x_1)\psi_2(x_2) + j(\psi_2(x_1)\psi_1(x_2) + \psi_1(x_1)\psi_2(x_2)) \quad (1.49)$$

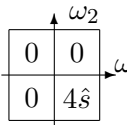
Du fait de la transformée de Hilbert entre ψ_1 et ψ_2 (voir dual-tree 1D, section 1.2.3), cette nouvelle ondelette complexe 2D correspond à un signal analytique au sens de Hahn [63], construit à partir de l'ondelette réelle $\psi_1(x_1)\psi_1(x_2)$, qui consiste à annuler les fréquences négatives

dans trois quadrant du spectre de Fourier ($\omega_1, \omega_2 < 0$) :

$$s_{A_1}(\mathbf{x}) = s(\mathbf{x}) - \overset{\text{part}}{\mathcal{H}_0} \overset{\text{part}}{\mathcal{H}_{\frac{\pi}{2}}} s(\mathbf{x}) + \mathbf{j} \left(\overset{\text{part}}{\mathcal{H}_0} s(\mathbf{x}) + \overset{\text{part}}{\mathcal{H}_{\frac{\pi}{2}}} s(\mathbf{x}) \right) \quad (1.50)$$


Il s'agit donc d'une ondelette dont le support fréquentiel est un carré dyadique dans le premier quadrant. En conséquence, cette ondelette ne concerne qu'une seule des deux diagonales (45°), ce qui est un premier pas vers une meilleure directionnalité.

La définition 2D du signal analytique de Hahn pose cependant un problème important : le signal de départ ne peut pas être retrouvé à partir de sa représentation analytique. Dans notre cas, cela se traduit pas « la perte d'une diagonale ». La solution de Hahn est de considérer un deuxième signal analytique correspondant au quadrant voisin :

$$s_{A_2}(\mathbf{x}) = s(\mathbf{x}) + \overset{\text{part}}{\mathcal{H}_0} \overset{\text{part}}{\mathcal{H}_{\frac{\pi}{2}}} s(\mathbf{x}) + \mathbf{j} \left(\overset{\text{part}}{\mathcal{H}_0} s(\mathbf{x}) - \overset{\text{part}}{\mathcal{H}_{\frac{\pi}{2}}} s(\mathbf{x}) \right) \quad (1.51)$$


Dans le cas de nos ondelettes, celà correspond à utiliser une seconde ondelette :

$$\psi_{\text{HH}_2} = \psi(x)\psi^*(y) = \psi_1(x)\psi_1(y) + \psi_2(x)\psi_2(y) + \mathbf{j}(\psi_2(x)\psi_1(y) - \psi_1(x)\psi_2(y)) \quad (1.52)$$

correspondant donc au quadrant voisin et à la deuxième diagonale (orientation 135°).

L'application du même raisonnement avec les deux autres ondelettes 2D aboutit à 6 ondelettes 2D complexes, orientées dans 6 directions différentes :

$$\psi_{\text{LH}_1} = \phi_1(x_1)\psi_1(x_2) - \phi_2(x_1)\psi_2(x_2) + \mathbf{j}(\phi_2(x_1)\psi_1(x_2) + \phi_1(x_1)\psi_2(x_2)) \quad \text{vers } 15^\circ \quad (1.53)$$

$$\psi_{\text{HH}_1} = \psi_1(x_1)\psi_1(x_2) - \psi_2(x_1)\psi_2(x_2) + \mathbf{j}(\psi_2(x_1)\psi_1(x_2) + \psi_1(x_1)\psi_2(x_2)) \quad \text{vers } 45^\circ \quad (1.54)$$

$$\psi_{\text{HL}_1} = \psi_1(x_1)\phi_1(x_2) - \psi_2(x_1)\phi_2(x_2) + \mathbf{j}(\psi_2(x_1)\phi_1(x_2) + \psi_1(x_1)\phi_2(x_2)) \quad \text{vers } 75^\circ \quad (1.55)$$

$$\psi_{\text{HL}_2} = \psi_1(x_1)\phi_1(x_2) + \psi_2(x_1)\phi_2(x_2) + \mathbf{j}(\psi_2(x_1)\phi_1(x_2) - \psi_1(x_1)\phi_2(x_2)) \quad \text{vers } 105^\circ \quad (1.56)$$

$$\psi_{\text{HH}_2} = \psi_1(x_1)\psi_1(x_2) + \psi_2(x_1)\psi_2(x_2) + \mathbf{j}(\psi_2(x_1)\psi_1(x_2) - \psi_1(x_1)\psi_2(x_2)) \quad \text{vers } 135^\circ \quad (1.57)$$

$$\psi_{\text{LH}_2} = \phi_1(x_1)\psi_1(x_2) + \phi_2(x_1)\psi_2(x_2) + \mathbf{j}(\phi_2(x_1)\psi_1(x_2) - \phi_1(x_1)\psi_2(x_2)) \quad \text{vers } 165^\circ \quad (1.58)$$

La décomposition séparable complexe à partir du dual-tree réalise donc une analyse orientée « à la Gabor », dans les 6 directions $(1+2k) \times 15^\circ$.

La réalisation par banc de filtres est en filigrane dans les équations (1.53) à (1.58), à savoir qu'il suffit de recombinaison par sommes et soustractions les sorties de 4 bancs de filtres séparables définis par produit tensoriels des parties réelles et imaginaires du dual-tree 1D. Le schéma de décomposition complet est illustré à la figure 1.7. L'invariance par translation approximative est obtenue pour une redondance de $4\times$, et la phase des coefficients peut s'utiliser comme dans une décomposition de Gabor. Finalement, cette extension a permis d'améliorer la directionnalité et l'invariance des ondelettes classiques, grâce à l'intégration d'ondelettes *complexes*. Le type de représentation obtenu n'est pas nouveau, puisqu'il s'agit d'une analyse type Gabor avec 6 orientations, mais à son avantage la reconstruction est clairement définie, comme la moyenne des reconstructions des quatre bancs de filtres. Cette méthode de reconstruction revient à une modélisation de la structure locale de l'image comme une somme de 6 ondelettes orientées vers $(1+2k)15^\circ$, « pondérées et déphasées » par les coefficients de la décomposition. Ce modèle englobe bien-sûr les structures de type « bord » ou « ligne », pour lesquelles un seul des 6 coefficients complexes sera significatif (si l'orientation du contour correspond à une des orientations privilégiées). Mais de nombreuses autres structures sont implicitement prises en compte, comme des croisements et des coins par exemple.

Cependant, cette décomposition est liée à une définition ambiguë de signal analytique 2D, nécessitant deux signaux complexes pour une seule fonction de base. De plus, les 6 directions d'analyse sont rigides et peuvent ne pas être alignées avec les orientations réelles des structures de l'image.

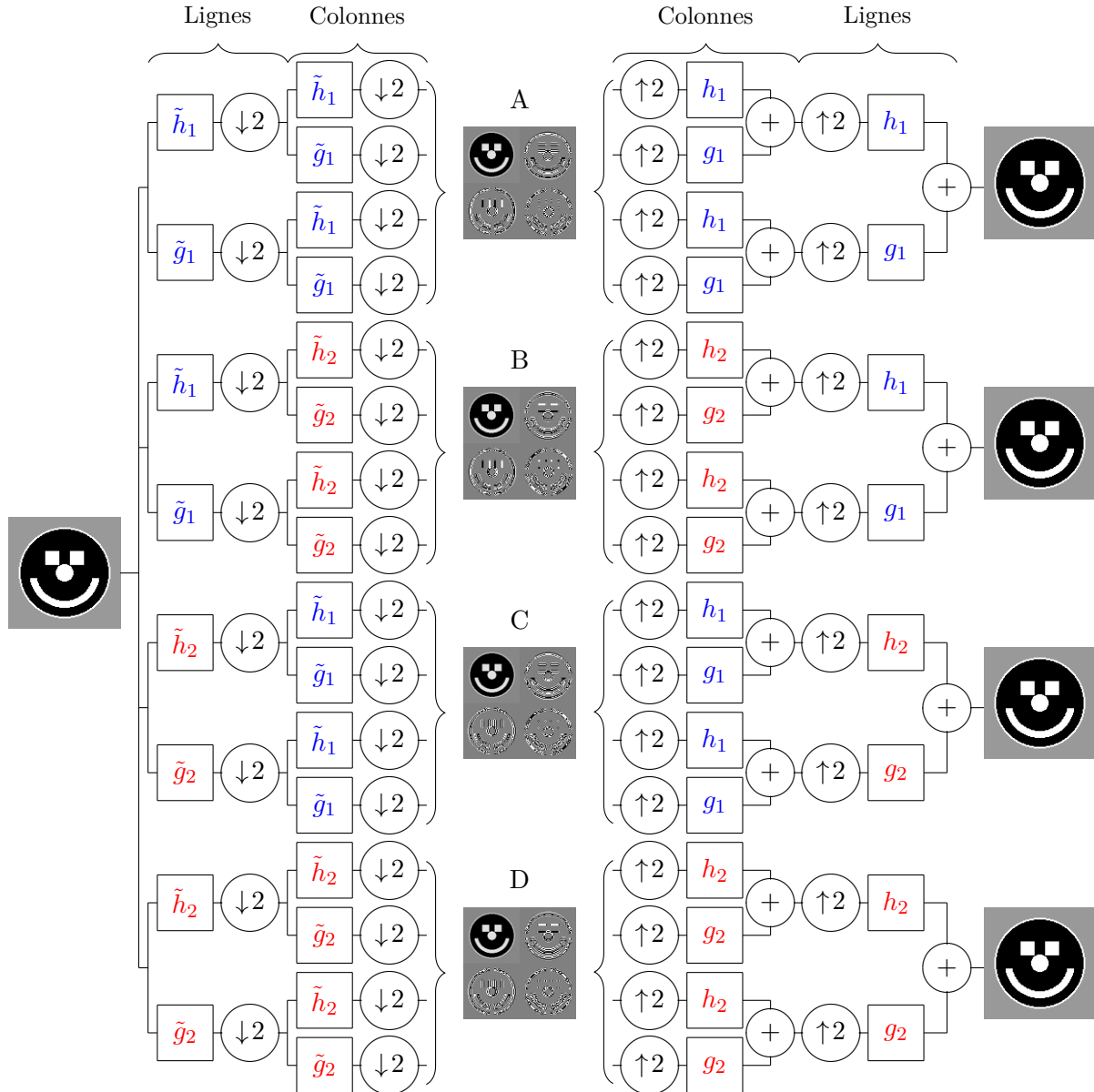


FIGURE 1.7 – Implantation séparable du banc de filtres dual-tree. Pour construire les sous-bandes complexes, les sorties des quatre bancs de filtres 2D doivent être combinées en deux sorties complexes $(A + D) + j(B - C)$ et $(A - D) + j(C - B)$.

1.4 RÉSUMÉ ET DISCUSSION

Nous avons présenté dans ce chapitre des outils d'analyse qui traduisent tout signal comme une somme d'oscillations, en termes d'amplitude, phase et fréquence. A la base de notre travail se trouvent les représentations en ondelettes analytiques peu redondantes et à reconstruction parfaite. Ces outils offrent une représentation *localisée* dans l'espace et dans les fréquences/échelles à partir d'une séquence *compacte* de coefficients dont les valeurs définissent *directement* les formes oscillantes en termes d'amplitude, phase et fréquence.

En 1D, l'algorithme dual-tree est présenté comme un très bon compromis entre :

- les bancs de filtres orthogonaux qui sont non-redondants et inversibles mais qui manquent de flexibilité car ils gèrent mal la directionnalité et sont instables par translation ;
- les outils d'analyse plus fins comme l'analyse de Gabor qui permettent d'extraire des données stables et physiquement interprétables, mais qui sont trop redondants et se dicrétisent mal.

On s'intéresse donc à une famille de représentations suffisamment compactes et cohérentes à la fois, tout en donnant accès directement aux informations d'amplitude, phase et fréquence.

En 2D, nous avons vu que l'extension des représentations analytiques reste un sujet ouvert. La première solution a été donnée au début des années 80 avec la transformée de Gabor 2D. L'idée est de décomposer le signal en une séquences de « sous-signaux à orientation unique » - par filtrages anisotropes - dont on peut ensuite facilement étudier l'amplitude et la phase avec les outils 1D. Ces représentations permettent une détection de contours de qualité grâce à la phase qui intègre la forme du contour (« plutôt ligne » ou « plutôt bord »), indépendamment de l'amplitude qui présentera un maximum au centre du contour. La proximité des filtres de Gabor avec le système visuel humain en fait un outil intéressant pour les applications où la perception est à prendre en compte.

Néanmoins, pour une analyse fine, on doit utiliser de nombreuses orientations, ce qui augmente considérablement la taille des données, et rend confuse l'interprétation de la phase qui se retrouve, pour une échelle fixée, représentée par autant d'angles « φ » qu'il y a d'orientations. Avec peu d'orientations, l'analyse est concise et compacte, mais l'interprétation de la phase est plus incertaine car l'orientation locale sous-jacente est très approximative. De plus, l'aspect numérique, notamment à travers la notion de reconstruction parfaite, peut être mal défini. L'extension 2D du dual-tree réalise le même genre d'analyse avec un nombre d'orientations fixe et une redondance raisonnable.

Mais la représentation « 2D dual-tree CWT » doit faire face à la forte concurrence des transformées dites « géométriques » [72], qui n'offrent certes pas de notion de phase⁵, mais qui prennent bien en compte la géométrie locale de l'image (curvelets, bandlets, ondelettes non-séparables, représentations adaptatives avec apprentissage de dictionnaire *etc.*). Il apparaît que la notion de signal analytique « à la Gabor », c'est-à-dire dans une série de directions fixées, n'est pas suffisante pour l'image. Depuis une petite quinzaine d'années, deux définitions majeures mieux adaptées à l'image dans le cadre du signal analytique, ont été proposées : le formalisme *quaternionique* de T. Bülow et le formalisme *monogène* de M. Felsberg. Ces nouveaux outils partent d'une réflexion de fond sur l'approche *signal* des images, et proposent une extension complète du concept de phase. La particularité de ces approches est de sortir de l'algèbre des nombres complexes, qui apparaît trop limitée pour étendre correctement les notions oscillatoires en dimension 2.

L'utilisation d'algèbres plus générales que $\mathbb{C}$ permet de manipuler les signaux multidimensionnels avec plus de flexibilité, et ouvre la voie vers les signaux *vectoriels* comme les *images couleur*, ou même les images multi-/hyper- spectrales, dont les pixels sont associés non plus à un nombre réel mais à un vecteur, comme les trois intensités des canaux rouge, vert,

5. Sauf pour les pyramides orientables de Freeman et Adelson dans le cas de filtres en quadratures [54], nous y reviendrons à la section 3.1.3. Notons également que des curvelets complexes ont été appliquées à l'étude d'ondes sismiques [110].

bleu. En effet, la littérature sur la représentation des images couleur se résume à quelques articles isolés, et en pratique ces images sont traitées de façon sous-optimale car il n'existe pas de représentation multi-échelle bien établie pour la décomposition de signaux vectoriels.

Nous proposons maintenant de nous concentrer sur les modèles quaternionique et monogène, qui nous semblent prometteurs, en tant qu'outils « signal » dédiés à la représentation des images et offrant de potentielles extensions à la couleur.

APPROCHES QUATERNIONIQUES

Ce chapitre concerne la modélisation des signaux 2D à travers le formalisme quaternionique de Thomas Bülow [15], et les ondelettes quaternioniques à reconstruction parfaite (QWT) de Chan *et al.* qui lui ont succédé [24]. Ces dernières constituaient encore il y a deux ans la seule décomposition analytique associée à une réflexion poussée sur le concept de phase 2D et permettant de reconstruire l'image à partir d'un ensemble raisonnablement redondant de coefficients. Les références associées à [15] concernent la redéfinition des représentations de Fourier et de Gabor dans l'algèbre des quaternions $\mathbb{H}$, comme solution de dépassement des limites de l'algèbre complexe $\mathbb{C}$, avec toutefois une absence de dimension numérique et algorithmique. L'aspect banc de filtres à reconstruction parfaite qui a été apporté dans [24] avec la QWT comble ce défaut, mais souffre encore aujourd'hui à notre sens d'un manque d'analyse, qui explique peut-être que peu de travaux ont été proposés pour l'application de cette transformée. C'est pourquoi une partie substantielle du travail de cette thèse a été l'étude approfondie de ce cadre, pour aboutir à une modélisation des données construites par la QWT. Cette analyse a été illustrée à travers deux travaux expérimentaux qui montrent que la QWT est avantageuse par rapport aux ondelettes bi-orthogonales séparables. La première proposition est d'appliquer la parcimonie de la QWT dans un contexte de codage d'images, qui constitue la seule application de la littérature à utiliser la reconstruction de la transformée. La seconde concerne la classification de textures, application très classique des ondelettes.

Après avoir présenté l'algèbre de quaternions et quelques travaux connexes sur les images couleur qui ne seront pas traités d'avantage dans cette thèse, nous développons le formalisme quaternionique avec le même regard que dans le chapitre précédent, c'est-à-dire en se concentrant sur les atomes de décompositions, et sur la compréhension des données amplitude/phase. Dans un second temps, nous présenterons les deux études expérimentales que nous avons conduites pour mettre en lumière l'apport de la QWT sur les ondelettes bi-orthogonales séparables (notées « DWT »).

2.1 IMAGE ET QUATERNIONS

2.1.1 Algèbre des quaternions

L'algèbre des quaternions $\mathbb{H}$ proposée par W. Hamilton en 1844 [65], et relativement connue pour la manipulation des rotations 3D, généralise celle des nombres complexes $\mathbb{C}$. On se donne trois nombres imaginaires $\mathbf{j}_1$, $\mathbf{j}_2$ et $\mathbf{j}_3$ tels que $\mathbf{j}_1\mathbf{j}_2 = -\mathbf{j}_2\mathbf{j}_1 = \mathbf{j}_3$ et $\mathbf{j}_1^2 = \mathbf{j}_2^2 = \mathbf{j}_3^2 = -1$, et un

quaternion s'écrit :

$$q = a + \mathbf{j}_1 b + \mathbf{j}_2 c + \mathbf{j}_3 d \quad \text{avec} \quad a, b, c, d \in \mathbb{R} \quad (2.1)$$

L'addition et la multiplication sont analogues au cas complexe, à ceci près que la multiplication n'est *pas commutative*. La norme d'un quaternion est définie par $|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$. La représentation exponentielle d'un quaternion quelconque est définie par :

$$q = |q|e^{u\theta} \quad \text{avec} \quad \mathbf{u} = \frac{\mathbf{j}_1 b + \mathbf{j}_2 c + \mathbf{j}_3 d}{\sqrt{b^2 + c^2 + d^2}} \quad \text{et} \quad \theta = \text{atan} \frac{\sqrt{b^2 + c^2 + d^2}}{a} \quad (2.2)$$

Un quaternion unitaire ($|q| = 1$) et « imaginaire pur » ($a = 0$) noté $\mathbf{u}$ vérifie la formule d'Euler $e^{u\theta} = \cos \theta + \mathbf{u} \sin \theta$. Il généralise ainsi la notion classique de nombre imaginaire complexe. Par contre, la propriété usuelle $e^{u_1 \theta_1} e^{u_2 \theta_2} = e^{u_1 \theta_1 + u_2 \theta_2}$ n'est pas toujours vérifiée (elle l'est en particulier si $\mathbf{u}_1 = \mathbf{u}_2$). Dans un but de généralisation des outils de traitement du signal, la question de la forme exponentielle est fondamentale, puisqu'elle doit porter les informations d'amplitude et de phase.

Une forme dite « polaire » peut être définie par $q = |q|e^{u_1 \varphi} e^{u_2 \theta} e^{u_3 \psi}$, où les $\mathbf{u}_i$ sont des quaternions imaginaires purs unitaires. La norme est unique, mais les angles φ , θ et ψ dépendent de la base des $\mathbf{u}_i$ choisie. Une méthode de calcul inspirée des angles d'Euler consiste à choisir $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = (\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3)$ [15] :

$$q = |q|e^{j_1 \varphi} e^{j_2 \theta} e^{j_3 \psi} \quad \text{avec} \quad (\varphi, \theta, \psi) \in [-\pi, \pi[\times \left[\frac{-\pi}{2}, \frac{\pi}{2} \right[\times \left[\frac{-\pi}{4}, \frac{\pi}{4} \right] \quad (2.3)$$

Les composantes Cartésiennes de q sont alors définies par :

$$a = |q|(\cos \varphi \cos \psi \cos \theta + \sin \varphi \sin \psi \sin \theta) \quad (2.4)$$

$$b = |q|(-\cos \varphi \sin \psi \sin \theta + \sin \varphi \cos \psi \cos \theta) \quad (2.5)$$

$$c = |q|(\cos \varphi \cos \psi \sin \theta - \sin \varphi \sin \psi \cos \theta) \quad (2.6)$$

$$d = |q|(\cos \varphi \sin \psi \cos \theta + \sin \varphi \cos \psi \sin \theta) \quad (2.7)$$

On appellera $|q|$ le « module » de q et (φ, θ, ψ) les « angles de phase ». Le calcul des angles est procédural. Il faut d'abord calculer la norme $|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$, et normaliser les composantes : $(a, b, c, d) \leftarrow (a, b, c, d)/|q|$. Puis on calcule l'angle $\psi = -\frac{1}{2} \arcsin(2(bc - ad))$. Dans le cas régulier $\psi \in]-\frac{\pi}{4}, \frac{\pi}{4}[$, on a :

$$\varphi = \frac{1}{2} \arg\{a^2 - b^2 + c^2 - d^2 + \mathbf{j}2(cd + ab)\} \quad (2.8)$$

$$\theta = \frac{1}{2} \arg\{a^2 + b^2 - c^2 - d^2 + \mathbf{j}2(bd + ac)\} \quad (2.9)$$

Dans le cas particulier $\psi = \pm \frac{\pi}{4}$, φ et θ ne sont pas uniques (seule la quantité $(\varphi - \theta)$ l'est). On peut choisir par exemple $\varphi = 0$ et $\theta = \frac{1}{2} \arg\{a^2 - b^2 - c^2 + d^2 + \mathbf{j}2(ac - bd)\}$. On peut aussi simplement conserver les formules du cas régulier, mais toute interprétation des données devra considérer leur caractère singulier.

Finalement, on doit vérifier $q = e^{j_1 \varphi} e^{j_3 \psi} e^{j_2 \theta}$ en recalculant (a, b, c, d) comme indiqué plus haut, et si $q = -e^{j_1 \varphi} e^{j_3 \psi} e^{j_2 \theta}$, alors on décale φ : $\varphi \leftarrow \varphi \pm \pi$; de façon à garder $\varphi \in [-\pi, \pi[$. Cette forme polaire sera celle utilisée dans ce chapitre.

Une autre représentation polaire est celle de Cayley-Dickson, qui considère un « module complexe » $A_q = (a_0 + a_1 \mathbf{j}_1)$ et une phase complexe $B_q = (b_0 + b_1 \mathbf{j}_1)$, pour définir $q = A_q e^{B_q \mathbf{j}_2}$. On obtient :

$$A_q = \cos(\sqrt{c^2 + d^2})^{-1}(a + \mathbf{j}_1 b) \quad (2.10)$$

$$B_q = \arctan(\sqrt{c^2 + d^2})\sqrt{c^2 + d^2} \left(\frac{ad + bc}{a^2 + b^2} + \mathbf{j}_1 \frac{ad - bc}{a^2 + b^2} \right) \quad (2.11)$$

Cette forme a été utilisée dans [93], dont nous parlerons à la section 4.2.2, mais également pour la prise en compte de la dimension *couleur*.

2.1.2 Outils quaternioniques pour les images couleur

Avant de commencer avec les représentations quaternioniques de T. Bülöw pour les images en niveaux de gris, une ambiguïté doit être levée quant à l'utilisation des quaternions pour l'image. En effet, une série de travaux indépendants de ceux que nous allons présenter, initiés par S. Sangwine *et al.* [122, 115, 45] et poursuivis ensuite dans notre équipe [59, 41, 117, 118, 20], ont étudié la définition d'une transformée de Fourier quaternionique pour les images *couleur*.

Le but est de définir un outil vectoriel qui manipule la couleur comme une entité, afin de développer des traitements couleur dits « non marginaux ». L'idée de départ est de stocker les trois composantes - rouge, vert, bleu - dans les trois parties imaginaires d'un quaternion : $s(\mathbf{x}) = s_R(\mathbf{x})\mathbf{j}_1 + s_G(\mathbf{x})\mathbf{j}_2 + s_B(\mathbf{x})\mathbf{j}_3$. Des manipulations géométriques de *vecteurs couleur* sont alors possible avec des additions et multiplications quaternioniques, permettant d'écrire de façon intuitive des opérations ayant un sens physique (extraction de la teinte, saturation *etc.*). En particulier, la représentation de Cayley-Dickson permet de séparer l'information d'intensité de l'information chromatique.

Etant donné un signal 2D quaternionique $s(\mathbf{x})$, la transformée de Fourier couleur quaternionique est définie par :

$$\hat{s}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} e^{-\boldsymbol{\mu}\boldsymbol{\omega}^T \mathbf{x}} d\mathbf{x} \quad (2.12)$$

et son inverse s'obtient en inversant le signe de l'exponentielle. Les coefficients de Fourier sont composés d'un module, d'un angle de « phase » et des deux angles d'une donnée appelée « axe ». Cette transformée est paramétrée par le quaternion unitaire $\boldsymbol{\mu}$, qui détermine un axe d'analyse dans l'espace couleur, permettant par exemple de détecter les contours selon un critère colorimétrique dépendant de l'application [117, 118].

La démarche a ensuite été étendue pour proposer des outils quaternioniques permettant de *manipuler* les images couleur [41], notamment à travers la transformée de Fourier quaternionique *inverse*. Ainsi, des procédures de *filtrage* couleur ont pu être définies.

Enfin, il faut noter la définition d'un *banc de filtres* quaternionique couleur il y a maintenant six ans [20, 40] (toujours dans notre équipe), dans le but de définir une représentation multi-échelle bien adaptée aux images couleur. Le schéma numérique est inversible, et la méthode est non-marginale. Les couleurs sont traitées de façon cohérente par rapport à l'axe couleur d'analyse fixé. Au final, il apparaît que la méthode peut également être définie par les outils classiques dans des espaces euclidiens, même si le formalisme des quaternions donne une ré-écriture élégante.

Nous nous intéresserons aux images couleur dans la suite de cette thèse, toutefois le présent chapitre sera réservé aux images en niveaux de gris, pour lesquelles les « dimensions » de l'algèbre des quaternions vont plutôt servir à enrichir l'information *géométrique*, portée par des coefficients quaternioniques munis d'une phase à *trois angles*, par opposition à l'angle unique des coefficients complexes.

2.2 LE FORMALISME DE T. BÜLOW

Dans la continuité des travaux de Hahn sur le signal analytique 2D, autour d'une réflexion sur l'annulation des fréquences négatives par transformées de Hilbert partielles [63, 64], T. Bülöw [15] a proposé un modèle de signal 2D construit dans l'algèbre des quaternions qui généralise de façon élégante les notions de signal analytique et de phase.

2.2.1 Transformée de Fourier quaternionique (QFT)

Soit un signal bidimensionnel à valeurs quaternioniques $s(\mathbf{x})$, la transformée de Fourier quaternionique « QFT » et son inverse sont définies dans [15] par :

$$\hat{s}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} e^{-j_1 \omega_1 x_1} s(\mathbf{x}) e^{-j_2 \omega_2 x_2} d\mathbf{x} = A(\boldsymbol{\omega}) e^{\varphi(\boldsymbol{\omega}) j_1 + \theta(\boldsymbol{\omega}) j_2 + \psi(\boldsymbol{\omega}) j_3} \quad (2.13)$$

$$s(\mathbf{x}) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} e^{+j_1 \omega_1 x_1} \hat{s}(\boldsymbol{\omega}) e^{+j_2 \omega_2 x_2} d\boldsymbol{\omega} \quad (2.14)$$

Cette définition est très analogue à la transformée de Fourier classique, et les équations (1.30) et (1.31) peuvent être retrouvées en identifiant j_1 et j_2 avec l'imaginaire complexe $j \in \mathbb{C}$. Ici, les deux dimensions spatiales x et y sont liées à deux entités algébriques distinctes j_1 et j_2 , ce qui va permettre une décomposition plus riche en termes d'analyse du signal. Les coefficients de Fourier ont désormais *trois* angles de phase, obtenus par la méthode des angles d'Euler décrite à la section 2.1.1, comme illustré sur la figure 2.1.

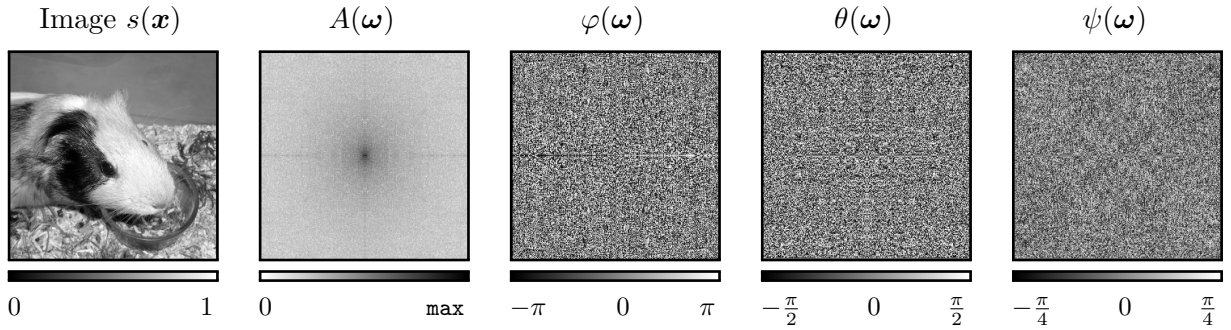


FIGURE 2.1 – Transformée de Fourier quaternionique de T. Bülow. La fréquence $\boldsymbol{\omega} = [0 \ 0]$ est au centre des illustrations, le contraste de A a été renforcé.

Propriétés

La QFT vérifie les propriétés habituelles de conservation de l'énergie, linéarisation des opérateurs différentiels, et lien entre convolution dans un domaine et multiplication dans l'autre [15]. Le test de reconstruction après neutralisation de l'amplitude/la phase donne visuellement le même résultat qu'avec la transformée de Fourier (voir figure 1.5), et d'une manière générale, la phase quaternionique de cette nouvelle représentation sera supposée porter l'information structurelle de l'image. La propriété la plus importante dans notre contexte est le théorème du décalage (du « retard » en 1D), qui se traduit désormais de façon plus riche à travers la phase quaternionique. Le spectre quaternionique d'un signal translaté de $\boldsymbol{\tau}$ noté $s(\mathbf{x} - \boldsymbol{\tau})$ est $e^{-j_1 \omega_1 \tau_1} \hat{s}(\boldsymbol{\omega}) e^{-j_2 \omega_2 \tau_2}$, et donc :

$$s(\mathbf{x}) \leftarrow s(\mathbf{x} - \boldsymbol{\tau}) \Leftrightarrow \begin{cases} \varphi(\boldsymbol{\omega}) \leftarrow \varphi(\boldsymbol{\omega}) - \omega_1 \tau_1 \\ \theta(\boldsymbol{\omega}) \leftarrow \theta(\boldsymbol{\omega}) - \omega_2 \tau_2 \end{cases} \quad (2.15)$$

On retrouve donc dans la phase de la QFT une information de translation 2D, plus flexible que dans le cas classique, et complétée par une troisième donnée ψ .

Plaçons nous maintenant dans le cas où s est réel, ce qui sera presque toujours vrai (on exclut l'étude très particulière des gains fréquentiels de filtres à valeurs quaternionique). Le spectre QFT possède alors une forme de symétrie « hermitienne » - au sens des quaternions. Plus particulièrement, la partie réelle de $\hat{s}$ est paire selon les axes x_1 et x_2 , sa partie j_1 -imaginaire est x_1 -impaire et x_2 -paire, sa partie j_2 -imaginaire est x_1 -paire et x_2 -impaire, et sa partie j_3 -imaginaire est impaire selon les axes x_1 et x_2 [15]. En utilisant la parité des différents termes intégrés sur $\mathbb{R}^2$, on peut réduire la formule (2.14) à sa version « réelle » de cette façon :

$$s(\mathbf{x}) = \frac{1}{\pi^2} \int_{\mathbb{R}^+ \times \mathbb{R}^+} \Re \left\{ e^{+j_1 \omega_1 x_1} \hat{s}(\boldsymbol{\omega}) e^{+j_2 \omega_2 x_2} \right\} d\boldsymbol{\omega} \quad (2.16)$$

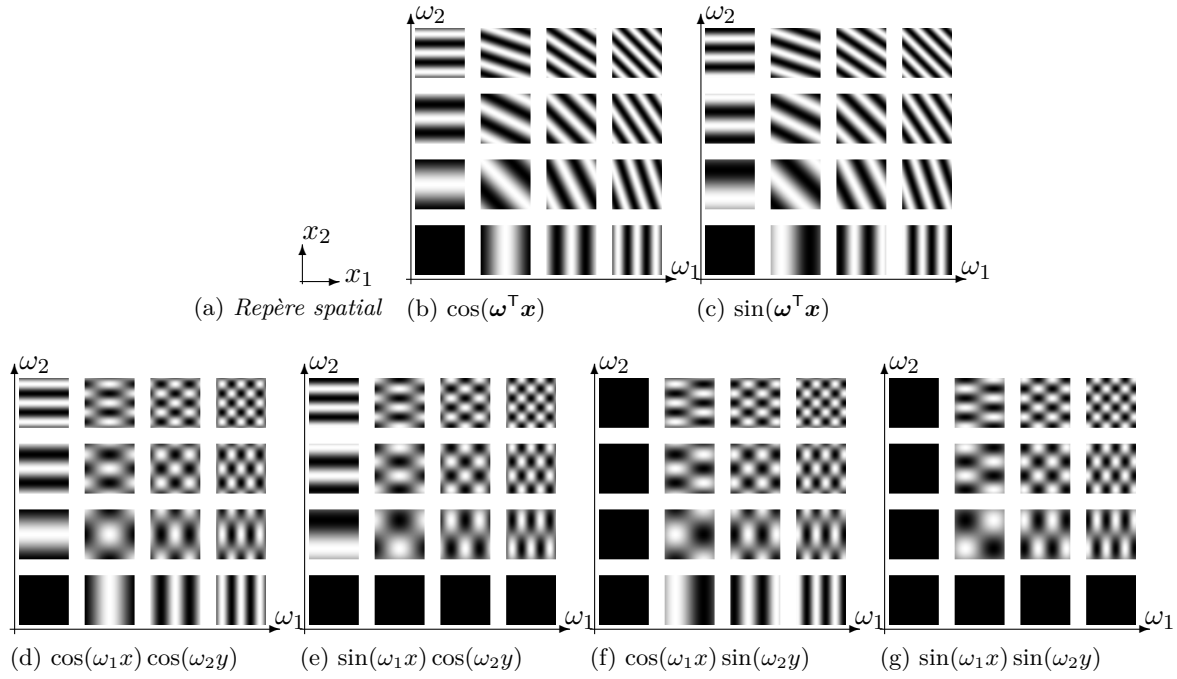


FIGURE 2.2 – (b) et (c) : Parties réelle et imaginaire des ondes planes orientées, base de la transformée de Fourier classique. (d,e,f,g) : Parties réelle et imaginaires quaternioniques des boîtes à œufs étirées, atomes de la QFT. Chaque image est une fonction du repère spatial (a), oscillant autour du « gris moyen ».

Alors qu'avec la transformée de Fourier classique, la moitié des coefficients de Fourier suffisent à reconstruire un signal réel, limitant l'intégrale à un *demi-plan* du domaine fréquentiel (voir eq. (1.33)), ici c'est le *premier quadrant* du spectre qui suffit. L'information de l'ensemble du domaine fréquentiel est déplacée dans les quatre composantes des coefficients quaternioniques.

On peut donc regrouper les deux exponentielles dans l'intégrale (2.13), et identifier les atomes quaternioniques analogues aux exponentielles complexes classiques :

$$e^{j_1 \omega_1 x_1 + j_2 \omega_2 x_2} = \cos(\omega_1 x_1) \cos(\omega_2 x_2) + j_1 \sin(\omega_1 x_1) \cos(\omega_2 x_2) + j_2 \cos(\omega_1 x_1) \sin(\omega_2 x_2) + j_3 \sin(\omega_1 x_1) \sin(\omega_2 x_2) \quad (2.17)$$

Les composantes cartésiennes de ces fonctions sont illustrées figure 2.2. Afin d'appréhender cette nouvelle famille d'atomes, et de mieux comprendre le sens géométrique de la phase quaternionique, nous proposons d'étudier la QFT sous l'angle que nous avons adopté au chapitre précédent - celui de la reconstruction d'image par sommation d'oscillations.

Analyse des atomes quaternioniques

Dans la perspective de l'étude des *ondelettes* quaternioniques, nous proposons d'approfondir l'analyse des atomes de la QFT.

Rappelons que dans le cas classique, les atomes de Fourier sont identifiés comme la partie réelle du produit dans $\mathbb{C}$ d'une exponentielle et d'un coefficient $\Re\{\hat{s}(\mathbf{x})e^{j\omega^T \mathbf{x}}\}$. Il s'agit comme nous l'avons déjà vu d'une *onde plane orientée* dont l'amplitude et la phase sont déterminées par $\hat{s}(\mathbf{x})$. Nous venons de voir que les atomes quaternioniques peuvent être identifiés de la même façon, mais il reste à éclaircir le lien entre d'une part l'amplitude et la phase du coefficient de Fourier, et d'autre part la forme de l'*atome réel* reconstruit dans l'intégrale de l'équation (2.16).

Pour cela, on considère les composantes cartésiennes et polaires d'un coefficient QFT quelconque :

$$\hat{s} = \hat{s}_1 + j_1 \hat{s}_2 + j_2 \hat{s}_3 + j_3 \hat{s}_4 = A e^{j_1 \varphi + j_2 \theta + j_3 \psi} \quad (2.18)$$

Dans un premier temps, une analyse cartésienne

$$\begin{aligned} \Re \left\{ e^{j_1 \omega_1 x_1} \hat{s} e^{j_2 \omega_2 x_2} \right\} \\ = \hat{s}_1 \cos(\omega_1 x_1) \cos(\omega_2 x_2) - \hat{s}_2 \sin(\omega_1 x_1) \cos(\omega_2 x_2) - \hat{s}_3 \cos(\omega_1 x_1) \sin(\omega_2 x_2) + \hat{s}_4 \sin(\omega_1 x_1) \sin(\omega_2 x_2) \end{aligned} \quad (2.19)$$

nous montre une forte analogie avec le cas complexe : la multiplication de $\mathbb{H}$ qui est en jeu correspond à faire une somme pondérée des « parties paires et impaires » de l'atome. Dans $\mathbb{C}$, le lien avec l'amplitude et la phase du signal réel résultant était obtenu avec une identité trigonométrique. Nous proposons donc de prendre en compte directement les *composantes polaires* de $\hat{s}$, à travers la ré-écriture suivante :

$$\begin{aligned} \Re \left\{ e^{j_1 \omega_1 x_1} \hat{s} e^{j_2 \omega_2 x_2} \right\} &= \frac{A}{\sqrt{2}} (a_1 \cos(\omega_1 x_1 + \omega_2 x_2 + \varphi_1) + a_2 \cos(\omega_1 x_1 - \omega_2 x_2 + \varphi_2)) \quad (2.20) \\ \text{avec} \quad a_1 &= \cos(\psi + \pi/4) \in [0; 1] \quad \varphi_1 = \varphi + \theta \\ a_2 &= \sin(\psi + \pi/4) \in [0; 1] \quad \varphi_2 = \varphi - \theta \end{aligned}$$

Cette écriture va nous permettre d'identifier la forme du signal reconstruit, en fonction des valeurs de la phase quaternionique. Tout d'abord, on remarque que l'atome consiste en une superposition de deux ondes planes dont les amplitudes et phases respectives sont déterminées par φ , θ et ψ . Il s'agit d'une traduction nouvelle de l'information contenue dans *deux* coefficients de Fourier classiques, aux coordonnées $[\omega_1 \ \omega_2]$ et $[\omega_1 \ -\omega_2]$. Cela illustre comment les coefficients du premier quadrant du spectre contiennent également l'information du quadrant voisin.

Nous proposons d'étudier la forme des atomes en fonction des valeurs de ψ - nous rappelons que cet angle est restreint à l'intervalle $[-\pi/4; \pi/4]$ (cf. eq. (2.3)). L'illustration de la figure 2.3 montre des atomes de QFT pour une fréquence et une phase fixées, à l'exception de ψ qui varie de $-\pi/4$ à $\pi/4$.

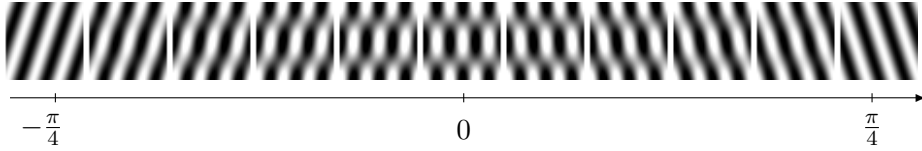


FIGURE 2.3 – Atomes de QFT (équation (2.20)) pour différentes valeurs de ψ ($\omega = 2\pi[0.025 \ 0.075]$, $\varphi = 0$, $\theta = 0$).

Dans le cas où ψ est proche de $\pm\pi/4$ l'atome correspond à une onde plane pure $\frac{A}{\sqrt{2}} \cos(\omega_1 x_1 \pm \omega_2 x_2 + \varphi \pm \theta)$, dont l'orientation est déterminée d'une part par la coordonnée fréquentielle ω , et d'autre part par le signe de ψ , qui bascule avec le quadrant voisin. On retrouve alors un atome de Fourier classique. Ce cas correspond en quelque sorte à une situation dégénérée, puisque l'atome ne dépend que de la somme (ou la différence) de φ et θ , ce qui rappelle l'ambiguïté déjà soulevée dans le calcul de la représentation polaire d'un quaternion pour $\psi = \pm\pi/4$ (voir section 2.1.1). L'information de φ et θ , réduite à une seule valeur, est une phase de Fourier classique.

Dans le cas où $\psi \approx 0$, l'atome prend la forme plus complexe d'une « boîte à œufs ». En faisant varier la fréquence ω et la phase (φ, θ), la variété des formes obtenues est bien reflétée par la figure 2.2 (d-g). Notons que lorsque $\omega_1 \approx 0$ ou $\omega_2 \approx 0$, l'atome est à nouveau réduit à une onde plane, horizontale ou verticale. Au contraire, pour les fréquences *diagonales*, l'atome peut présenter une structure *isotrope*. Plus généralement, on constate qu'un coefficient quaternionique permet d'encoder une grande variété de structures, contrairement aux coefficients complexes de la transformée de Fourier classique, toujours liés à une ligne ou un bord orienté. Ici, des structures « i2D » (intrinsèquement 2D) peuvent être représentées en un seul coefficient, et caractérisées par la phase quaternionique.

Notons que cette analyse que nous proposons explique un résultat expérimental présenté dans [15], qui consistait à étudier la phase du signal suivant :

$$s_\lambda(\mathbf{x}) = (1 - \lambda) \cos(\omega_1 x_1 + \omega_2 x_2) + \lambda \cos(\omega_1 x_1 - \omega_2 x_2) \quad (2.21)$$

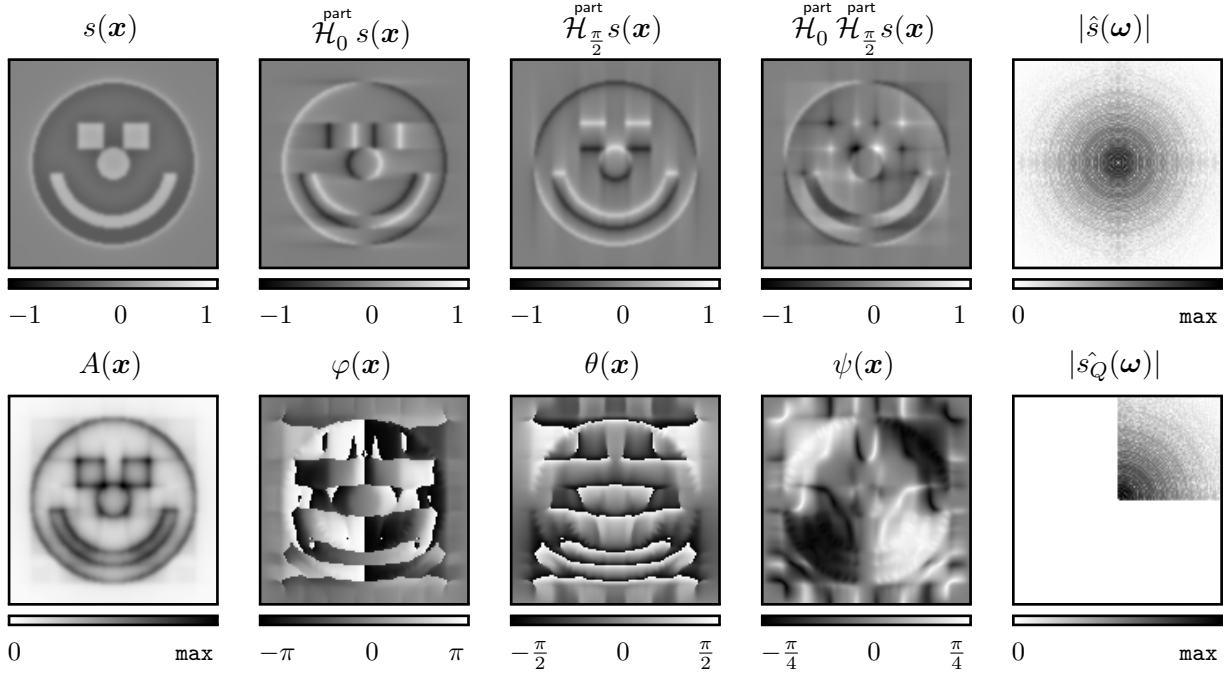


FIGURE 2.4 – *Signal quaternionique analytique* $s_Q = Ae^{j_1\varphi+j_2\theta+j_3\psi}$ associé à s . Le contraste des spectres a été renforcé.

pour différentes valeurs de λ dans $[0; 1]$. Ce signal est sensiblement de la même forme que l'atome réel de QFT que nous avons mis en évidence (avec $\varphi = \theta = 0$), c'est pourquoi T. Bülow obtient empiriquement une relation presque linéaire entre λ et ψ .

Dans tous les cas, les angles φ et θ déterminent la position précise de l'atome, de manière analogue à la phase de Fourier classique, et conformément au théorème du décalage (équation (2.15)).

Finalement, l'algèbre des quaternions apporte un nouveau concept de « pondération et déphasage » construit par le produit quaternionique d'une « boîte à œufs » et d'un coefficient de Fourier, qui permet de modéliser les atomes de base de façon plus riche que le traditionnel déphasage directionnel d'ondes planes. On s'attend à ce que des structures plus complexes - par exemple isotropes - puissent être mieux représentées. Notons toutefois que les structures de type « boîtes-à-œufs » qui sont introduites sont définies de façon séparable, et risquent donc de souffrir de leur dépendance aux axes x_1 et x_2 . Voyons maintenant le *signal quaternionique analytique*, qui définit la contrepartie spatiale de la QFT.

2.2.2 Le signal quaternionique analytique

Suite à plusieurs tentatives d'extensions à l'image du signal analytique [63, 64, 67] basées sur différentes utilisations de la transformée de Hilbert partielle (section 1.3.4), et vu les symétries de la transformée de Fourier quaternionique, T. Bülow propose le signal « quaternionique analytique » s_Q suivant :

$$s_Q(\mathbf{x}) = s(\mathbf{x}) + j_1 \mathcal{H}_0^{\text{part}} s(\mathbf{x}) + j_2 \mathcal{H}_{\frac{\pi}{2}}^{\text{part}} s(\mathbf{x}) + j_3 \mathcal{H}_0^{\text{part}} \mathcal{H}_{\frac{\pi}{2}}^{\text{part}} s(\mathbf{x}) \quad (2.22)$$

$$= A(\mathbf{x}) e^{j_1 \varphi(\mathbf{x}) + j_2 \theta(\mathbf{x}) + j_3 \psi(\mathbf{x})} \quad (2.23)$$

Ce signal construit par ses transformées de Hilbert partielles et totale est illustré en représentation cartésienne et polaire sur la figure 2.4. La QFT $\hat{s}_Q$ d'un tel signal est nulle dans les trois quadrant du spectre où $\omega_1 < 0$ ou $\omega_2 < 0$, ce qui correspond aux « fréquences négatives » au sens large. Comme nous l'avons vu, les symétries de la QFT sont telles que le premier quadrant contient toute l'information de l'image. En conséquence, et contrairement au modèle de Hahn, le signal s est contenu dans la partie réelle de son extension quaternionique s_Q .

À nouveau, on s'intéresse à la correspondance entre la forme du signal et les valeurs de son amplitude/phase. Puisque s est contenu dans la partie réelle de s_Q , il vient qu'il peut être exprimé sous une forme de « modulation » de manière analogue au cas 1D :

$$s = \Re\{s_Q\} = \frac{A(\mathbf{x})}{\sqrt{2}} (\cos(\psi(\mathbf{x}) + \pi/4) \cos(\omega_1 x_1 + \omega_2 x_2 + \varphi(\mathbf{x}) + \theta(\mathbf{x})) + \sin(\psi(\mathbf{x}) + \pi/4) \cos(\omega_1 x_1 - \omega_2 x_2 + \varphi(\mathbf{x}) - \theta(\mathbf{x}))) \quad (2.24)$$

(voir aussi l'équation (2.20)). Cette écriture va nous permettre de comprendre l'information portée par la phase quaternionique. Il s'agit donc, comme dans le cas classique, d'un atome de Fourier dont l'amplitude et la phase ne sont plus des constantes mais des fonctions de $\mathbf{x}$. Ici, pour interpréter correctement le signal quaternionique, on devra faire l'hypothèse que A et ψ sont *localement* constants et φ et θ sont *localement* linéaires. Le signal s analysé sera ainsi modélisé localement par un atome de QFT, c'est-à-dire une boîte à œufs combinée à une onde plane. Ces hypothèses sur les paramètres du modèle sont semblables à celle que nous avons présentées pour le signal analytique 1D (section 1.1.2), et convergent vers l'idée que le signal à analyser doit être à *bande-étroite*.

Voyons maintenant concrètement l'information portée par l'amplitude et la phase *locale*. La figure 2.4 illustre s_Q pour un signal s à bande-étroite, que nous avons obtenu par filtrage passe-bande d'une image en niveaux de gris. En conséquence, les contours de l'image d'origine sont traduits ici par des oscillations locales, qui vont être caractérisées par le calcul de $A(\mathbf{x})$, $\varphi(\mathbf{x})$, $\theta(\mathbf{x})$ et $\psi(\mathbf{x})$. On voit que l'amplitude A est forte là où des éléments oscillants sont présents dans s , à l'instar d'un détecteur de contours. Les angles φ et θ sont très analogues à la phase classique 1D, et présentent le même type de comportement « monotone », mais cette fois séparément dans les directions de x_1 et de x_2 . Le troisième angle est une donnée nouvelle, qui présente clairement des valeurs autour de $\pm\pi/4$ (blanc et noir) au niveau des structures *directionnelles*. D'une manière générale, la visualisation de la phase est peu intuitive, et il faut garder à l'esprit que ses valeurs n'ont de sens que lorsque l'amplitude est forte. En effet, pour les mêmes raisons qu'en 1D, les coefficients de faible amplitude sont peu significatifs et donc leur phase n'est pas à prendre en compte. L'usage (ou plutôt l'usage dans le cas 1D) veut que l'on se concentre principalement sur les valeurs de phase au niveau des maxima de $A(\mathbf{x})$, là où elles traduisent complètement la forme locale comme un atome QFT déphasé. Le but de notre travail est d'utiliser cette phase comme un descripteur de la *géométrie locale* de l'image.

Puisque cette analyse quaternionique nécessite un filtrage passe-bande, et que nous souhaitons avoir une représentation complète, il nous faut collecter *toutes* les bandes d'un signal. C'est pourquoi nous allons maintenant présenter une analyse quaternionique par bancs de filtres.

2.2.3 Ondelettes quaternioniques

Une approche multi-résolution de l'analyse quaternionique a déjà été présentée par T. Bülow [15] à travers une extension quaternionique des filtres de Gabor. Des ondelettes quaternioniques continues dans le même esprit ont ensuite été présentées dans [5].

Approche type Gabor

La construction d'une décomposition « à la Gabor » est parfaitement analogue au cas classique, puisque les atomes sont définis par un fenêtrage spatial gaussien des exponentielles quaternioniques :

$$g_{\sigma,\xi}(\mathbf{x}) = h(\mathbf{x})e^{(j_1\xi_1x_1+j_2\xi_2x_2)} \xleftrightarrow{\mathcal{F}} h(\boldsymbol{\omega} - \boldsymbol{\xi}) \quad (2.25)$$

L'expression de la fonction gaussienne h est de la forme $ae^{-\|\mathbf{x}\|^b}$ où a est un facteur de normalisation et b détermine la largeur de la fonction, qui dépend de σ . Nous omettons volontairement de préciser les valeurs de a et b car elles diffèrent selon les sources et ne sont pas indispensables

pour la compréhension de notre travail. On englobe donc dans $g_{\sigma,\xi}$ à la fois le cas des *filtres* de Gabor dont le support spatial est fixe et celui des *ondelettes* de Gabor construites par dilatation d'une forme unique. Le spectre de $g_{\sigma,\xi}$ - calculé par la QFT à partir de la réponse impulsionnelle du filtre à valeurs quaternioniques - est réduit à une gaussienne centrée en $\omega = \xi$, et concentre donc son énergie dans le premier quadrant du spectre si $\xi_1, \xi_2 > 0$. C'est alors un signal approximativement *analytique* au sens de T. Bülöw (l'emploi d'une fenêtre gaussienne entraîne le même type d'approximation que nous avons déjà expliqué en 1D - section 1.2.1 et figure 1.2).

Finalement, on s'intéressera au signal de sortie d'un filtrage de Gabor autour de la fréquence centrale ξ :

$$s_\xi(\mathbf{x}) = s(\mathbf{x}) * g_{\sigma,\xi}(\mathbf{x}) = A_\xi(\mathbf{x}) e^{j_1 \varphi_\xi(\mathbf{x}) + j_2 \theta_\xi(\mathbf{x}) + j_3 \psi_\xi(\mathbf{x})} \quad (2.26)$$

Puisque l'opérateur de Hilbert partiel commute avec la convolution, on sait que la sous-bande s_ξ est elle aussi approximativement analytique. Son l'amplitude A_ξ et sa phase locales ($\varphi_\xi, \theta_\xi, \psi_\xi$) pourront alors être étudiées en termes d'informations géométriques. En revanche, comme pour la transformée de Gabor classique, la reconstruction exacte n'est pas possible en pratique.

Avant de passer à la définition d'une transformée en ondelettes quaternioniques peu redondante et à reconstruction parfaite, nous souhaitons introduire un lien intéressant avec les filtres de Gabor complexes. A partir des composantes cartésiennes d'un filtre de Gabor quaternionique, on peut identifier deux filtres de Gabor complexes :

$$g_{\sigma,\xi} = g_1 + j_1 g_2 + j_2 g_3 + j_3 g_4 = h(\mathbf{x}) e^{(j_1 \xi_1 x_1 + j_2 \xi_2 x_2)} \in \mathbb{H} \quad (2.27)$$

$$g^+ = (g_1 - g_4) + j(g_2 + g_3) = h(\mathbf{x}) e^{j(\xi_1 x_1 + \xi_2 x_2)} \in \mathbb{C} \quad (2.28)$$

$$g^- = (g_1 + g_4) + j(g_2 - g_3) = h(\mathbf{x}) e^{j(\xi_1 x_1 - \xi_2 x_2)} \in \mathbb{C} \quad (2.29)$$

dont les fréquences centrales sont respectivement $[\xi_1 \ \xi_2]$ et $[\xi_1 \ -\xi_2]$. Ce lien étroit entre le monde de Fourier classique et le monde Fourier quaternionique suggère qu'une analyse quaternionique pourrait être construite à partir d'une transformée en ondelettes complexe (section 1.3.4), suivie de recombinaisons des coefficients.

Transformée en ondelettes quaternionique

La *transformée en ondelettes quaternionique* (QWT) de W. L. Chan *et al.* [24] était la première décomposition 2D *analytique* à reconstruction parfaite lorsqu'elle a été proposée en 2004 [23]. Elle correspond à une transformée séparable dyadique utilisant trois ondelettes $\psi^a(\mathbf{x})$, $\psi^b(\mathbf{x})$, $\psi^c(\mathbf{x})$ et une fonction d'échelle $\phi(\mathbf{x})$ définies de manière séparable par les fonctions 1D $\phi(t)$ et $\psi(t)$ et leur transformées de Hilbert respectives :

$$\phi(\mathbf{x}) = \phi(x_1)\phi(x_2) + j_1 \mathcal{H}\phi(x_1)\phi(x_2) + j_2 \phi(x_1)\mathcal{H}\phi(x_2) + j_3 \mathcal{H}\phi(x_1)\mathcal{H}\phi(x_2) \quad (2.30)$$

$$\psi^a(\mathbf{x}) = \phi(x_1)\psi(x_2) + j_1 \mathcal{H}\phi(x_1)\psi(x_2) + j_2 \phi(x_1)\mathcal{H}\psi(x_2) + j_3 \mathcal{H}\phi(x_1)\mathcal{H}\psi(x_2) \quad (2.31)$$

$$\psi^b(\mathbf{x}) = \psi(x_1)\phi(x_2) + j_1 \mathcal{H}\psi(x_1)\phi(x_2) + j_2 \psi(x_1)\mathcal{H}\phi(x_2) + j_3 \mathcal{H}\psi(x_1)\mathcal{H}\phi(x_2) \quad (2.32)$$

$$\psi^c(\mathbf{x}) = \psi(x_1)\psi(x_2) + j_1 \mathcal{H}\psi(x_1)\psi(x_2) + j_2 \psi(x_1)\mathcal{H}\psi(x_2) + j_3 \mathcal{H}\psi(x_1)\mathcal{H}\psi(x_2) \quad (2.33)$$

Il vient deux choses de cette définition. La première est qu'il est facile de vérifier que les fonctions $\phi(\mathbf{x})$, $\psi^a(\mathbf{x})$, $\psi^b(\mathbf{x})$ et $\psi^c(\mathbf{x})$ sont approximativement analytiques au sens de T. Bülöw, et donc offrent une analyse quaternionique multi-échelle de l'image dans trois directions globales séparées : horizontale, verticale et diagonale. La seconde est que l'implantation de cette décomposition en ondelettes peut être effectuée par quatre bancs de filtres 2D séparables distincts, un pour chaque composante cartésienne des coefficients quaternioniques. La particularité par rapport aux bancs de filtres séparables habituels est que les filtres « pour les lignes » peuvent être différents des filtres « pour les colonnes ». Les deux ensembles de filtres concernés doivent correspondre à deux ondelettes formant une paire de Hilbert, ce qui est exactement la définition d'une transformée en ondelettes complexes 1D. Il est donc naturel d'utiliser la partie réelle et

la partie imaginaire d'une transformée complexe - comme le dual-tree présenté précédemment - d'une manière séparable pour implanter la QWT de [24].

L'avantage de ce lien avec le dual-tree est que toute la réflexion sur l'approximation de la transformée de Hilbert entre les filtres reste valide. Il suffit donc de *combiner* correctement les sous-bandes obtenues, de façon à correspondre au modèle quaternionique. Un schéma complet de l'implantation du banc de filtres QWT est donné figure 2.5. Les coefficients (notés A , B , C et D sur la figure 2.5) sont d'abord obtenus sous forme *cartésienne* et doivent être convertis en représentation polaire pour obtenir l'amplitude et la phase :

$$w_{LL,j}[\mathbf{k}] = \left\langle s(\mathbf{x}), 2^{-j} \phi(2^{-j} \mathbf{x} - \mathbf{k}) \right\rangle = A_{LL,j} e^{j_1 \varphi_{LL,j} + j_2 \theta_{LL,j} + j_3 \psi_{LL,j}} \quad (2.34)$$

$$w_{LH,j}[\mathbf{k}] = \left\langle s(\mathbf{x}), 2^{-j} \psi^a(2^{-j} \mathbf{x} - \mathbf{k}) \right\rangle = A_{LH,j} e^{j_1 \varphi_{LH,j} + j_2 \theta_{LH,j} + j_3 \psi_{LH,j}} \quad (2.35)$$

$$w_{HL,j}[\mathbf{k}] = \left\langle s(\mathbf{x}), 2^{-j} \psi^b(2^{-j} \mathbf{x} - \mathbf{k}) \right\rangle = A_{HL,j} e^{j_1 \varphi_{HL,j} + j_2 \theta_{HL,j} + j_3 \psi_{HL,j}} \quad (2.36)$$

$$w_{HH,j}[\mathbf{k}] = \left\langle s(\mathbf{x}), 2^{-j} \psi^c(2^{-j} \mathbf{x} - \mathbf{k}) \right\rangle = A_{HH,j} e^{j_1 \varphi_{HH,j} + j_2 \theta_{HH,j} + j_3 \psi_{HH,j}} \quad (2.37)$$

On obtient donc une décomposition en ondelettes analogue à la très habituelle transformée bi-orthogonale dyadique, mais dont les coefficients « $w_{...}[\mathbf{k}]$ » ont une valeur quaternionique.

Transformée inverse et interprétation

La reconstruction du signal de départ est définie comme la somme des quatre reconstructions cartésiennes [24] - ce qui était déjà le cas avec les ondelettes complexes. En fait, n'importe lequel des quatre bancs de filtres peut à lui seul reconstruire l'image, et on peut s'interroger sur la justification de cette méthode de reconstruction finalement très simple. Dans la littérature sur le dual-tree et sur la QWT, cette question n'est pas vraiment traitée¹, c'est pourquoi nous avons déjà apporté une interprétation autour des ondelettes complexes aux sections 1.2.3 et 1.3.4. La compréhension de cette reconstruction réside dans l'analogie avec les « parties paires et impaires » des atomes *réels* de Fourier (equations (2.16) et (2.19)). Les coefficients doivent *conférer* à l'atome son amplitude et sa phase. Cette opération de pondération et déphasage est exprimée dans l'intégrale de Fourier par une *multiplication* quaternionique, mais est en fait équivalente, comme nous l'avons vu, à la somme des « parties paires et impaires » de l'atome non déphasé, pondérée par les composantes cartésiennes du *coefficient* (eq. (2.19)). Voilà pourquoi la reconstruction de la QWT fait la somme des « parties paires et impaires » de l'ondelette (les différents bancs de filtres), pondérée par les composantes cartésiennes des coefficients d'ondelettes (A , B , C et D , figure 2.5) - ce qui revient à sommer les quatre reconstructions individuelles.

Ce principe très analogue à la transformée de Fourier - et totalement nouveau pour une transformée en ondelettes 2D - est à l'origine de l'invariance par translation de l'amplitude des coefficients, car dans la phase est contenue une informations structurelle de précision sous-pixellique. Les structures sont reconstruites par un « dosage local de cosinus et de sinus » bien précis, qui ne dépend pas comme avec les ondelettes classiques d'un équilibre avec les coefficients voisins dont les oscillations se compensent mutuellement. Cela évitera comme on le verra l'introduction des artefacts oscillants classiques après traitement non linéaire dans le domaine des ondelettes.

En revanche, on peut déplorer l'existence de trois sous-bandes orientées pour une échelle - horizontal, vertical, diagonal - à l'origine d'un schéma « ligne-colonne » (séparable) qui risque de *privilégier* les structures alignées aux axes du repère. A une échelle donnée, et en comparaison avec le dual-tree 2D, le défi de la QWT sera de faire des ses *trois* coefficients quaternioniques (pour les trois sous-bandes) une meilleure représentation que les *six* coefficients complexes du dual-tree.

L'illustration d'un exemple de données obtenues sur une image réelle est donnée figure 2.6. La visualisation des angles de phase n'est pas très intuitive à cause du sous-échantillonnage des

1. Notons toutefois la reconstruction par dual-tree 2D définie dans [27].

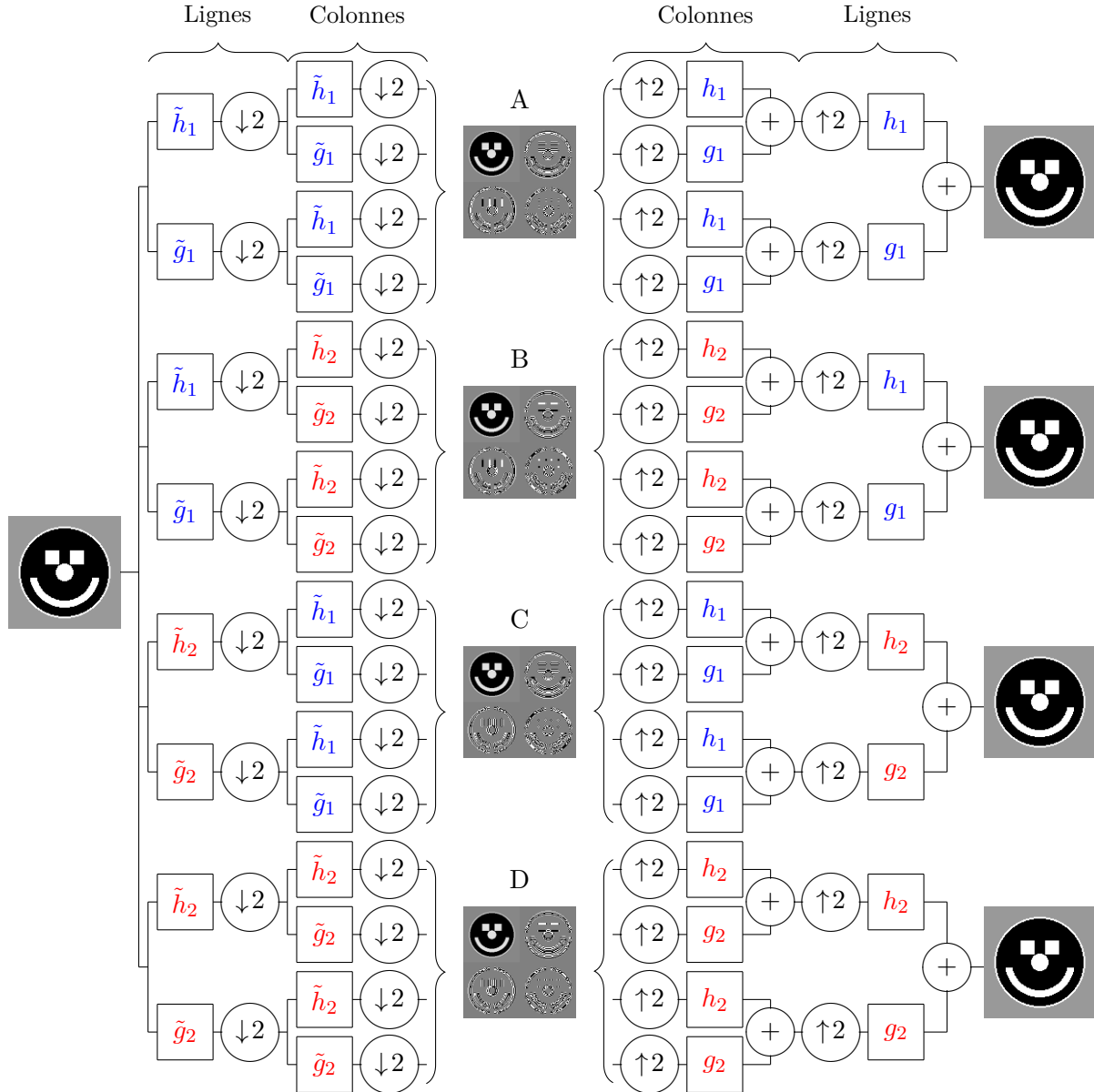


FIGURE 2.5 – Implantation séparable du banc de filtres dual-tree. Les 4 bancs de filtres 2D résultants offrent deux interprétations possibles : 1) Une décomposition par trois ondelettes quaternioniques orientées horizontalement, verticalement et diagonalement [24]. Les sous-bandes à valeurs quaternioniques sont donc $A + \mathbf{j}_1 B + \mathbf{j}_2 C + \mathbf{j}_3 D$. 2) Une décomposition par 6 ondelettes complexes type Gabor orientées vers $k\pi/6$ [124]. Les sous bandes sont donc $(A + D) + \mathbf{j}(B - C)$ et $(A - D) + \mathbf{j}(C - B)$.

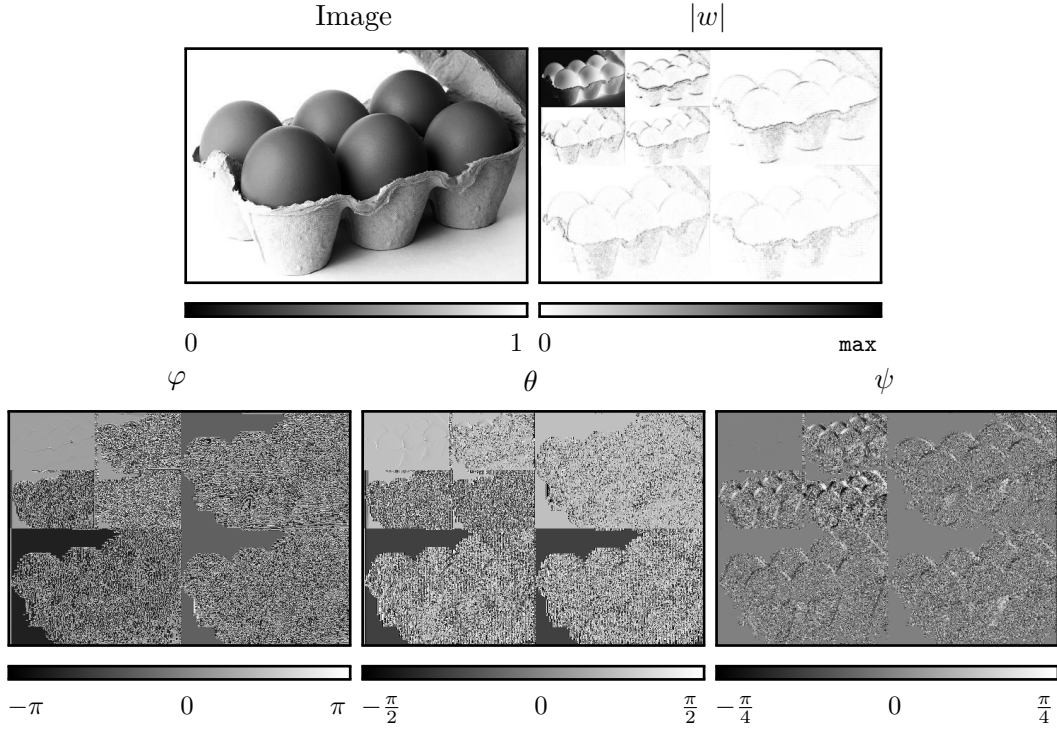


FIGURE 2.6 – QWT d’une image réelle. Le contraste de l’amplitude a été renforcé.

sous-bandes, comme le montre plus précisément la figure 2.7. En contrepartie, on espère que ces données de phase moins nombreuses portent beaucoup d’information visuelle.

Pour conclure sur la définition de la QWT, l’algorithme hérite des ondelettes complexes une invariance par translation approximative pour une redondance raisonnable, et ses filtres sont approximativement liés par des transformées de Hilbert partielles. Les coefficients des sous-bandes sont quaternioniques et s’étudient donc en représentation polaire, afin d’extraire les information de *présence* et de *structure locale* à travers l’amplitude et la phase. On s’attend à ce qu’elle surpasse les ondelettes utilisées habituellement grâce à ses propriétés nouvelles qui pallient les défauts classiques de variance, effet pseudo-Gibbs, et qui combinent le besoin théorique d’une notion de phase - fondamentale en traitement du signal.

2.3 UTILISATION DE LA QWT

Les données quaternioniques ont été peu étudiées jusqu’à maintenant. D’abord, T. Bülow a proposé des applications en estimation de flot optique basé sur ses filtres de Gabor quaternioniques, travail ensuite repris dans le cas de la QWT dans [5, 158, 24]. T. Bülow a également proposé un algorithme de segmentation de texture s’appuyant sur le troisième angle ψ , et un calcul original de représentation en « wedgelets » a été proposé dans [24]. Cette section synthétise l’ensemble de ces propositions.

2.3.1 Les angles φ et θ et le flot optique

Le flot optique est un champ de *vecteurs de déplacement* qui se calcule par exemple à partir de deux images successives d’une vidéo. L’hypothèse principale dans la démarche de calcul d’un flot optique est que l’intensité lumineuse des éléments de la scène est conservée d’une image à l’autre. Alors, on peut utiliser des procédures de calcul sur l’intensité des pixels pour estimer le champ de vecteurs correspondant aux déplacements dans l’image. L’utilisation de la phase locale par décomposition de Gabor est classique [51], et permet une précision sous-pixelique.

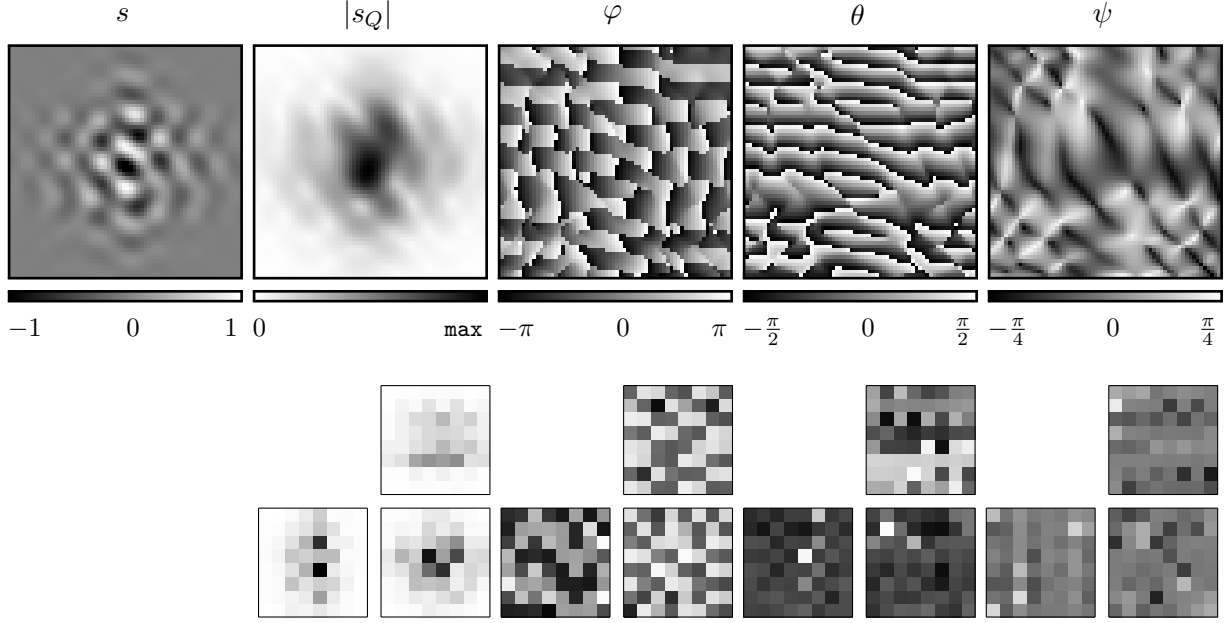


FIGURE 2.7 – Haut : Signal analytique associé à un signal 2D s à bande étroite. Bas : Sous-bandes horizontales, verticales et diagonales de l'échelle de la QWT de s contenant la majorité de l'énergie. Les sous-bandes QWT sont zoomées.

T. Bülow propose d'utiliser le formalisme quaternionique pour améliorer la méthode, ce qui a été repris dans [5, 158], puis transposé à la QWT dans [24].

Le théorème du décalage existant pour la QFT (eq. (2.15)) est approximativement conservé pour une analyse quaternionique spatiale. Afin d'unifier le cas Gabor avec le cas ondelettes, on considère d'une manière générale les quatre coordonnées polaires de l'analyse quaternionique du signal $s(\mathbf{x})$ autour de la fréquence ξ que l'on note :

$$\{ A_\xi(\mathbf{x}), \varphi_\xi(\mathbf{x}), \theta_\xi(\mathbf{x}), \psi_\xi(\mathbf{x}) \} \quad (2.38)$$

obtenue par filtrage de Gabor ou comme une sous-bande particulière de la QWT. Un décalage de s entraîne le changement suivant :

$$s(\mathbf{x}) \leftarrow s(\mathbf{x} - \boldsymbol{\tau}) \Leftrightarrow \begin{cases} \varphi_\xi(\mathbf{x}) \leftarrow \varphi_\xi(\mathbf{x}) - \tau_1 \xi_1^l(\mathbf{x}) \\ \theta_\xi(\mathbf{x}) \leftarrow \theta_\xi(\mathbf{x}) - \tau_2 \xi_2^l(\mathbf{x}) \end{cases} \quad \text{avec } \xi^l(\mathbf{x}) \text{ à estimer localement} \quad (2.39)$$

Le déphasage dépend de la « fréquence locale » $\xi^l(\mathbf{x})$ du signal - qui n'est pas connue *a priori*, contrairement aux coefficients de Fourier qui sont attachées à une fréquence précise. Ici, ξ^l peut varier dans une certaine bande liée au gabarit de l'ondelette.

Soient deux images $s_A(\mathbf{x})$ et $s_B(\mathbf{x})$, considérant leurs phases quaternioniques respectives $(\varphi_{\xi,A}, \theta_{\xi,A})$ et $(\varphi_{\xi,B}, \theta_{\xi,B})$, la carte de vecteurs de déplacement relatif $[d_{x_1}(\mathbf{x}) \ d_{x_2}(\mathbf{x})]^T$ entre s_A et s_B sera calculée par :

$$d_{x_1}(\mathbf{x}) = \frac{\varphi_{\xi,B}(\mathbf{x}) - \varphi_{\xi,A}(\mathbf{x}) + k_1\pi}{\xi_1^l(\mathbf{x})} \quad d_{x_2}(\mathbf{x}) = \frac{\theta_{\xi,B}(\mathbf{x}) - \theta_{\xi,A}(\mathbf{x}) + k_2\pi}{\xi_2^l(\mathbf{x})} \quad (2.40)$$

où k_1 et k_2 sont tels que les numérateurs soient dans $[-\pi; \pi]$. Ce calcul est clairement conditionné par la qualité de l'estimation de la fréquence locale. Dans le cas le plus simple, on décide que cette fréquence locale est égale à la fréquence centrale du filtre de Gabor utilisé. Mais à cause du principe d'incertitude sur la localisation spatiale et fréquentielle, l'utilisation de ξ comme fréquence locale de référence n'est pas très précise. De plus, dans le cas des ondelettes quaternioniques, les fréquences centrales des filtres ne sont pas forcément explicites, et les bandes passantes - par octaves - sont trop larges pour une telle simplification de la fréquence locale.

La solution est d'approximer les dérivées partielles de la phase, ce qui était déjà classique avec les filtres de Gabor complexes [51, 12]. Dans le cas quaternionique, le calcul est simplifié par deux estimations 1D, puisque φ a tendance à varier le long de x_1 et θ le long de x_2 , contrairement à la phase de Gabor complexe qui doit être dérivée dans la direction de ξ , ce qui va être mal défini sur une grille discrète. Dans le cas de la QWT, nous avons déjà vu (figure 2.7) que le sous-échantillonnage rend la lecture de la phase difficile. C'est également un problème pour l'estimation des dérivées de phase car les variations sont grandes entre deux coefficients voisins. Une astuce proposée par [24] consiste à créer *manuellement* un décalage de 1 pixel en x_1 et en x_2 , par exemple sur l'image s_A . On obtient une image $s_{A_2}(\mathbf{x}) = s_A(\mathbf{x} - [1 \ 1]^T)$ dont l'analyse quaternionique donnera la phase suivante :

$$\varphi_{\xi, A_2}(\mathbf{x}) = \varphi_{\xi, A}(\mathbf{x}) - \xi_1^l(\mathbf{x}) \quad \theta_{\xi, A_2}(\mathbf{x}) = \theta_{\xi, A}(\mathbf{x}) - \xi_2^l(\mathbf{x}) \quad (2.41)$$

qui donne à son tour des valeurs précises de $\xi^l(\mathbf{x})$. Ceci repose sur l'hypothèse que la fréquence locale est la même pour l'image s_A et l'image s_B , et évite l'étape d'approximation numérique des dérivées partielles de la phase.

L'algorithme d'estimation multi-résolutions proposé dans [24] utilise également un seuillage sur l'amplitude des coefficients quaternioniques pour éliminer les déplacements qui concerneraient des structures négligeables (la phase étant invariante par changement d'intensité), et privilégie les sous-bandes verticales (resp. horizontales) pour l'estimation de d_{x_1} (resp. d_{x_2}). L'information extraite à une échelle est utilisée à l'échelle plus grossière pour le choix des constantes k_1 et k_2 . Cette méthode exploite donc pleinement la structure de décomposition en ondelettes quaternioniques, à l'exception de l'angle ψ qui n'est pas utilisé. Les résultats expérimentaux sont concordants avec notre perception [15, 24], même si les déplacements aux échelles grossières semblent plus difficiles à obtenir. L'avantage de cette méthode est dans la rapidité de calcul, grâce au calcul efficace des dérivées et à la compacité de la représentation par rapport à une approche « à la Gabor ».

2.3.2 L'angle ψ

Nous avons vu à la section 2.2.1 (autour de l'équation (2.21)) que ψ prend des valeurs extrêmes $\pm\pi/4$ si la structure locale est une onde plane, et reste autour de 0 sinon.

T. Bülow a montré que cette information permet d'enrichir le procédé classique de segmentation de texture à l'origine basé sur des filtres de Gabor complexes [12]. La méthode consiste à mesurer les variations de phase pour discriminer les différentes textures, et le fait d'avoir trois angles dans le cas quaternionique au lieu d'un dans le cas complexe apparaît comme un avantage. Finalement, le troisième angle ψ apporte une information expérimentalement utile pour l'analyse de textures.

Une application proposée dans [24] utilise les trois angles de QWT pour estimer une représentation en « wedgelets » de l'image. L'algorithme découpe l'image en blocs dyadiques, associés à trois coefficients QWT $w_{LH,j}$, $w_{HL,j}$ et $w_{HH,j}$, dans lesquelles on estime le segment de droite (wedgelet) qui correspond le mieux à la structure locale. L'estimation revient à trouver les deux paramètres de la droite à partir des trois coefficients.

L'orientation globale - horizontal, vertical ou diagonal - est obtenue en repérant le plus fort des trois coefficients. Si l'orientation est plutôt diagonale, le signe de ψ permet de savoir laquelle des deux diagonales possible est concernée. Une fois l'orientation connue, les angles φ et θ donnent la position de la droite par rapport au centre du bloc. Pour raffiner les mesures, les auteurs ont préalablement généré des valeurs de phase pour une séquences d'images de synthèse dont l'orientation est connue, puis établi une relation expérimentale par ajustement de courbe (curve fitting). Les détails sont présentés dans [24].

Nous venons de recenser ce qui est proposé dans la littérature concernant l'utilisation des ondelettes quaternioniques. L'invariance de l'amplitude aux translations, l'invariance de la phase aux changements d'intensité, et le sens physique des valeurs de la phase, font de la QWT

une amélioration claire des ondelettes séparables pour une redondance raisonnable de $4\times$. Afin d'analyser plus finement les qualités et limites de cette nouvelle représentation, nous proposons de l'incorporer dans deux cadres applicatifs faisant appel à des possibles qualité de parcimonie et de description de l'information.

2.4 ETUDE DE LA PARCIMONIE DE LA QWT

La notion de parcimonie [102, 139] se développe depuis quelques années dans des domaines du traitement du signal et des images tels que la compression, la détection, la séparation de sources et le débruitage. Il s'agit d'un outil mathématique qui permet de mesurer la puissance de représentation d'un certain « dictionnaire » constitué d'atomes. Suite à l'essor en image des transformées en ondelettes orthogonales (non-redondante), cette nouvelle notion a amené les chercheurs à s'intéresser à des représentations plus *redondantes* - c'est-à-dire liées à un dictionnaire plus riche. Il vient que certaines de ces représentations redondantes sont plus *parcimonieuses*, en ce sens qu'elles concentrent l'information utile dans un nombre plus réduit de coefficients. On peut alors définir de nouvelles méthodes de compression, mais aussi et plus généralement analyser ou manipuler les éléments d'un signal de façon plus efficace. Ce gain de qualité de représentation, obtenu par le compromis d'un dictionnaire plus grand, est analogue à la capacité qu'à une langue riche d'exprimer beaucoup d'idées en peu de mots. La contrepartie est que les « systèmes communicants » doivent *connaître* un grand dictionnaire, en d'autres termes, les algorithmes de décomposition et de reconstruction du signal sont plus coûteux.

Prenons par exemple l'algorithme JPEG-2000 [2]. La transformée biorthogonale séparable (DWT) est utilisée pour sa non-redondance et l'orthogonalité de sa base, qui en font un outil intéressant d'un point de vue compactage des données et d'un point de vue statistique. D'un autre côté, la rigidité de cette transformée la rend inapte à représenter efficacement les contours, car elle est optimale pour les singularités ponctuelles. De plus, plusieurs travaux montrent l'efficacité des représentations *redondantes* [50, 154], lorsque l'on dispose d'un petit budget de données et que l'on s'autorise à avoir quelques dégradations. Dans ce contexte (« lossy compression »), les transformées redondantes offrent un codage de meilleure qualité, car les coefficients sont par exemple invariants par translation, ou mieux localisés, et correspondent à une reconstruction de l'image par des structures élémentaires plus adaptées et permettant de mieux restituer les contours (moins d'oscillations, moins d'effets rectangulaires...).

La QWT se place clairement dans ce contexte car elle est redondante par $4\times$. Contrairement à d'autres représentations redondantes comme les pyramides ou les ondelettes non-décimées, la QWT n'implique pas un « plus grand nombre » de coefficients qui augmenterait la résolution des sous-bandes, mais plutôt augmente la *dimension* de ces coefficients, à la manière des coefficients de Fourier, qui deviennent des vecteurs dont l'amplitude et la phase s'interprètent comme des descripteurs structurels plus riches, incluant la notion de résolution fine. On s'attend à ce que la QWT soit *plus parcimonieuse* que la DWT, c'est-à-dire permette des reconstructions partielles plus acceptables, pour une quantité équivalente de données conservées.

Ce travail ne prétend pas présenter un schéma de compression d'image mais bien d'appréhender la parcimonie de la QWT, par des reconstructions partielles, dans une approche comparative par rapport à la transformée en ondelettes (DWT), qui est la plus utilisée en image. En effet la définition d'un schéma de compression complet nécessite l'utilisation de plusieurs « blocs » comme la transformée (DCT, ondelettes *etc.*), la quantification, un algorithme de codage spécialement optimisé pour la transformée utilisée (EZW, SPIHT, SPECK *etc.*), le codage entropique, et un travail global très important d'optimisation, ce qui sort de notre sujet : l'étude des transformées. L'étude de la parcimonie reste néanmoins fondamentale pour déterminer le potentiel d'une transformée face à différentes applications comme la compression d'image. Dans toute la partie expérimentale sur la QWT, nous appellerons « DWT » (pour discrete wavelet transform) la transformée biorthogonale séparable munie des filtres Daubechies 9/7, qui fera référence.

Nous allons mettre la DWT et la QWT à l'épreuve de la quantification des coefficients (au sens nombre de bits alloués pour représenter l'information), étape incontournable en compression et significative de la robustesse de la transformée. Les tests seront volontairement très simples, pour ne pas noyer les résultats dans des justifications trop diverses, et se concentrer sur la transformée.

2.4.1 Codage de l'amplitude

Dans un premier temps, nous allons observer l'effet de la quantification de l'amplitude A . Nous avons déjà dit que l'amplitude QWT était analogue au module de la DWT, ce test est l'occasion de voir si l'invariance de la QWT et l'information de la phase apportent quelque chose en termes de qualité visuelle.

Nous appliquons une quantification *uniforme* sur K valeurs de l'amplitude A de la façon suivante :

$$A \leftarrow \text{round} \left[\frac{K-1}{\max(A)} \times A \right] \frac{\max(A)}{K-1} \quad (2.42)$$

La phase QWT ainsi que le signe des coefficients DWT sont gardés intacts pour reconstruire l'image. Les quelques résultats présentés à la figure 2.8 montrent l'avantage très clair de la QWT, qui confirme que la phase contient une information structurelle bien plus riche que le signe DWT. Les oscillations introduites par la DWT autour des contours (**monarch**, **face**) sont remplacés par une structure géométrique bien préservée par la QWT. En revanche, on remarque que les textures (comme la surface de l'eau dans l'image **sailing1**) sont très lissées par la QWT, alors que la DWT arrive à reproduire de la texture par des oscillations. Cela est dû à l'orthogonalité de la DWT, qui permet de mieux modéliser les structures aléatoires comme les textures.

Les courbes de type « débit-distortion » (dernière colonne figure 2.8) confirment également la supériorité de la QWT en termes de PSNR, mais nous ne nous concentrerons pas sur cette mesure qui ne reflète pas de façon complète la qualité perçue (En témoignent les reconstructions de l'image **sailing1** qui ont le même PSNR). Nous privilégierons donc en général les observations qualitatives. Parmi toutes les images que nous avons pu tester, la DWT n'a jamais vraiment montré de résultat supérieur à la QWT, excepté pour des images où certaines textures sont très importantes comme l'image **sailing1**.

Grâce à l'information complémentaire portée par la phase, nous avons constaté que l'amplitude de la QWT peut être quantifiée de façon plus dure que la DWT. Quantitativement, il est possible d'économiser en moyenne environ 1 bit par coefficient. Voyons maintenant le cas de la phase.

2.4.2 Codage de la phase quaternionique

Le problème de quantification pour des données angulaires comme la phase QWT n'est pas classique. Si des travaux sur la modélisation statistique de variables circulaires existent (distributions de Von Mises par exemple), la communauté image et compression n'est pas très riche en ce qui concerne la donnée « phase ».

On peut néanmoins trouver quelques références sur le codage de la phase de la transformée de Fourier [119], où il est montré que la finesse nécessaire pour le pas de quantification des phases est inversement proportionnelle à la fréquence concernée. Cela peut s'expliquer d'une part par le fait qu'une erreur de Δ_φ sur une phase φ s'interprète comme une erreur de position de $\Delta_x = \Delta_\varphi / \omega$ dans le domaine spatial (voir théorème du décalage) - donc minime si la fréquence ω est grande ; et d'autre part par l'importance réduite des hautes fréquences dans les images naturelles qui, comme on le sait, concentrent majoritairement leur énergie dans les basses fréquences.

Notons qu'un travail a également été conduit pour la modélisation statistique des « différences

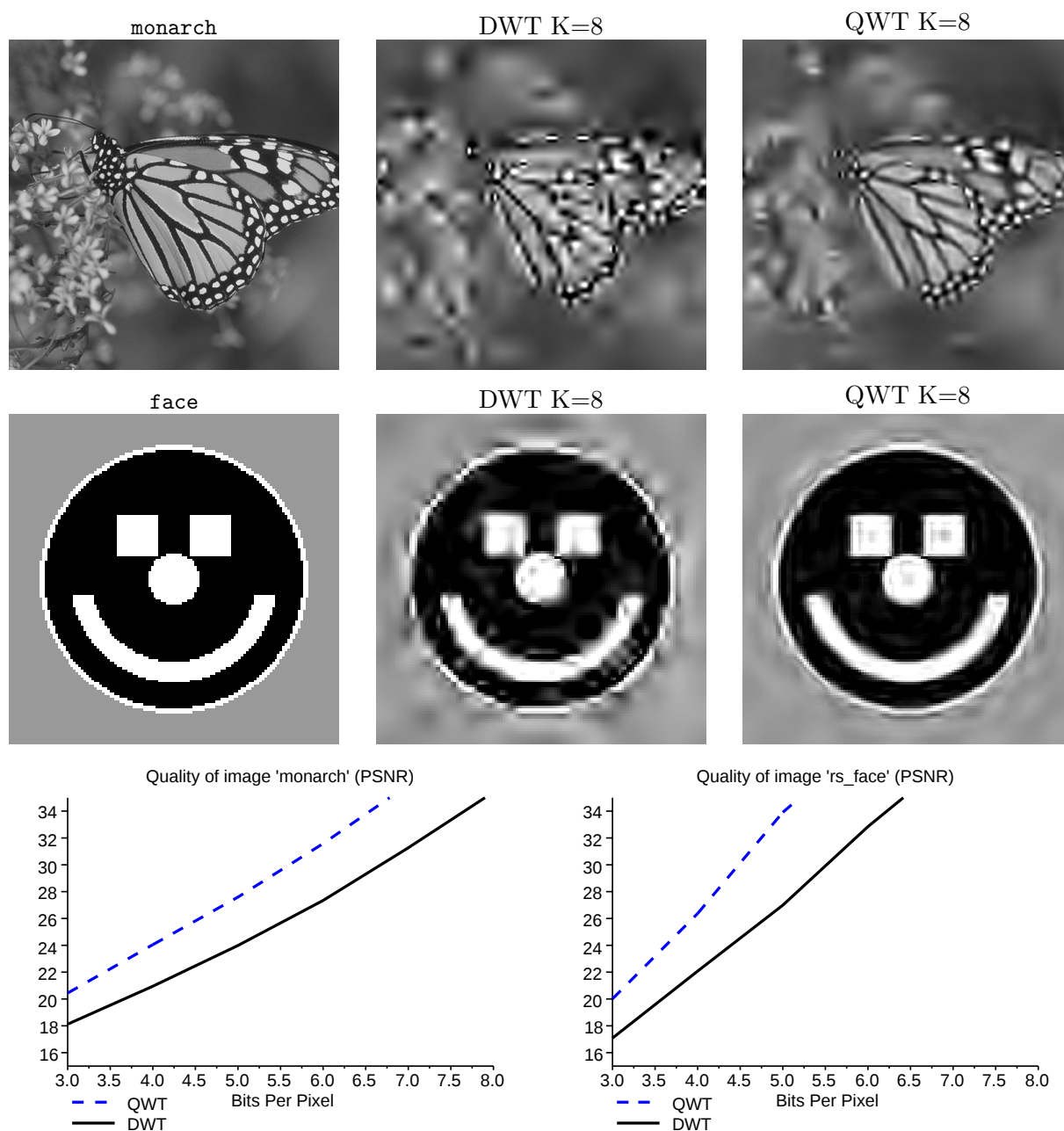


FIGURE 2.8

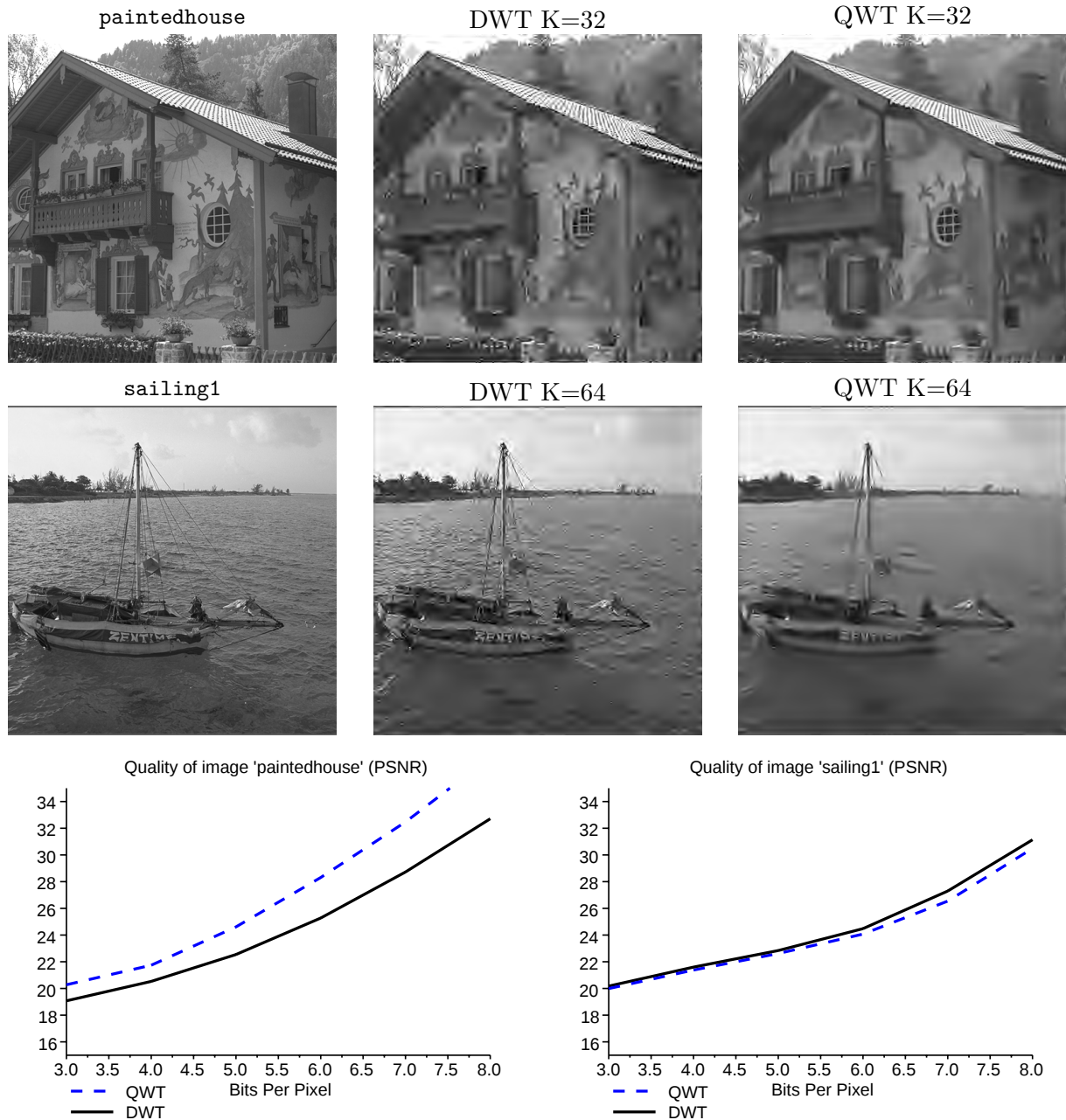


FIGURE 2.8 – Effet de la quantification uniforme de l'amplitude de la DWT/ QWT. L'abscisse des courbes débit-distortion est en bits par pixels, correspondant à $\log_2(K)$.

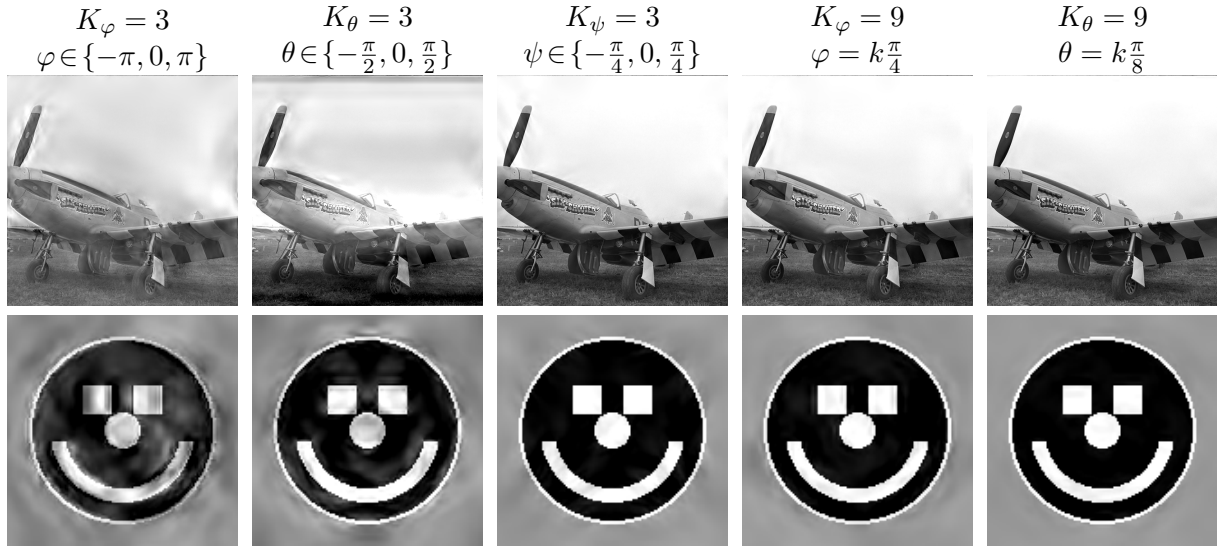


FIGURE 2.9 – *Effet de la quantification uniforme de la phase QWT. Chaque reconstruction correspond à la quantification d'un seul des trois angles.*

locales de phase » dans le domaine des ondelettes complexes [152], mais cela concerne uniquement l'extraction de descripteur, et ne peut pas être utilisé directement pour le codage².

Dans le cadre de ce travail, le but est d'appréhender la qualité de l'information conservée par quantification des composantes quaternioniques. Nous allons donc définir une quantification uniforme des trois angles de phase. L'idée est de visualiser l'effet de la quantification, ce qui correspond à reconstruire l'image après avoir *arrondi* les valeurs. Cet arrondi est plus ou moins grossier, selon le nombre d'intervalles choisi pour cette quantification. Nous avons choisi dans un premier temps de fixer un nombre d'intervalles différent pour chaque angle : K_φ , K_θ et K_ψ . L'effet de la quantification de chacun est défini par :

$$\varphi \leftarrow \text{round} \left[\frac{K_\varphi - 1}{2\pi} \times \varphi \right] \frac{2\pi}{K_\varphi - 1} \quad (2.43)$$

$$\theta \leftarrow \text{round} \left[\frac{K_\theta - 1}{\pi} \times \theta \right] \frac{\pi}{K_\theta - 1} \quad (2.44)$$

$$\psi \leftarrow \text{round} \left[\frac{K_\psi - 1}{\pi/2} \times \psi \right] \frac{\pi/2}{K_\psi - 1} \quad (2.45)$$

Notons que les intervalles de quantifications sont répartis sur les domaines de définitions respectifs des trois angles : $[-\pi; \pi]$, $[-\frac{\pi}{2}; \frac{\pi}{2}]$ et $[-\frac{\pi}{4}; \frac{\pi}{4}]$. Dans cette partie préliminaire, la circularité des données n'est pas exploitée, ce qui implique que les deux intervalles de quantifications proches des valeurs π et $-\pi$ seront superposés, et donc redondants.

Expérimentalement, pour des valeurs de K correspondant à un pas de quantification de $\pi/16$, la reconstruction est visuellement parfaite. La figure 2.9 montre des reconstructions pour des valeurs de K plus faibles. On constate qu'une quantification assez dure peut être appliquée sans dégradation majeure sur la reconstruction. Ce point est très positif d'un point de vue codage car cela signifie que le redondance de la transformée pourrait être largement compensée par un codage très léger de la phase - qui constitue trois quart des données. Afin d'étudier la phase plus finement, il peut être intéressant d'étudier la distribution statistique des angles. Pour cela, nous avons étudié les histogrammes de φ , θ et ψ dans chaque sous-bande QWT. Nous devons rappeler que les phases ne sont pas toutes pertinentes, selon l'amplitude associée. En effet, les phases liées à des coefficients quaternioniques de faible amplitude ont d'une part peu d'importance en termes d'information portée, et d'autre part risquent d'introduire un biais

2. Les références [120, 146, 105], que nous n'avions pas remarquées jusqu'à maintenant, traitent de la quantification polaire pour le codage de signaux audio et mériteraient d'être étudiées.

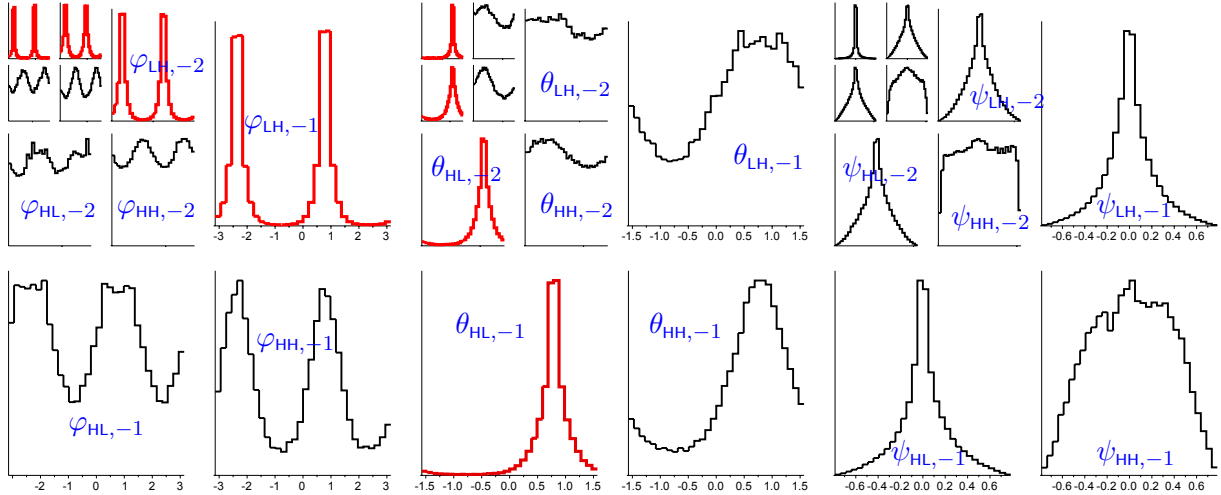


FIGURE 2.10 – Histogrammes moyens des angles de phase QWT dans chaque sous-bande. Les courbes sont disposées comme dans l’affichage des sous-bandes QWT figure 2.6. Les coefficients ont été sélectionnés selon un seuillage sur l’amplitude avec $T = 0.02$.

à cause d’éventuelles instabilités numériques, liées aux petites valeurs de leurs composantes cartésiennes. C’est pourquoi nous proposons d’utiliser une procédure de sélection des coefficients, à travers un seuillage, afin d’alimenter le calcul d’histogramme avec une grande majorité de données *pertinentes*. La procédure de calcul est donc la suivante :

- Pour chaque échelle $i \in \{-1, -2, -3\}$, faire :
 - Pour chaque sous-bande $k \in \{HH, HL, LL\}$, faire :
 - sélectionner les coefficients dont l’amplitude normalisée $A_{k,i} / \max(A_{k,i})$ dépasse le seuil T ;
 - extraire les valeurs de phase de ces coefficients $\varphi_{k,i}$, $\theta_{k,i}$ et $\psi_{k,i}$;
 - calculer les histogrammes correspondants $h_{\varphi_{k,i}}$, $h_{\theta_{k,i}}$ et $h_{\psi_{k,i}}$;

Lorsque $T = 0$, les distributions obtenues expérimentalement sont la plupart du temps uniformes. En augmentant ce seuil, nous avons constaté que les distributions devenaient plus singulières, alors même que le nombre de coefficients considérés était toujours substantiel. Cela confirme que les phases non pertinentes noient l’information recherchée si on ne les élimine pas du calcul des histogrammes. Nous avons retenu le seuil $T = 0.02$ qui semblait suffisant pour éliminer le bruit tout en permettant de conserver environ 25% des coefficients dans chaque sous-bande.

Dans un second temps, nous avons observé que les distributions de la phase quaternionique étaient très semblables d’une image à l’autre. C’est pourquoi, afin de rendre les données plus robustes, nous proposons de calculer dans chaque sous-bande un histogramme *moyen*, à partir des histogrammes obtenus avec une base d’images naturelles. Nous avons choisi la base LIVE [125], qui contient des photographies variées.

L’ensemble des histogrammes moyens avec sélection des coefficients dans chaque sous-bande QWT est illustré figure 2.10. La première remarque est que les distributions obtenues sont soit uniformes, unimodales ou bimodales, ce qui est intéressant car des modèles classiques peuvent être appliqués.

Plus particulièrement, on constate que les angles φ et θ ont une distribution assez « dégénérée » dans les sous-bandes concernant les détails respectivement horizontaux et verticaux. Ces distributions ont été coloriées en rouge dans l’illustration. Cela rejoint l’observation déjà faite à propos de l’estimation de flot optique, à savoir que l’information de « déplacement local » horizontal (resp. vertical) portée par φ (resp. θ) est ambiguë si elle concerne une structure localement horizontale (resp. verticale). En effet, quel que soit son décalage vers la droite ou la gauche, une structure horizontale ressemblera toujours à une structure horizontale, la donnée de décalage n’est donc pas une information significative. En conséquence, la phase prend quasiment systématiquement la même valeur : $\varphi \approx \pi/4 \pm \pi/2$ pour les structures horizontales, $\theta \approx \pi/4$

pour les structures verticales. A ce jour néanmoins, nous ne savons pas expliquer pourquoi ces valeurs *particulières* sont révélées.

Le cas de la première échelle (la plus fine) est vraiment particulier car les structures encodées par les coefficients ont un support spatial de l'ordre de 4 pixels, et sont donc éloignées du modèle *continu* associé à la représentation quaternionique. Ceci est à notre avantage car il y a de grandes chances pour que la phase de la première échelle puisse être quantifiée très grossièrement sans influence sur la reconstruction.

Finalement, devant la variété de cas possible à propos des valeurs de la phase QWT, il paraît nécessaire de mettre en place une quantification *adaptive*.

2.4.3 Schéma complet de quantification

Nous présentons ici une proposition qui s'inscrit dans le cadre du projet ANR CAIMAN, auquel cette thèse est associée. Ce projet a pour but d'étudier un nouveau codeur d'images prenant en compte conjointement les contraintes de réseau, de sécurité, et de qualité, à travers une méthode de compression « scalable », c'est-à-dire dans laquelle les informations visuelles sont hautement hiérarchisées. Ce dernier aspect nécessite une méthode efficace de représentation de l'information visuelle, ce qui est l'objet de cette thèse. Nous pensons que les transformées en ondelettes analytiques sont prometteuses pour ce genre d'application, c'est pourquoi nous avons proposé, comme une première étape vers un codage hiérarchisé de l'information, un schéma de quantification adaptatif des coefficients d'ondelettes quaternioniques, spécialement adapté à la QWT.

L'idée est d'exploiter les connaissances que nous avons sur la phase pour quantifier grossièrement là où on suppose qu'il n'y aura pas de conséquence sur la reconstruction, et plus finement pour les coefficients importants. Plusieurs paramètres entrent en jeu.

- Le premier est un paramètre de *pertinence* des valeurs, directement lié à l'amplitude A . Les coefficients de faible amplitude codent une structure *peu visible*, et donc ne requièrent pas une grande précision sur la phase. A plus forte raison, la quantification de l'amplitude aura un effet de seuillage, soit la mise à zéro de tous les coefficients en dessous du pas de quantification choisi pour A . Dans ce cas, la phase ne sera simplement pas codée.
- Le second est un paramètre d'*échelle*. Nous avons vu qu'à la première échelle, la phase ne porte que très peu d'information. On décidera donc de la quantifier plus grossièrement que pour les échelles suivantes.
- Nous proposons enfin de prendre en compte les valeurs particulières révélées par les histogrammes, afin que la quantification des phases soit centrée autour de ces valeurs. Ainsi, de *nombreux* coefficients seront bien restitués malgré une quantification dure.

Etant donné un nombre de valeur K_A pour la quantification de l'amplitude A , on considère le schéma particulier suivant pour la quantification de la première échelle :

	$\varphi \in$	$\theta \in$	$\psi \in$
LH	$\{-\frac{3\pi}{4}; \frac{\pi}{4}\}$ (1 bit)	$\{-\frac{\pi}{4}; \frac{\pi}{4}\}$ (1 bit)	$\{0\}$ (0 bit)
HL	$\{-\frac{3\pi}{4}; -\frac{\pi}{4}, \frac{\pi}{4}, \frac{3\pi}{4}\}$ (2 bits)	$\{\frac{\pi}{4}\}$ (0 bit)	$\{0\}$ (0 bit)
HH	$\{-\frac{3\pi}{4}; -\frac{\pi}{4}, \frac{\pi}{4}, \frac{3\pi}{4}\}$ (2 bits)	$\{-\frac{\pi}{4}; \frac{\pi}{4}\}$ (1 bit)	$\{-\frac{\pi}{8}; \frac{\pi}{8}\}$ (1 bit)

(2.46)

Cela donne pour chaque coefficient de la première échelle dont l'amplitude dépasse le premier intervalle de quantification un codage sur un total de $\log_2(K_A)$ bits pour A et 8 bits pour la phase. Les « petits » coefficients sont donc codés sur seulement $\log_2(K_A)$ bits.

Pour les échelles suivantes, on utilisera une quantification plus fine :

	$\varphi \in$	$\theta \in$	$\psi \in$
LH	$\{-\frac{3\pi}{4}; -\frac{\pi}{4}, \frac{\pi}{4}, \frac{3\pi}{4}\}$ (2 bits)	$\{k \frac{\pi}{2^{1+ j }}\}$ $(1+ j)$ bits)	$\{-\frac{\pi}{4}; 0; \frac{\pi}{4}\}$ (2 bits)
HL	$\{k \frac{\pi}{2^{1+ j }}\}$ $(2+ j)$ bits)	$\{-\frac{\pi}{4}; \frac{\pi}{4}\}$ (1 bits)	$\{-\frac{\pi}{4}; 0; \frac{\pi}{4}\}$ (2 bits)
HH	$\{k \frac{\pi}{4}\}$ (3 bits)	$\{k \frac{\pi}{4}\}$ (2 bits)	$\{k \frac{\pi}{2^{ j }}\}$ ($ j $ bits)

(2.47)

Les coefficients aux échelles plus grossières consommeront donc $\log_2(K_A) + 3|j| + 15$ bits, lorsque leur amplitude dépasse le premier intervalle. Nous considérons que les angles $(\varphi_{HL}, \theta_{LH}, \psi_{HH})$ sont les plus importants car ils codent respectivement la position horizontale des détails verticaux, la position verticale des détails horizontaux, et « l'isotropie » des détails diagonaux. Ils se voient donc quantifiés de plus en plus finement avec l'échelle. Nous nous sommes attachés à inclure dans les valeurs quantifiées les valeurs majoritaires selon les histogrammes, afin de minimiser statistiquement l'impact de la quantification.

Notons que dans ce test le nombre total de bits alloués dépend du nombre de coefficients dont l'amplitude est sous le premier intervalle - c'est-à-dire dépend du contenu de l'image. Le nombre effectif de bits alloués pour la quantification d'une image n'est calculé et donc connu qu'à la fin du processus.

Afin de comparer avec la DWT, une astuce est nécessaire car nous souhaitons observer les deux reconstructions d'une même image ayant été codée sur le même nombre de bits par chaque transformée. En plus du fait que ce nombre ne soit pas prévisible par les valeurs de $K_{...}$, la DWT ne possède pas de phase, et donc pas le même nombre de données de départ. Nous avons donc utilisé une méthode itérative qui ajuste le seuil en fonction d'un débit fixé. Nous avons ainsi pu expérimentalement obtenir des nombres de bits très proches. Il faut juste noter que pour cela, le pas de quantification du module est probablement légèrement différent pour la DWT et la QWT. Globalement, on observe peu de différence, en remarquant que le nombre de bits alloués pour la phase QWT est comparable à celui utilisé pour coder le *signe* de la DWT - c'est-à-dire très faible.

Des résultats comparatifs sont affichés figure 2.11. Les comparaisons que nous faisons sur la base LIVE donnent largement l'avantage à la QWT pour un nombre de bits par pixels inférieur à 6 environ, comme illustré par les courbes débit-distortion des images **monarch** et **sailing2**. Ces courbes sont données à titre indicatif mais sont à considérer avec une certaine « distance » puisque la mesure de PSNR n'est pas toujours corrélée avec la qualité perçue des images. C'est pour cela que malgré l'apparente supériorité de la DWT pour un budget de bits plus grand, la différence visuelle entre les images reconstruites par les deux méthodes (voir les reconstructions de **monarch** à 7.58 bpp) est négligeable en pratique.

On voit clairement que la QWT permet de mieux restituer les contours, à la manière de ce qu'on a pu voir avec la quantification de l'amplitude seule (figure 2.8). Par exemple, les motifs sur l'aile du monarque à 4.08bpp, ou bien le mât du bateau, sur lequel nous avons zoomé, ne présentent pas avec la QWT les oscillations indésirables produites par la DWT. Cette fois-ci, la comparaison est équitable puisque la phase est également codée, et pour cette raison l'amélioration est aussi moins marquante. La capacité de la phase QWT à restituer des structures précises en utilisant peu de données binaires permet donc de surpasser la DWT. Les distortions introduites par la QWT ne sont pas dérangeantes à l'œil, contrairement aux artefacts oscillants de la DWT, qui dénaturent l'aspect de la voile dans l'image **sailing2**. La DWT restitue néanmoins généralement mieux les textures, moins bien reproduites par les atomes lisses de la QWT.

Pour une reconstruction de haute qualité, la DWT devrait être préférée car il n'est plus question d'approximer les structures de façon lisse, mais plutôt de retrouver la valeur exacte de tous les pixels, ce qui n'est pas le but ici.

Une piste d'amélioration pour ce type d'application est sans doute vers une étude statistique plus poussée, qui partirait par exemple d'une hypothèse gaussienne sur les pixels, pour dériver analytiquement les distributions de A , φ , θ et ψ . On pourrait alors optimiser statistiquement la quantification avec des outils type Lloyd-Max. Une autre piste serait d'exploiter la dépendance inter-échelle des coefficients, d'autant plus qu'il existe probablement une dépendance à la fois de l'amplitude et de la phase. Un stage de Master est en cours dans notre laboratoire sur ce sujet.

Pour conclure sur le codage par QWT, cette nouvelle transformée légèrement redondante par une notion de phase portant des informations géométriques surpasse la représentation très utilisée par ondelettes séparables grâce à des atomes moins oscillants et mieux localisés, dont les coefficients quaternioniques peuvent être codés de façon efficace par peu de données. Notons que

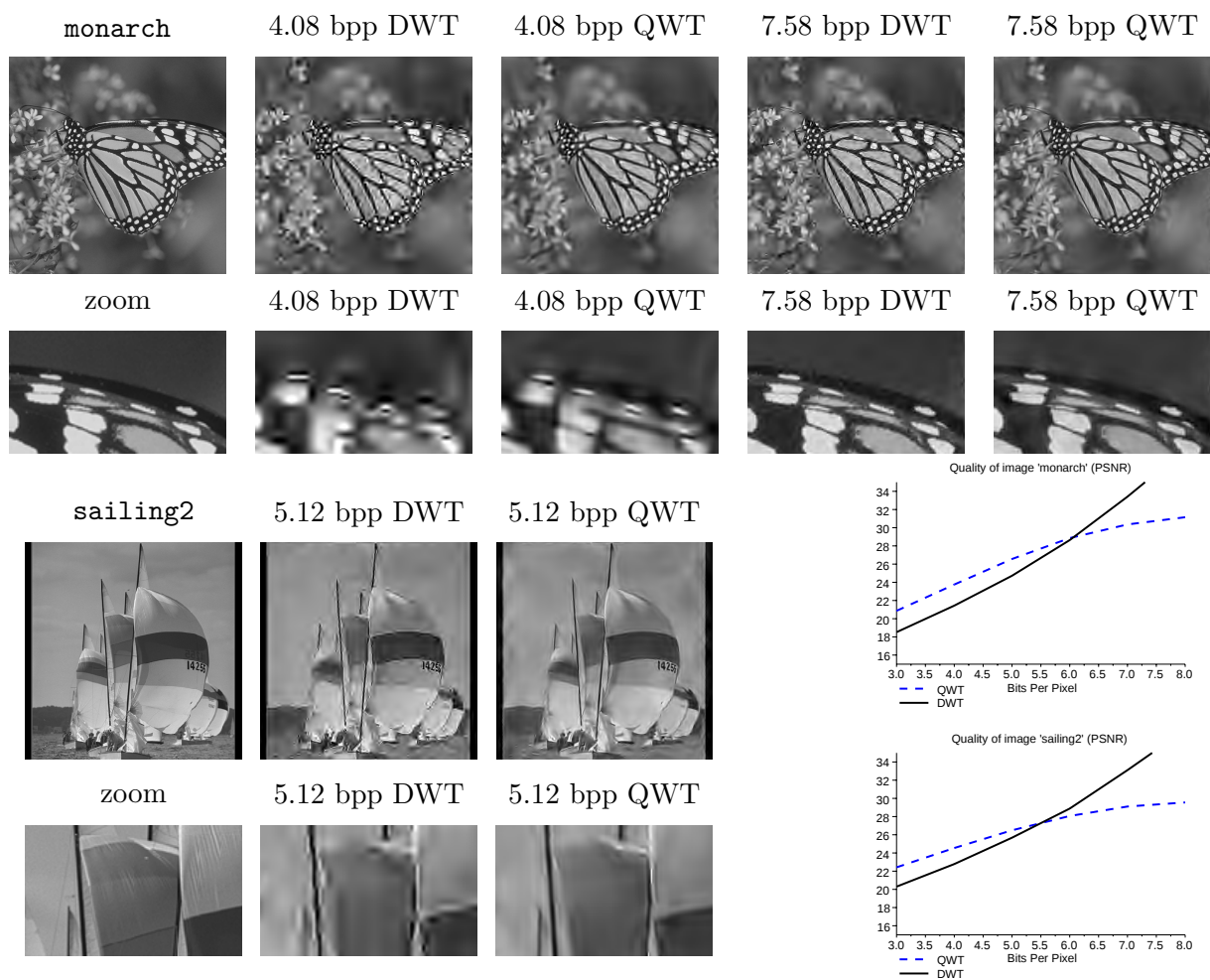


FIGURE 2.11 – Résultats comparatifs de quantification DWT vs. QWT, à quantités de données égales. « bpp » (bits per pixel) est le nombre moyen de bits utilisés pour un pixel de l'image.

nous avons proposé la seule application de la littérature qui utilise la reconstruction de la QWT, et donc la puissance de représentation des images en termes d'atomes quaternioniques. Voyons maintenant une autre application classique des ondelettes, la classification de textures, pour laquelle le troisième angle de la QWT pourrait constituer une nouvelle information discriminante permettant d'améliorer les méthodes classiques.

2.5 MISE À L'ÉPREUVE EN CLASSIFICATION DE TEXTURES

Suite à la segmentation de texture proposée par T. Bülow [15], nous proposons d'analyser ce que la QWT peut apporter en *classification* de textures. Cette dernière diffère de la segmentation en ce que les textures sont sur des images différentes qu'il faut discriminer les unes des autres. De plus, la classification implique souvent de grandes bases de données, ce qui ajoute une contrainte de performance de l'algorithme, et justifie l'utilisation d'une transformée en ondelettes peu redondante comme la DWT ou la QWT plutôt qu'une décomposition de Gabor. Le cœur de la méthode réside dans l'extraction d'un vecteur descripteur pour chaque image, afin de mesurer les ressemblances/différences par des distances dans des espaces vectoriels. Nous pensons que l'invariance par translation de l'amplitude et la richesse de l'information géométrique contenue dans la phase font de la QWT un outil potentiellement efficace pour cette tâche. Comme précédemment, la recherche ne va pas porter sur l'optimisation des différents blocs entourant la transformée, mais plutôt sur la mise en valeur du pouvoir de caractérisation des mesures quaternioniques.

Le travail est de trouver automatiquement, à partir d'une image texturée, la *classe* à laquelle elle appartient la plus vraisemblablement. La notion de « texture » possède différentes définitions dans la littérature, on s'accordera à la présenter par des exemples classiques comme des photographies macroscopiques de goudron, surface de l'eau, sable, ou des vues satellitaires de forêt, zone urbaine, océan - tous caractérisés par une sorte de complexité relativement uniforme à l'allure plus ou moins périodique. On appelle « classe » une sorte de texture, selon une classification que l'on peut faire visuellement de façon intuitive.

L'extraction de caractéristiques texturales sous forme de *descripteurs* à partir des représentations en ondelettes est classique depuis une vingtaine d'années [143, 3, 91]. À partir de chaque sous-bande, une ou plusieurs mesures globales comme des mesures statistiques ou d'énergie sont extraites, et stockées dans un vecteur plus ou moins grand que l'on appelle *descripteur*. On se restreint donc à la famille de méthodes basées ondelettes, puisque c'est notre objet d'étude, mais d'autres techniques d'analyse de texture existent, notamment par les matrices de co-occurrences [66].

Dans notre cas il est important de noter que les ondelettes complexes dual-tree 2D [124] ont été utilisées pour la classification de textures [36, 21, 22]. D'après les auteurs, l'invariance par translation de l'amplitude et la finesse d'orientation des sous-bandes $((1 + 2k) \times 15^\circ)$ permettent de surpasser les ondelettes classiques grâce à des descripteurs plus stables et plus riches. Dans un souci de positionnement dans la littérature, nous ajouterons dans nos expérimentations les descripteurs CWT de [21, 22], même si l'approche reste très différente puisque la phase n'est pas utilisée dans ces références.

La QWT peut être vue comme une extension acceptable des ondelettes analytiques 1D, contrairement à la CWT qui correspond plutôt à une décomposition type Gabor. En ce sens nous proposons d'exploiter cette propriété en utilisant l'amplitude et la phase des sous-bandes QWT pour construire un descripteur de texture meilleur qu'avec les ondelettes classiques.

Les algorithmes utilisés sont celui de Mallat [101] avec les filtres de Daubechies « 9/7 » pour la DWT, et celui de Kingsbury [124] avec les filtres « Q-Shift » pour la QWT. La taille des filtres Q-Shift sera de 9 coefficients par souci d'équité avec la taille des filtres de Daubechies, mais des versions plus longues seront également traitées dans les résultats.

Nous utiliserons des décompositions sur L échelles, produisant $3L$ sous-bandes plus une sous-bande basse fréquence que nous n'utiliserons pas. Il est usuel de considérer que l'information

basse-fréquence - décrivant des structures très lisses - n'est pas importante pour l'analyse des textures.

2.5.1 Bases d'images de textures

Afin d'obtenir des mesures de performance relativement robustes, nous utiliserons trois bases de textures connues :

- La base de Phil Brodatz [14], qui contient 111 photographies de textures diverses de dimensions 640×640 pixels. Nous avons découpé chacune de ces images en 25 imagerie de dimensions 128×128 dont nous considérerons qu'elles appartiennent à la même classe.
- Les bases Outex TC12 et Outex TC14 (<http://www.outex.oulu.fi>), qui ont été créées spécialement pour tester les méthodes de classification de texture. La première contient 24 classes de textures comprenant chacune 380 images de dimension 128×128 , aux orientations et illuminations différentes. La seconde contient 68 classes de 60 échantillons aux illuminations différentes.

2.5.2 Protocole expérimental

Rappelons que le but de ces expérimentations est en partie d'approfondir la compréhension des coefficients QWT. Le schéma complet doit être suffisamment élémentaire pour que l'on puisse identifier l'influence du choix de la transformée et des descripteurs sur la qualité de reconnaissance. C'est pourquoi nous proposons d'utiliser l'algorithme dit des k plus proches voisins (k -ppv, ou k -NN en anglais), qui considère des distances euclidiennes entre les descripteurs à travers une procédure de vote assez rudimentaire.

Etant donné une méthode de calcul d'un descripteur qui sera expliquée plus bas, et qui associe à une image un vecteur $x = [x_1 \ x_2 \ \dots \ x_n]^T$ caractérisant la texture qu'elle contient, voici l'algorithme k -ppv. On constitue d'abord une *base d'apprentissage* faite d'images indexées par $i \in \{1..N_A\}$, et étiquetées par leur classe c_i , en sélectionnant N_A images dans les bases disponibles. Un descripteur x_i est extrait de chacune d'entre elles, et se trouve donc associé à une classe connue c_i . Etant donnée une nouvelle image texturée, dont la classe est *a priori* inconnue, on extrait son descripteur y qui peut être comparé par distances normées ou non avec les descripteurs x_i connus. L'algorithme k -ppv consiste d'abord à calculer toutes les distances $d(x_i, y)$ et à garder les k plus petites. Les k descripteurs connus les plus proches de y nous donnent donc accès à k classes $\{c_{i_1}, \dots, c_{i_k}\}$. Si la méthode d'extraction de descripteur est efficace, alors la distance euclidienne doit être corrélée à la ressemblance visuelle entre les deux images. La classe estimée est donc identifiée à celle qui a le plus d'occurrences dans les k classes retenues.

Pour évaluer la qualité de nos descripteurs, on décompose les images disponibles en deux groupes :

- Une base d'apprentissage de N_A images dont le programme connaîtra la classe ;
- Une base de test de N_T images dont le programme devra estimer la classe.

Le principe du découpage est commenté par la suite. L'évaluation consiste à compter le nombre d'erreur commises sur la base de test, en comparant avec les vraies classes. On considérera que le descripteur est satisfaisant s'il permet d'obtenir un fort taux de reconnaissance avec l'algorithme k -ppv. Nous étudierons également des taux de reconnaissance individuels au sein des classes pour tenter de dégager certains types de textures plus ou moins adaptés à tel ou tel descripteur.

Enfin, il faut noter qu'ajouter des étapes de sophistication comme une sélection automatique de caractéristiques, des normalisations ou autres outils statistiques complexes, permettrait certes d'améliorer la reconnaissance globale mais noierait la contribution du descripteur. Nous ne pourrions alors pas vraiment conclure sur l'intérêt d'utiliser telle ou telle représentation. C'est pourquoi nous sommes plus intéressés par les *différences* de résultats que nous allons obtenir, que par les résultats en eux-mêmes.

Validation croisée

Les résultats d'une telle expérimentation dépendent du choix de la base d'apprentissage. Puisqu'il existe différentes manières de le paramétrer, nous avons choisi le principe de la validation croisée. Afin d'affiner l'évaluation des performances, nous proposons de faire l'expérimentation 100 fois, pour 100 choix de base différents, puis de mesurer la variation des taux de reconnaissance parmi ces 100 essais, pour apprécier leur robustesse vis à vis du choix des bases. L'affichage d'un taux de reconnaissance pour une classe sera donc sous forme d'un rectangle centré autour du taux moyen et dont la hauteur vaudra deux fois l'écart-type du taux. Une ligne verticale traversera également ce rectangle depuis le taux minimum obtenu jusqu'au taux maximum. Voir par exemple sur la figure 2.12. Grâce à cette méthode, les résultats que nous affichons sont très stables, à savoir que lorsque l'on relance 100 nouvelles fois l'algorithme, les taux moyens et les écart-types varient de moins de 1%.

Précisons enfin la procédure de sélection des échantillons de la base d'apprentissage. Pour éviter de privilégier certaines classes, nous utilisons un choix aléatoire de N_A individus dans la base complète. Les images restantes forment la base de test. De plus, aucune de nos images ne fait partie à la fois de la base d'apprentissage et de la base de test.

2.5.3 Extraction des descripteurs

L'extraction de descripteurs est l'algorithme qui calcule une séquence de nombres à partir d'une image, sensés la décrire au mieux. Dans notre cas, il s'agit de caractéristiques de texture. Le descripteur le plus classique parmi les méthodes basées ondelettes est une mesure d'énergie dans chaque sous-bande, qui permet de caractériser le contenu fréquentiel de l'image - très lié par hypothèse au type de texture qu'elle contient. Des mesures statistiques globales comme la moyenne et l'écart-type (ce dernier étant sensiblement le même type de mesure que l'énergie) sont aussi très utilisées. Ici encore, nous nous garderons d'utiliser des méthodes plus complexes, afin de rester dans l'état de l'art.

Mesures sur l'amplitude

Les mesures classiques faites sur des coefficients d'ondelettes réels se transposent naturellement sur l'amplitude des coefficients quaternioniques (ou même complexes). On considère le module du coefficient réel ou quaternionique $A_s[\mathbf{k}]$ de la sous-bande s à la coordonnée spatiale $\mathbf{k}$. La mesure d'énergie relative est donnée par :

$$m_s = \frac{1}{E} \sum_{\mathbf{k}} (A_s[\mathbf{k}])^2 \quad (2.48)$$

où E correspond à la somme des énergies de toutes les sous-bandes excepté la sous-bande basses-fréquences. La mesure m_s s'interprète donc comme la quantité relative d'énergie dans une certaine bande de fréquence.

La mesure d'écart-type est donnée par :

$$m_s = \sqrt{\frac{1}{N} \sum_{\mathbf{k}} (A_s[\mathbf{k}] - \mu)^2} \quad \mu = \frac{1}{N} \sum_{\mathbf{k}} A_s[\mathbf{k}] \quad (2.49)$$

où N est le nombre de coefficients dans la sous-bande s . Expérimentalement, la mesure d'énergie et celle d'écart-type donnent des performances équivalentes. Nous choisirons la mesure d'écart-type qui est indépendante d'un possible offset dans les données.

Mesures sur la phase

Le premier avantage de la QWT est dans l'invariance de son amplitude, et sera mis à l'épreuve avec les descripteurs précédents. Le second avantage est dans cette nouvelle information de phase

qui apporte une analyse plus riche de l'image, par rapport à la DWT. Comme nous l'avons déjà mentionné, la littérature est assez pauvre au sujet d'une mesure globale sur des phases. On notera la « cohérence globale de phase » [10] qui mesure la netteté d'une image à partir de sa phase locale mais dépend d'un modèle mathématique trop restreint pour notre application. Une modélisation des différences locales de phase dans la CWT a été également proposée [152] mais reste spécifique aux ondelettes complexes.

L'intuition nous suggère qu'une mesure globale comme la moyenne ou la variance des angles φ et θ ne formera pas nécessairement une caractéristique intéressante. En effet, l'information φ/θ définit un recalage fin des structures autour de la position du coefficient, ce qui a du sens localement, mais la façon d'étendre ce genre d'information à une mesure globale sur toute une sous-bande n'est pas immédiate. En revanche, le troisième angle ψ , qui discrimine directement les structures *directionnelles* des autres, peut à travers une mesure globale comme l'écart-type traduire une notion de « quantité de structures orientées ». L'utilisation de ψ en segmentation de texture par T. Bülow nous influence dans le même sens, et les quelques expérimentations que nous avons faites à partir de mesures sur φ et θ ont abouti à une reconnaissance très mauvaise (inférieure à 3%).

Nous allons donc pour exploiter la phase QWT en classification de textures utiliser ψ uniquement. Le choix de la mesure globale de ψ est assez intuitif. On imagine facilement qu'une moyenne risque de tendre analytiquement vers 0, puisque $\psi \in [-\pi/4; \pi/4]$, et au vu des histogrammes étudiés dans le cas de la quantification (figure 2.10), il semble que la distribution de cet angle soit relativement symétrique autour de 0. La moyenne ne porte donc pas d'information si elle vaut toujours 0. Une mesure d'énergie a du sens lorsque les données correspondent plus ou moins à une notion d'énergie, ce qui est le cas de l'amplitude des coefficients (lorsque la conservation de l'énergie est plus ou moins satisfaite par la transformée), mais qui s'interprète difficilement dans le cas d'une phase. Finalement, une mesure à la fois simple et adaptée est celle de l'écart-type, qui permettra de mesurer la largeur de la distribution dans chaque sous-bande, témoignant ainsi du profil global des structures présentes : plutôt directionnelles si l'écart-type est grand. La mesure de ψ est donc définie ainsi :

$$m_s = \sqrt{\frac{1}{N} \sum_{\mathbf{k}} (\psi_s[\mathbf{k}] - \mu)^2} \quad \mu = \frac{1}{N} \sum_{\mathbf{k}} \psi_s[\mathbf{k}] \quad (2.50)$$

Notons que ψ ne pose aucun « problème de modulo », puisqu'il est borné à $\pm\pi/4$, et que l'information portée à chacune de ses bornes n'est pas la même : diagonale dans un sens ou dans l'autre, donc pas de « circularité » autour de $\pm\pi/4$. Bien qu'il s'agisse d'un angle, l'écart-type se calcule donc de façon directe.

Une idée supplémentaire est de pondérer la mesure par l'amplitude, afin de privilégier les angles liés à des coefficients importants. Cela permettrait de ne pas parasiter la mesure avec des valeurs insignifiantes liées aux coefficients de petite amplitude. On peut alors modifier la mesure précédente de la façon suivante :

$$m_s = \sqrt{\frac{1}{N} \sum_{\mathbf{k}} w_s[\mathbf{k}] (\psi_s[\mathbf{k}] - \mu_w)^2} \quad \mu_w = \frac{1}{N} \sum_{\mathbf{k}} w_s[\mathbf{k}] \psi_s[\mathbf{k}] \quad w_s[\mathbf{k}] = \frac{A_s[\mathbf{k}]}{\sum_{\mathbf{k}} A_s[\mathbf{k}]} \quad (2.51)$$

Expérimentalement, cette seconde mesure donne de meilleurs résultats que l'écart-type standard, nous retiendrons donc la version pondérée pour la suite. Voyons maintenant comment combiner ceci avec la mesure sur l'amplitude.

Combinaison des mesures

Le descripteur QWT complet doit exploiter les informations d'amplitude et de phase ensemble. Il est commun en classification d'avoir à combiner des données de nature différente - ici, une « dispersion d'énergie » et une « dispersion de phase ». Cela peut produire un manque de

cohérence dans le descripteur lors du calcul des distances dans l'algorithme k -ppv. L'utilisation d'une métrique permet en général d'équilibrer les différences d'ordre de grandeur entre les données.

Une méthode classique est de normaliser chaque élément du descripteur par l'écart-type de cet élément dans la base complète. Nous avons constaté expérimentalement que cet ajustement n'augmente pas significativement les performances. Il apparaît qu'en pratique, nos mesures sont d'un ordre de grandeur similaire et ne posent pas de problème particulier quant à leur cohabitation dans un seul descripteur. Les mesures d'écart-type sur A et d'écart-type pondéré sur ψ seront donc simplement concaténées en un descripteur de taille double sans aucune normalisation.

Mesures par les ondelettes complexes

Il reste à définir les descripteurs basés sur la représentation en ondelettes complexes. Cette partie est un peu marginale dans ce travail car l'approche conceptuelle des ondelettes complexes 2D n'est pas celle du signal analytique et de la QWT. Cependant, l'algorithme étant très similaire, une confrontation pratique peut être intéressante.

Nous reprendrons les descripteurs définis dans [21, 22], mesurant la variance et l'entropie de l'amplitude et la variance de la phase des 6 sous-bandes de la CWT. Notons que les résultats présentés dans ces articles sont meilleurs que ceux exposés dans ce travail, puisque les algorithmes de classification sont optimisés pour maximiser les performances, notamment avec une sélection/pondération des mesures basée sur une analyse en composantes principale. Comme expliqué plus haut, ce n'est pas l'esprit de notre travail.

Voici également quelques remarques sur ces descripteurs. Contrairement à l'angle ψ de la phase quaternionique, la phase des ondelettes complexe est bien une donnée circulaire définie dans $[-\pi; \pi]$. Une mesure de moyenne ou bien d'écart-type sur une telle donnée doit en général gérer cette circularité, à travers une notion de « modulo π ». Cet aspect n'est pas traité dans [22], et les équations indiquent que les écart-types de phase complexe sont calculés indifféremment. En outre, la notion de pondération par l'amplitude que nous avons proposée n'est pas présente dans la définition de ces descripteurs. Etant donné le schéma de classification très sophistiqué qui est utilisé conjointement, notamment la sélection des descripteurs par ACP, et le fait que les poids résultants de cette ACP ne soient pas indiqués, il est difficile d'établir un lien entre ces différents descripteurs et les types de texture mieux reconnues. Il est donc probable que les mesures sur la phase CWT soient éliminées dans les expérimentations de [21, 22].

2.5.4 Résultats expérimentaux

Maintenant que les descripteurs basés ondelettes classiques/complexes/quaternioniques sont définis, étudions les différents résultats de classification.

De nombreux résultats sont obtenus, dépendant de la profondeur de décomposition des transformées L , des descripteurs choisis, du nombre k de voisins considérés dans l'espace des descripteurs, de la taille de la base d'apprentissage N_A et de la base de texture choisie parmi Brodatz, TC12 et TC14.

Les meilleurs résultats sont obtenus avec $L = 3$ et $k = 3$. En ce qui concerne N_A , plus la base d'apprentissage est grande, meilleurs sont les résultats. Cependant, puisque nous nous intéressons aux *différences* de résultats entre les différents descripteurs, on choisira une valeur de N_A telle que les taux de reconnaissance ne soient pas trop élevés - mais tout de même supérieurs à 50%.

Le choix de la taille des filtres Q-Shift pour le dual-tree qui génère la CWT et la QWT, se situe entre 9, 14 et 18 coefficients [80]. Nous n'avons pas observé de différence flagrante de résultats en fonction de la taille des filtres. Cela s'explique en partie par le fait que les filtres de la première échelle (un tiers des mesures pour $L = 3$) sont différents, et ne changent pas lorsque l'on change de filtres Q-Shift. Ensuite, et surtout, les filtres sont tous de longueur importante, ce

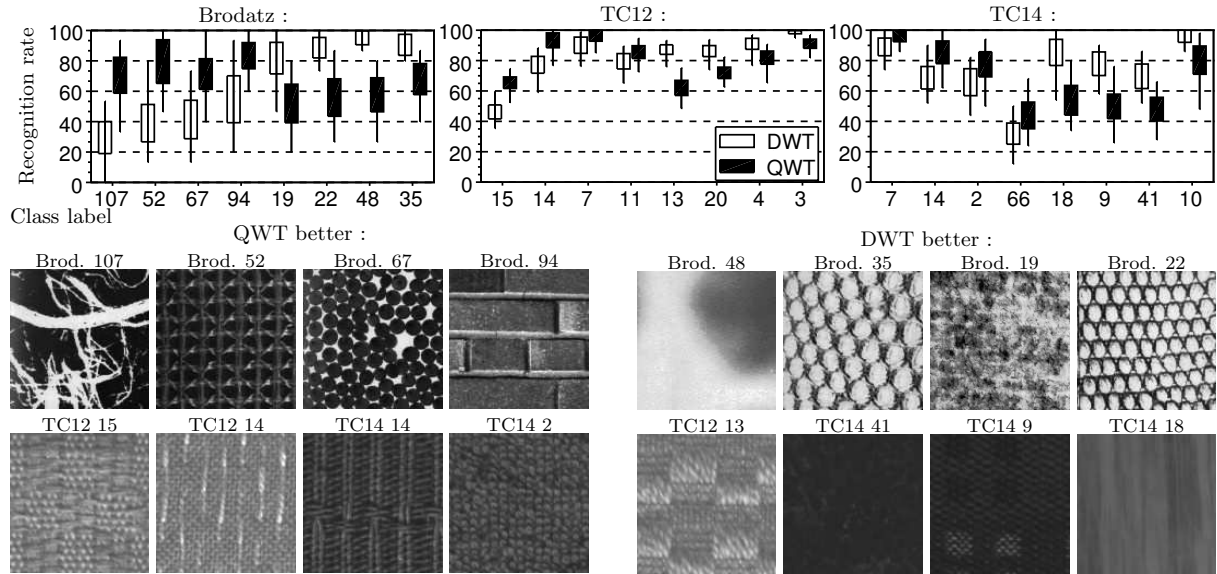


FIGURE 2.12 – Performances de la classification basée sur la mesure de l'amplitude : résultats de validation croisée pour quelques classes des trois bases.

qui assure une bonne séparation fréquentielle. Ceci est au détriment de la localisation temporelle, mais qui est de peu d'importance dans notre cas puisque nous utilisons des mesures globales de sous-bandes.

Finalement, nous nous fixerons les paramètres suivants :

- Nombre d'échelles $L = 3$;
- le descripteur d'amplitude DWT/QWT est une mesure d'écart-types (9 valeurs) ;
- le descripteur de phase QWT est une mesure d'écart-types pondérés (9 valeurs) ;
- le descripteur complet QWT concatène les mesures amplitude/phase (18 valeurs) ;
- le premier descripteur CWT concatène variances et entropies de l'amplitude (36 valeurs) ;
- le second descripteur CWT concatène variances et entropies de l'amplitude et de la phase (72 valeurs) ;
- le nombre de voisins considérés dans le k -ppv est $k = 3$;
- la taille de la base d'apprentissage pour chaque classe est $N_A = 10$ (resp. $N_A = 30$, $N_A = 10$) pour la base de Brodatz (resp. TC12, TC14) ;
- les filtres du dual-tree pour la CWT et la QWT sont les Q-Shift de taille 9.

Les résultats de validation croisée pour chaque classe et chaque descripteur sont recensés en annexe dans la table A.1. Nous allons maintenant les discuter.

Comparaison des mesures sur l'amplitude

Nous commençons par comparer les performances de classification en utilisant seulement les mesures d'écart-type sur l'amplitude de la DWT/QWT. Cela permet de faire une première comparaison totalement équitable, puisque les deux descripteurs en compétition sont analogues et de même dimension.

Une sélection des résultats pour une vingtaine de classes sont illustrés figure 2.12. D'une manière globale, les performances sont équivalentes entre les deux transformées, mais nous avons choisi d'illustrer des classes particulières, pour lesquelles les différences de performance entre DWT et QWT sont les plus grandes, que nous considérons particulièrement significatives pour notre travail.

Nous proposons d'analyser ces résultats à travers ces cas particuliers.

La texture Brod. 67 contient de nombreux petits disques noirs très similaires et disposés de façon apparemment aléatoire. L'invariance par translation de l'amplitude QWT est un atout certain pour decrire de façon stable ce type de texture, là où la DWT a sans doute une réponse

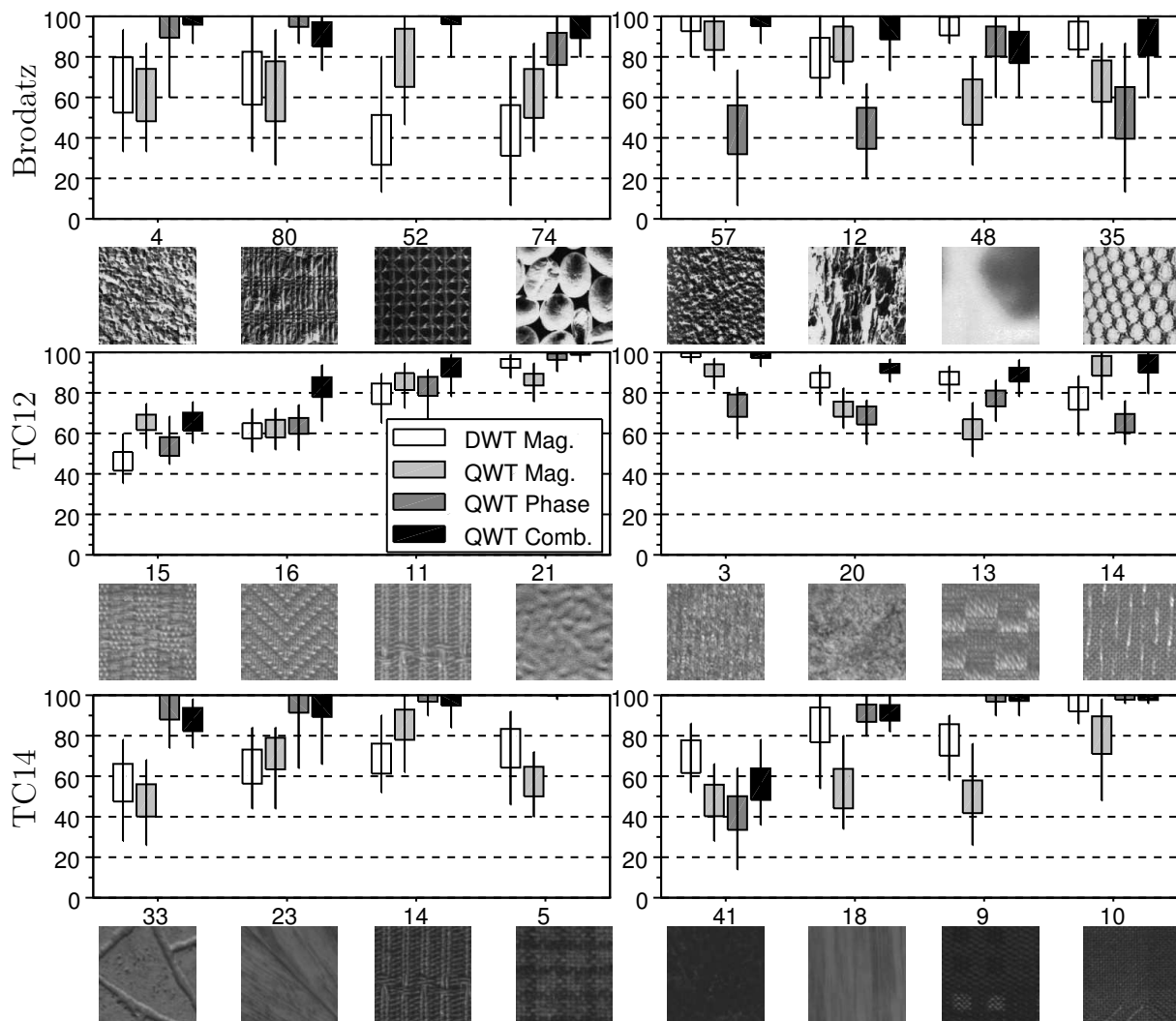


FIGURE 2.13 – Comparaison des performances de reconnaissance pour les mesures basées sur l'amplitude DWT (blanc), l'amplitude QWT (gris clair), la phase QWT (gris foncé) et la combinaison amplitude/phase QWT (noir).

plus variable, ce qui explique très certainement le fort avantage de la QWT pour cette classe de textures.

Les différentes orientations contenues dans les textures semblent décisives pour leur bonne adéquation à la QWT. On remarque que les exemples Brod. 52, Brod. 94, TC12 15, TC12 14, TC14 14 et TC14 2 présentent des structures alignées horizontalement, verticalement et diagonalement. Ces structures aux orientations particulières sont bien prises en compte par l'implantation *séparable* de la QWT, ce qui explique son avantage pour ces classes. A l'inverse, les classes Brod. 35, Brod. 22, et TC14 18 contiennent des structures orientées dans des directions plutôt « intermédiaires », et sont mieux reconnues par la DWT.

Finalement, la seule utilisation de l'amplitude QWT offre des performances similaires à la DWT, avec un avantage en ce qui concerne les textures faites de structures répétées à des positions aléatoires, qui nécessitent l'invariance de la QWT pour une description stable. ainsi que pour les textures « alignées sur la grille ». Voyons maintenant si l'information de la phase permet d'améliorer la reconnaissance.

Contribution de la phase quaternionique

De façon surprenante, les performances de classification basées sur la seule utilisation de la phase QWT - illustrées à la figure 2.13 - donne à nouveau des performances similaires à la

	Brodatz		TC12		TC14	
	Moy.	Ec.-Type	Moy.	Ec.-Type	Moy	Ec.-Type
Amplitude DWT	64%	9	83%	4	56%	8
Amplitude QWT	64%	9	82%	4	51%	8
Phase QWT	65%	8	79%	4	58%	6
QWT Complet	76%	7	91%	4	63%	6
Amplitude CWT	65%	8	87%	4	62%	7

TABLE 2.1 – *Taux de reconnaissance moyens et leurs variations sur les 100 différents choix de base d'apprentissage, pour chaque descripteur et chaque base de texture.*

DWT (cette figure contient également des résultats que nous discuterons plus loin). Ce résultat confirme clairement que l'angle ψ contient une information importante pour la discrimination des textures, ce qui est très encourageant pour l'utilisation conjointe de l'amplitude et de la phase.

En regardant des classes particulières, les résultats sont très hétérogènes. Observons sur la figure 2.13 les performances des descripteurs **QWT Mag.** et **QWT Phase.** (gris clair et gris foncé). Certaines textures, comme Brod. 52, TC12 11 et TC14 14, sont bien détectées à la fois par l'amplitude et la phase. D'autres, comme TC12 20, TC14 41 ou Brod. 35, ne sont bien reconnues ni par l'une ni par l'autre. A l'inverse, de nombreux cas montrent une grande différence de performance entre l'amplitude et la phase (Brod. 80, Brod. 57, TC12 14, TC14 33, TC14 5, TC14 9), ce qui montre que les deux descripteurs sont *complémentaires*. Le premier donne une signature énergétique fréquentielle alors que le second ajoute une information géométrique.

Il est difficile de dégager des types de textures plus ou moins bien reconnues par tel ou tel descripteur, mais la complémentarité entre l'amplitude et la phase QWT est bien présente. C'est pourquoi nous proposons maintenant de combiner les deux descripteurs QWT.

Performances du descripteur QWT complet

Les résultats globaux en termes de moyenne et variation du taux de reconnaissance sur les 100 essais sont donnés à la table 2.1. On y retrouve les résultats déjà présentés, à savoir que les descripteurs d'amplitude DWT/QWT et le descripteur de phase QWT seul donnent des taux de reconnaissance semblables. On remarque également que la base TC12 semble plus facile à classer que les deux autres base, ce qui s'explique facilement par le fait qu'elle a, par rapport aux autres, peu de classes différentes (24 contre 111 et 68), qui sont très représentées (380 échantillons par classe contre 60 et 25); et que nous avons choisi une base d'apprentissage plus grande (30 échantillons contre 10 et 10).

Le descripteur QWT complet, qui contient les 9 mesures sur l'amplitudes et les 9 mesures sur ψ , obtient des performances de classification très satisfaisantes, avec un gain moyen d'environ 10% sur le taux de reconnaissance par rapport à la DWT. De plus, la variation de cette performance en changeant de base d'apprentissage est plus faible que celle des autres descripteurs, caractérisant ainsi une certaine robustesse.

D'une manière générale, à part une douzaine de classes « problématiques », la QWT est toujours supérieure à la DWT. Quelques performances par classe pour les 4 descripteurs DWT/QWT sont illustrées à la figure 2.13, et la comparaison finale des deux descripteurs en compétition pour de plus nombreuses classes est donnée figure 2.14.

Voyons maintenant quelques cas particuliers.

Notons que la classe Brod. 52, obtient le gain de performance le plus fort entre les deux transformées (100% contre 40% environ). Cette texture ne contient que des structures horizontales, verticales, et diagonales, ces dernières étant plus importantes en termes de luminosité. Il n'est donc pas surprenant que la phase seule donne *exactement 100%* de reconnaissance.

Dans la plupart des cas, si le descripteur d'amplitude QWT n'est pas très efficace par rapport à la DWT, c'est le descripteur de phase qui sera bon. La complémentarité évoquée précédemment

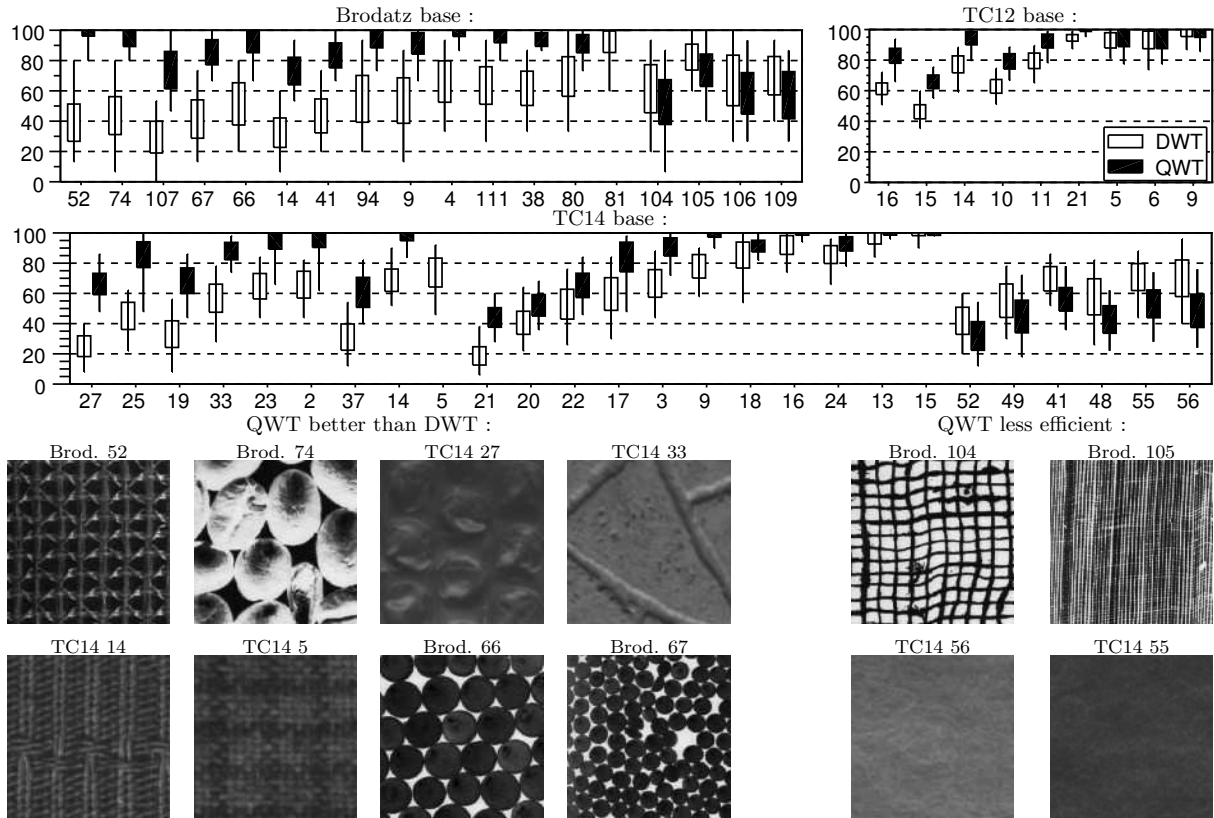


FIGURE 2.14 – Performances du descripteur final QWT basé sur la combinaison amplitude/phase, en comparaison avec le descripteur DWT.

entre les deux informations QWT est largement confirmée ici, et de nombreuses classes mieux reconnues par la QWT le doivent soit à l'amplitude, soit à la phase (Figure 2.13, TC14 33, TC14 9, Brod. 4, Brod. 57, Brod. 12).

Certains cas plus marquants (Figure 2.13, TC12 16, TC12 20) montrent des performances moyennes pour l'amplitude et la phase séparément, et de très bonnes performances avec la combinaison des deux.

Les classes TC14 55 et TC14 56, pour lesquelles la QWT est moins bonne (Figure 2.14), sont de type bruit blanc, purement aléatoires, sans contenu géométrique. On peut remarquer que la plupart des classes TC14 mieux reconnues par la DWT sont difficiles à distinguer à l'œil, et en conclure que la QWT est moins adaptée pour l'analyse des signaux type bruit blanc, par rapport à la DWT dont l'orthogonalité de la base assure une bonne caractérisation de ce genre de signal.

Finalement, nous avons montré expérimentalement que l'utilisation d'une représentation quaternionique des textures améliore fortement ce que l'on peut obtenir avec une transformée en ondelettes classique, en termes de taux de reconnaissance. Nous avons proposé un descripteur global simple à partir d'écart-types et d'écart-types pondérés dans les sous-bandes QWT, qui permet d'augmenter de 10% en moyenne le taux de reconnaissance, grâce à une information analogue à la DWT munie d'invariance par translation (l'amplitude) et une information géométrique complémentaire (la phase). Cette expérimentation valide l'intérêt de la QWT lorsqu'il s'agit de caractériser des textures contenant une certaine complexité géométrique.

Nous traitons maintenant les résultats donnés par les descripteurs CWT de [21, 22].

Performances des ondelettes complexes

Rappelons que la CWT est une approche totalement différente de la QWT dans la manipulation des coefficients, qui se focalise sur la directionnalité plutôt que la notion de signal analytique

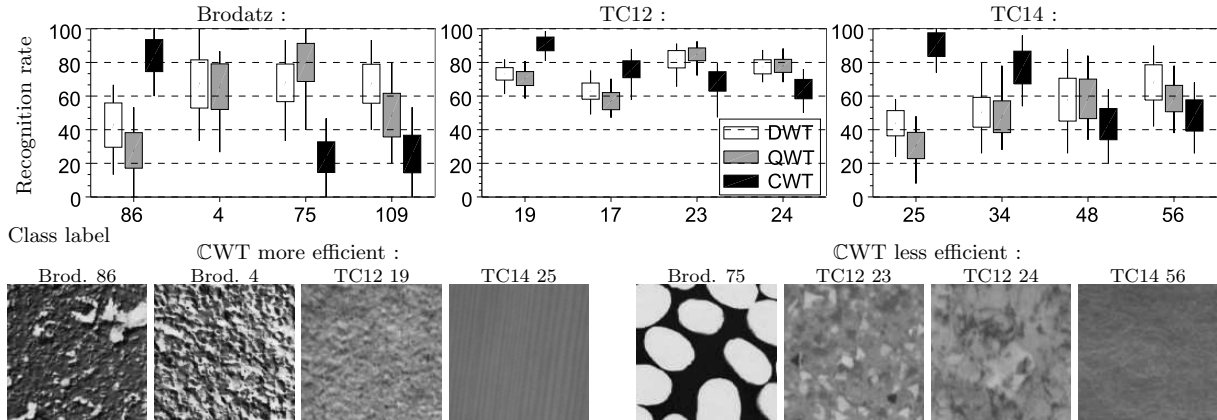


FIGURE 2.15 – Performances de classification DWT/CWT/QWT

2D. L'algorithme de décomposition utilisé est néanmoins presque le même, puisqu'il ne diffère que dans la combinaison des 4 sorties de bancs de filtres dual-tree (voir figure 2.5).

Les résultats liés à la CWT doivent être traités avec précaution. Rappelons que nous avons repris les descripteurs définis dans [21, 22], à l'origine utilisés dans un contexte visant une grande performance de classification. Nous n'avons pas répliqué les autres blocs utilisés dans ces références, à savoir le processus de sélection et normalisation des descripteurs, qui ont pour but d'accroître le taux de reconnaissance, mais qui nous auraient empêché d'identifier la contribution des différents descripteurs. Nous nous plaçons volontairement dans une situation difficile afin de comparer les différentes analyses. C'est pourquoi les résultats que nous obtenons sont moins bons que ceux présentés dans [21, 22].

Globalement, le premier descripteur (amplitude) donne de meilleurs résultats que la DWT mais pas aussi bons que la QWT (voir table 2.1). Cela signifie que la multiplicité des orientations disponibles dans cette transformée est certes un atout, mais qui ne surpasse pas le couple d'informations amplitude/phase de la QWT. Nous présentons figure 2.15 une comparaison par classe des performances des descripteurs d'amplitude DWT, amplitude CWT et amplitude/phase QWT. Les classes illustrées sont celles qui présentent les plus grandes différences entre les descripteurs. La classe TC14 25 illustre bien l'avantage des 6 orientations de la CWT sur les autres transformées, car les structures de ces textures sont alignées environ à 75° , ce qui correspond à la fréquence centrale d'une des sous-bandes.

Le second descripteur CWT a donné de très mauvais résultats (autour de 5% de reconnaissance), ce qui peut s'expliquer par une certaine instabilité des mesures de phase dues à la non prise en compte de la circularité de cette donnée dans le calcul des différences, et probablement au fait que les phases des petits coefficients sont considérées avec la même importance (pas de pondération par l'amplitude comme nous avons fait avec la QWT). Il est possible que l'étage de sélection statistique des mesures utilisé dans [22] élimine les mesures de phase, ce qui expliquerait que les auteurs obtiennent un bon résultat, qui serait alors dû aux seules mesures sur l'amplitude.

Finalement, notre descripteur QWT est également supérieur aux descripteurs CWT proposées dans [21, 22].

2.6 CONCLUSION SUR LES APPROCHES QUATERIONIQUES

Suite à la limite évidente de l'algèbre des nombres complexes pour une théorie du signal 2D, caractérisée par le manque d'unité entre les coefficients des représentations type Gabor, l'algèbre des quaternions a permis de définir une théorie du signal 2D complète. A partir d'une nouvelle transformée de Fourier dont les atomes sont des combinaisons d'ondes planes - par opposition aux simples ondes planes, une nouvelle notion de phase est apparue, et comprend trois angles : φ , θ et ψ . Une définition de signal analytique 2D est déterminée par l'annulation des fréquences

négatives dans les trois quadrants du spectre, offrant un modèle temporel sous forme d'atome de Fourier modulé en amplitude et en phase - selon les trois angles. Les deux premiers angles sont analogues à la phase complexe et encodent des déplacements locaux, mais cette fois-ci de manière plus complète puisque tout déplacement 2D est entièrement caractérisé par (φ, θ) . Le troisième angle pourrait être associé à une mesure de « directionnalité » des structures, et se révèle intéressant pour la caractérisation des textures. Cette nouvelle notion de phase permet de former des atomes 2D plus élaborés que les ondes planes, incluant des structures locales isotropes, des coins *etc.*

La QWT réalise une décomposition en ondelettes quaternionique à reconstruction parfaite de redondance $4\times$ par un algorithme rapide, grâce à un lien algorithmique avec le dual-tree 2D. Cette nouvelle transformée transpose la théorie quaternionique dans le monde des ondelettes, et possède une amplitude invariante par translation. Elle constitue une amélioration des ondelettes séparables car elle est munie d'une phase spécialement définie pour les signaux 2D.

La littérature sur la QWT est très faible, et une partie de notre contribution a été de reprendre et de reformuler la définition et la construction de cette nouvelle approche, tout en apportant des éléments nouveaux sur la compréhension des coefficients.

Nous avons conduit deux expérimentations sur la QWT pour la comparer aux ondelettes classiques dans les contextes de codage d'image et de classification de textures. La QWT s'est montrée avantageuse en codage, car elle présente une bonne parcimonie, et permet donc de reconstruire une approximation visuellement plus satisfaisante que la DWT pour une quantité de données égale. Il vient que la phase peut être quantifiée sur un nombre de bits très faible - analogue au codage du signe des coefficients DWT. Ce résultat se place dans un contexte de codage de « moyenne qualité », et recoupe ce qu'on peut obtenir en général avec d'autres représentations redondantes [50]. Il était aussi attendu dans cette perspective une moins bonne restitution des textures, pour lesquelles la DWT est supérieure. La QWT respecte mieux les formes géométriques, les contours, en générant beaucoup moins d'artefacts autour de ces derniers.

La classification de textures a également montré que la QWT était avantageuse par rapport à la DWT, en apportant un gain moyen approximatif de 10% au taux de reconnaissance global. Notons que ce gain est très significatif pour les textures constituées de motifs *géométriques*. En ce qui concerne les textures « type bruit » en revanche, la QWT n'améliore pas les performances, ce qui rejoint la remarque que nous avons faite à propos du *codage*, dans le cadre duquel ce type de texture est moins bien restitué avec la QWT qu'avec la DWT. Cela est dû à l'orthogonalité de cette dernière qui assure une bonne caractérisation des signaux proches d'un bruit blanc. Les mesures sur l'amplitude QWT donnent des performances équivalentes à la DWT et aux ondelettes complexes, mais l'avantage de la phase QWT est clair car les mesures de phase viennent compléter l'information avec de nouvelles caractéristiques de « quantité de structures diagonales », qui permettent de surpasser la DWT et les ondelettes complexes.

Finalement, nous avons constaté expérimentalement que ce nouveau formalisme pour les signaux 2D permettait d'améliorer la représentation géométrique des images, grâce à une amplitude invariante par translation et une phase portant une information structurelle. Cependant, nous avons aussi pointé que la construction séparable de cette transformée rend les représentations trop dépendantes de l'orientation des structures, ce qui se traduit par une variance par rotation. Il en résulte également que la QWT présente une classification arbitraire des détails selon les directions horizontale, verticale et diagonale, nous obligeant à une échelle fixée à étudier trois coefficients quaternioniques pour une seule position spatiale. Ceci va à l'encontre de l'idée d'unification de l'information géométrique des structures locales. Enfin, la définition de signal analytique 2D par l'approche quaternionique de T. Bülöw n'est pas entièrement satisfaisante car très liée aux lignes et aux colonnes de l'image.

Une solution existe peut-être dans le formalisme « monogène » proposé par M. Felsberg [48], qui prend la suite du travail de T. Bülöw mais cette fois dans une perspective *isotrope*. Et puisqu'une transformée en ondelettes monogène [144] associée à ce modèle a été proposée en 2008, voici l'objet du chapitre suivant.

APPROCHES ISOTROPES ET EXTENSION COULEUR

DANS les chapitres précédents, nous avons évoqué deux méthodes pour appréhender les signaux 2D avec les notions d'amplitude et phase.

La méthode *directionnelle* (Gabor, Dual-Tree), applique K analyses 1D dans des directions différentes. Ces représentations manquent d'unification des informations, puisque chaque position est décrite par K amplitudes et K phases, à chaque échelle. Si K est grand, la redondance est grande et l'information est difficile à synthétiser. Si K est petit, les analyses 1D sous-jacentes se font souvent dans une direction qui ne coïncide pas avec l'orientation locale du signal, ce qui introduit un biais et rend l'analyse variante par rotation du signal.

La méthode *séparable* (ondelettes quaternioniques), offre une information plus unifiée, grâce à une nouvelle notion de phase 2D, mais souffre également de variance par rotation, liée à un schéma « ligne/colonne ». Les structures sont classifiées dans seulement 4 orientations possibles.

Il est pourtant bien connu dans le domaine de la vision par ordinateur, que l'*orientation des contours* est une information primordiale, et implicitement nécessaire à une analyse *isotrope*, c'est-à-dire *invariante par rotation*. Une méthode isotrope classique est le *tenseur de structure*, dont une évolution a été proposée à travers les *filtres orientables* [54], qui généralisent la notion d'orientation locale, pour aboutir à un calcul optimal de phase « à la Gabor ». Nous discuterons ces analyses.

Nous allons présenter dans ce chapitre le *signal monogène*, qui offre un formalisme plus efficace pour analyser conjointement la phase et l'orientation, à travers une nouvelle définition de signal analytique 2D. A l'instar du signal quaternionique analytique du chapitre 3, il s'inclut dans un nouveau formalisme général de traitement du signal, et ouvre la voie vers de nouvelles représentations par ondelettes.

La définition de transformées en *ondelettes monogènes* est encore aujourd'hui un problème ouvert, pour lequel une première solution a été donnée en 2008 par M. Unser *et al.* [144]. La méthode s'appuie sur plusieurs contributions précédentes des auteurs, notamment sur l'introduction de redondance légère dans des bancs de filtres 2D pour améliorer l'invariance et l'isotropie. La transformée en ondelettes monogène fournit des coefficients vectoriels invariants par translation contenant une amplitude et une phase invariants par rotation, et une donnée d'orientation locale. Le banc de filtres à reconstruction parfaite proposé dans [144] pour réaliser cette analyse est peu redondant, et donc se place tout-à-fait dans la catégorie des représentations étudiées dans notre travail. Nous proposons dans ce chapitre une analyse et synthèse concise de ces travaux.

Rappelons que nous nous intéressons aux *images couleur*. Alors que la plupart des applications couleur utilisent des transformées en ondelettes de façon sous-optimale, faute d'outil mathématique, nous pensons que le formalisme monogène est une bonne piste pour définir une représentation 2D étendue aux signaux à valeurs vectorielles. D'une part, la réflexion sur les signaux 2D a ouvert la voie vers des outils multidimensionnels en général, aboutissant à des représentations à *coefficients vectoriels* dans des algèbres plus appropriées que $\mathbb{C}$. Ces travaux forment un environnement propice, selon nous, à l'augmentation de la dimension des *valeurs des signaux*. D'autre part, les représentations monogènes nous semblent constituer une évolution fondamentale prometteuse de l'analyse de Fourier pour les images, à travers des représentations invariantes par translation *et* par rotation, donnant accès à une notion claire de phase 2D. En outre, nous avons déjà évoqué au chapitre précédent que le formalisme quaternionique s'est montré trop limité pour l'imagerie couleur.

Notre but est de définir une représentation des images couleur dont les coefficients permettent à la fois d'analyser les caractéristiques bas-niveau de l'image, et de l'écrire comme une *somme d'atomes*, reproduisant ces caractéristiques locales à travers une procédure de *reconstruction* cohérente. La littérature nous offre un seul article proposant une extension couleur du signal monogène [38], c'est pourquoi notre premier travail, présenté dans ce chapitre, a été de prendre cette méthode comme point de départ, et de mener une étude de faisabilité. Nous avons donc défini une transformée en ondelettes *monogène couleur*, conforme au modèle de [38]. Cette transformée est la première représentation par ondelettes spécialement définie pour les images couleur de façon non-marginale, et dont l'analyse est issue d'un formalisme innovant de représentation des signaux 2D.

Pour mettre à l'épreuve ces nouvelles méthodes, nous avons ensuite conduit une application de cette transformée en débruitage d'images couleur, en passant par une modélisation statistique des coefficients. Ce premier travail expérimental nous permettra de discuter de cette première extension des ondelettes monogènes à la couleur.

Le chapitre est organisé de la façon suivante. Nous rappelons d'abord les outils classiques d'analyse « bas-niveau » des images, qui anticipent l'interprétation physique du formalisme monogène avec les notions d'orientation et de phase locales, et qui ne sont pas toujours mis en avant dans les manuels de traitement d'image. Le formalisme du signal monogène est ensuite présenté, ainsi que la définition de transformées en ondelettes monogènes. Notre première proposition d'ondelettes monogènes couleur est ensuite décrite, pour finir avec son application en débruitage couleur.

3.1 ANALYSE LOCALE DES STRUCTURES

L'étude d'une généralisation *invariante par rotation* du signal analytique pour l'image - qui sera notre objet à la section suivante - nécessite d'analyser les concepts d'*orientation* des structures locales d'une part, ainsi que de *phase* d'un signal 2D. Or ces notions sont étudiées depuis une trentaine d'années dans le contexte de la *vision bas-niveau* par ordinateur. Ainsi, le *tenseur de structure* a d'abord offert une analyse de l'orientation locale bien établie. Ensuite, les méthodes par *filtres en quadrature* ont abouti à une analyse de Gabor optimisée, réduite en chaque point à une orientation précise et unique. Ces dernières permettent enfin d'extraire une notion de phase, qui anticipe l'interprétation du *signal monogène* que nous présenterons à la prochaine section.

3.1.1 Tenseur de structure basé gradient

Le formalisme du tenseur de structure dédié à la description géométrique locale a été initié vers 1986 [53, 9], et est maintenant un outil classique [61, 73]. Il consiste en l'extraction en chaque point de l'image et pour une échelle donnée, de *trois* mesures, extraites d'une matrice 2×2 symétrique, le tenseur. Ces mesures sont destinées à décrire les structures locales de l'image en termes de quantité de variation, de directionnalité, et d'orientation principale. Il

existe plusieurs méthodes de calcul d'un tenseur de structure, la plus connue étant celle basée sur l'estimation du *gradient* de l'image.

Gradient

Le gradient d'une fonction $s(\mathbf{x})$ de dimension 2 est défini par :

$$\nabla s(\mathbf{x}) = \begin{bmatrix} \frac{\partial s(\mathbf{x})}{\partial x_1} & \frac{\partial s(\mathbf{x})}{\partial x_2} \end{bmatrix}^T = \mathcal{N}_\nabla(\mathbf{x}) \begin{bmatrix} \cos(\theta_\nabla(\mathbf{x})) \\ \sin(\theta_\nabla(\mathbf{x})) \end{bmatrix} \quad (3.1)$$

où $\mathcal{N}_\nabla = \|\nabla s\|$ est appelé *norme du gradient* et $\theta_\nabla = \arg\left\{\frac{\partial s(\mathbf{x})}{\partial x_1} + \mathbf{j}\frac{\partial s(\mathbf{x})}{\partial x_2}\right\}$ *direction du gradient*. Cet opérateur différentiel est par définition lié aux *variations locales* du signal étudié. Associé aux dérivées partielles d'ordre 1, il correspond à une étude *linéaire* de ces variations, qui permet de les approximer en tout point par un *plan* tangent (cf. développement de Taylor). Si le signal est une image, alors les « plans » de pente forte correspondent à des *contours rectilignes* de type « bord ». La détection de bords par gradient a toujours été très utilisée en traitement d'images, mais nécessite en pratique une approximation numérique, dont il existe plusieurs définitions (Prewitt, Sobel, Canny [18], Dérivée, Laplaciens de Gaussiennes [104] etc.).

Afin de mieux comprendre le lien entre les contours et le gradient, plaçons nous dans le cas des signaux « i1D » (« intrinsèquement 1D », également appelés « signaux simples »). Un signal i1D $s(\mathbf{x})$ est égal à une fonction 1D *orientée* :

$$s(\mathbf{x}) = f_{1D}(x_1 \cos \alpha + x_2 \sin \alpha) \quad (\text{signal « i1D »}) \quad (3.2)$$

Il s'agit par exemple d'une onde plane, et les fortes variations 1D de f_{1D} se traduisent dans s par des bords rectilignes, orientés vers $\alpha + \pi$. Dans ce cas l'expression du gradient de s se simplifie comme ceci :

$$\nabla s(\mathbf{x}) = f'_{1D}(x_1 \cos \alpha + x_2 \sin \alpha) [\cos \alpha \quad \sin \alpha]^T \quad \text{avec} \quad f'_{1D}(t) = \frac{\partial f_{1D}(t)}{\partial t} \quad (3.3)$$

ce qui implique :

$$\mathcal{N}_\nabla(\mathbf{x}) = |f'_{1D}(x_1 \cos \alpha + x_2 \sin \alpha)| \quad \text{et} \quad \theta_\nabla = \alpha \quad (3.4)$$

En supposant que l'image étudiée soit localement i1D, alors son gradient permettra d'extraire *en tout point* et de façon exacte, l'intensité et l'orientation d'un éventuel « bord ».

Maximisation de la variation locale

Pour aborder le formalisme du tenseur de structure, il faut penser le gradient comme la solution d'un problème de maximisation. L'idée est de considérer une mesure *orientée* de variation locale, définie par rotation de la dérivée partielle $\frac{\partial}{\partial x_1}$:

$$D_\theta s(\mathbf{x}) = \frac{\partial}{\partial x_1} \left[s \left(\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) \right] \quad (3.5)$$

et de s'intéresser à la direction θ qui *maximise* cette mesure. Si le signal présente un bord, alors sa variation maximum doit être dans la direction perpendiculaire à ce bord. On montre facilement que $D_\theta = \cos \theta \frac{\partial}{\partial x_1} + \sin \theta \frac{\partial}{\partial x_2}$ et donc par identité trigonométrique :

$$D_\theta s(\mathbf{x}) = \mathcal{N}_\nabla(\mathbf{x}) \cos(\theta_\nabla(\mathbf{x}) - \theta) \quad (3.6)$$

Notre mesure de variation atteint un *maximum global* $\mathcal{N}_\nabla$ dans la direction du gradient $\theta = \theta_\nabla$ et un minimum $-\mathcal{N}_\nabla$ dans la direction opposée $\theta = \theta_\nabla + \pi$. Elle est nulle dans les directions perpendiculaires $\theta = \theta_\nabla \pm \pi/2$. Le gradient peut donc être vu comme la solution de la maximisation de la mesure de variation orientée D_θ . Notons que la restriction de cette analyse à un modèle

d'ordre 1 donnera toujours, et quelle que soit la structure locale réelle du signal, une variation angulaire *cosinusoidale* (eq. (3.6))¹. L'idée du tenseur de structure est de modifier cette mesure de variation afin de rendre l'analyse plus efficace.

Construction du tenseur

D'une part, il est souhaitable que les directions opposées θ et $\theta + \pi$ soient fusionnées dans une même information d'*orientation* car elles décrivent la même variation. Pour cela, on peut par exemple considérer $(D_\theta s)^2$, qui atteindra son maximum $\mathcal{N}_\nabla^2$ dans les deux directions opposées, éliminant au passage la notion ambiguë de variation négative puisque $(D_\theta s)^2 \geq 0$.

D'autre part, on introduit un calcul de *moyenne* sur la mesure de variation dans un *voisinage* du point étudié, afin de détecter les structures qui ne correspondent pas au modèle d'ordre 1 (bord rectiligne). Si la structure est bien un bord, les variations dans le voisinage seront semblables, et le moyennage ne changera pas du gradient classique. Dans le cas contraire, la variation angulaire moyenne déviara du modèle cosinusoidal, et ne présentera pas nécessairement un maximum univoque. On considère donc la mesure $(h * (D_\theta s)^2)(\mathbf{x})$, où $h(\mathbf{x})$ est une fenêtre.

Cette nouvelle mesure orientée est une forme quadratique associée à une matrice définie-positive T appelée *tenseur de structure* :

$$(h * (D_\theta s)^2)(\mathbf{x}) = \begin{bmatrix} \cos \theta & \sin \theta \end{bmatrix} \underbrace{\begin{bmatrix} h * (\frac{\partial s}{\partial x_1})^2 & h * (\frac{\partial s}{\partial x_1} \frac{\partial s}{\partial x_2}) \\ h * (\frac{\partial s}{\partial x_1} \frac{\partial s}{\partial x_2}) & h * (\frac{\partial s}{\partial x_2})^2 \end{bmatrix}}_T \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (3.7)$$

La maximisation de cette forme en fonction de θ est un problème classique qui consiste à trouver les valeurs et vecteurs propres de T . Les valeurs propres λ_+ et λ_- sont positives et définies par :

$$\lambda_{\pm} = \frac{1}{2} \left(T_{11} + T_{22} \pm \sqrt{(T_{22} - T_{11})^2 + 4T_{12}^2} \right) \quad (3.8)$$

La plus grande, λ_+ , correspond à la valeur maximum de $(h * (D_\theta s)^2)$, et est atteinte dans la direction θ_+ du vecteur propre associé² :

$$\theta_+ = \frac{1}{2} \arg\{T_{11} - T_{22} + j2T_{12}\} \quad (3.9)$$

Le second vecteur propre est orthogonal au premier ($\theta_- = \theta_+ \pm \pi/2$). La maximisation de la variation donne donc $h * (D_{\theta_+} s)^2 = \lambda_+$, et fournit également une troisième information λ_- , qui mesure la variation de la fonction dans la direction perpendiculaire à la variation principale - soit la variation *minimum*, qui est positive, et cette fois pas forcément nulle.

Dans le cas dégénéré où on ne considérerait pas de voisinage, $h(\mathbf{x}) = \delta(\mathbf{x})$, on a $T_{12}^2 = T_{11}T_{22}$ et donc les équations (3.8) et (3.9) sont réduites aux valeurs du gradient classique : $\lambda_+ = \mathcal{N}_\nabla^2$, $\lambda_- = 0$, et $\theta_+ = \theta_\nabla(\pi)$. La variation minimum λ_- est dans ce cas toujours nulle, car on ne capte plus les structures ne correspondant pas au modèle. Dans le cas général, on a toujours :

$$h * (\mathcal{N}_\nabla^2) = \lambda_+ + \lambda_- \quad (3.10)$$

Finalement, le tenseur de structure apporte au gradient une mesure supplémentaire, souvent utilisée à travers la notion de « cohérence » définie par :

$$\chi = (\lambda_+ - \lambda_-) / (\lambda_+ + \lambda_-) \in [0; 1] \quad (\text{cohérence}) \quad (3.11)$$

qui est proche de 1 si la structure locale correspond effectivement à un bord rectiligne, et proche de 0 sinon. Dans le second cas, la structure locale peut être un coin ou une jonction, pour lesquels la donnée d'orientation risque d'être moins pertinente.

1. Cette fonction de θ est égale à la partie angulaire de l'équation du plan tangent au signal, exprimée en coordonnées polaires.

2. Selon les références, on trouvera parfois ' $T_{11} - T_{22}$ ' [87, 47, 43, 123] et parfois ' $T_{22} - T_{11}$ ' [144, 73].

D'autres méthodes utilisent aussi les dérivées d'ordre supérieur [83], ce qui permet de caractériser de nouvelles structures comme les jonctions. Une méthode alternative de cette analyse des ruptures et de leur orientation consiste à utiliser des *filtres en quadrature*. Nous proposons de décrire cette approche dans la partie suivante.

3.1.2 Approche « signal » du tenseur

Nous étudions ici l'approche « signal » vers une détection optimale des contours, qui a été proposée en 1982 dans la thèse de H. Knutsson [82], et qui a donné naissance à la notion d'invariance par changement de phase. Cette propriété permet de détecter également les *bords* et les *lignes* dans une procédure unifiée, sans avoir recours à des dérivées d'ordre supérieur. La méthode a ensuite été étendue dans un formalisme de tenseur [61].

Il existe depuis la fin des années 70 une définition assez répandue des filtres « en quadrature 2D » : il s'agit des filtres dont la réponse fréquentielle est nulle dans un demi-plan. La partie imaginaire de ces filtres est donc la transformée de Hilbert partielle de la partie réelle (voir section 1.3.1) dans une certaine direction. Les décompositions type Gabor 2D classiques [60, 34] que nous avons vues à la section 1.3.2 sont implicitement construites sur cette définition. L'avantage de ces filtres est que l'amplitude de leurs sorties est invariante par changement de phase du signal, c'est-à-dire que le modèle pourra s'adapter aux *bords* aussi bien qu'aux *lignes* (voir la correspondance entre la phase et les structures 1D à la table 1.1).

Afin d'étudier l'orientation locale de l'image, on s'intéresse à des filtres g dont la réponse fréquentielle est séparable en coordonnées polaires, et dont on considère les rotations g_{θ_0} d'angle quelconque θ_0 :

$$g_{\theta_0}(\mathbf{x}) \xleftrightarrow{\mathcal{F}} r(\|\boldsymbol{\omega}\|) d(\arg\{\omega_1 + j\omega_2\} - \theta_0) \quad (3.12)$$

où r est la réponse dite *radiale*, qui déterminera le gabarit fréquentiel du filtre, et d la réponse *angulaire* (fonction périodique), qui indiquera la sélectivité du filtre en termes d'orientation. L'idée est de remarquer que le filtrage par g_{θ_0} d'un signal simple $s(\mathbf{x}) = f_{1D}(x \cos \alpha + y \sin \alpha)$ fournit une sortie qui ne dépend de θ_0 que par un facteur constant. Plus précisément, on a :

$$(s * g_{\theta_0})(\mathbf{x}) = (s * g_{\theta_1})(\mathbf{x}) \frac{d(\alpha - \theta_0)}{d(\alpha - \theta_1)} \quad (3.13)$$

Ceci est dû au fait que le support du spectre d'un signal simple est réduit à une ligne orientée vers α . En se fixant une réponse angulaire d et plusieurs directions de filtrage, on peut donc retrouver exactement α . L'analyse est toujours valable si le signal est *localement* simple en tout point, et donnera une orientation variable $\alpha(\mathbf{x})$. La partie radiale r contrôle indépendamment l'échelle d'analyse.

Prenons maintenant le cas particulier :

$$d(\theta) = \begin{cases} \cos^2(\theta) & \text{si } |\theta| < \pi/2 \\ 0 & \text{sinon} \end{cases} \quad (3.14)$$

Le filtre correspondant est orienté, et *en quadrature*. Il est montré dans [82] que pour couvrir toutes les orientations possibles du signal dans ce cas particulier, au moins trois filtrages doivent être utilisés. En prenant *quatre* filtres orientés vers $\theta_0 \in \{0, \pi/4, \pi/2, 3\pi/4\}$ (choix qui simplifie les expressions et l'implantation), on peut considérer l'*amplitude* des signaux de sortie :

$$s_1 = |(s * g_0)(\mathbf{x})| \quad s_2 = |(s * g_{\pi/4})(\mathbf{x})| \quad s_3 = |(s * g_{\pi/2})(\mathbf{x})| \quad s_4 = |(s * g_{3\pi/4})(\mathbf{x})| \quad (3.15)$$

et vérifier par identités trigonométriques que :

$$\alpha = \frac{1}{2} \arg\{(s_1 - s_3) + j(s_2 - s_4)\} \quad (3.16)$$

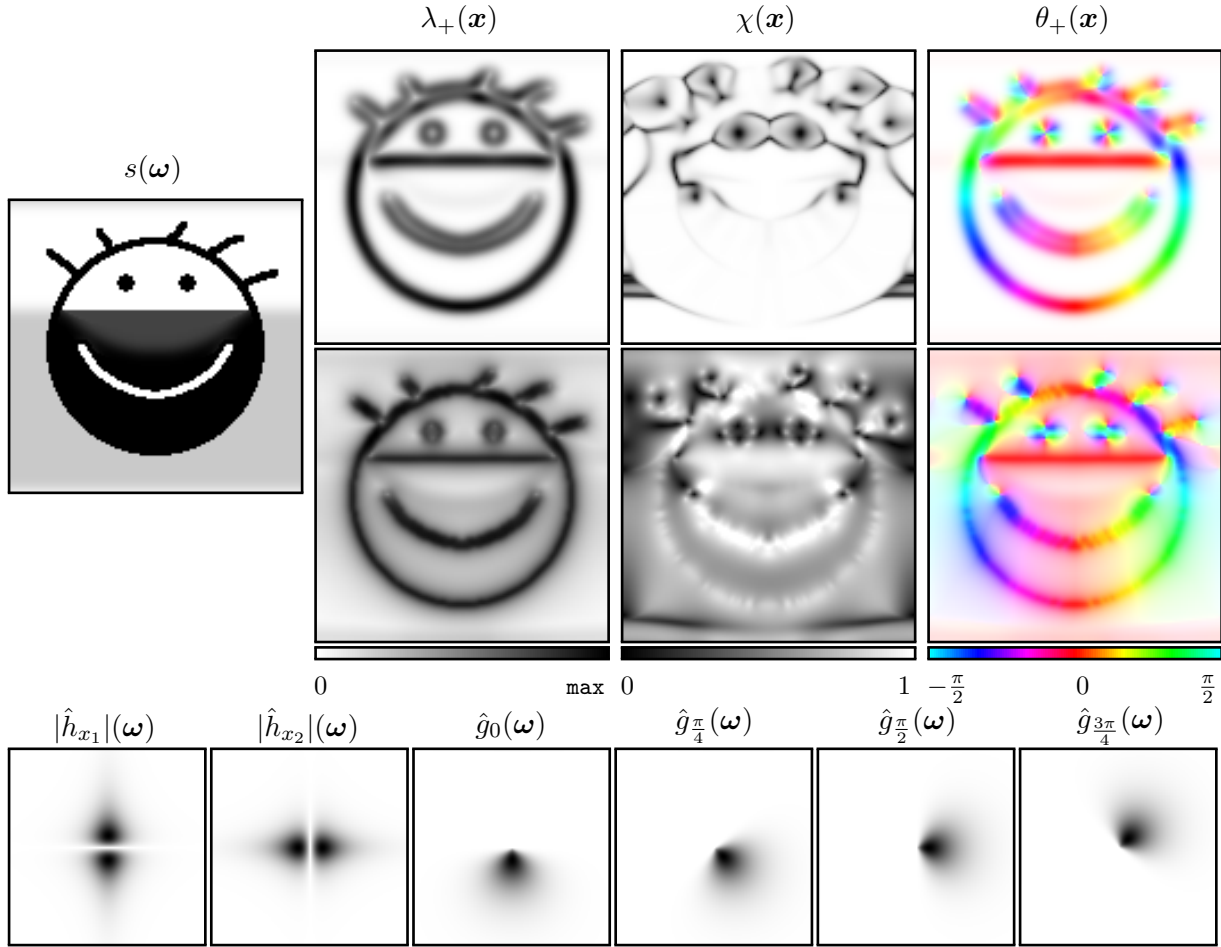


FIGURE 3.1 – Analyse par tenseur de structure de l'image s . Méthode par gradient (1ère ligne), méthode par filtres en quadrature séparables en coordonnées polaires (2ème ligne), et gains fréquentiels des filtres utilisés (3ème ligne).

Cette technique a été étendue à la construction complète de tenseur de structure dans [61]. Toujours dans le cas de nos quatre filtres, on peut construire le tenseur suivant :

$$T = \frac{1}{4} \begin{bmatrix} 3s_1 + s_2 - s_3 + s_4 & 2s_2 - 2s_4 \\ 2s_2 - 2s_4 & -s_1 + s_2 + 3s_3 + s_4 \end{bmatrix} \quad (3.17)$$

et montrer que ses valeurs/vecteurs propres portent le même type d'information que celles du tenseur basé gradient. Nous proposons d'analyser les différences entre les deux approches par un exemple.

Calcul d'un exemple

Afin de discuter les propriétés du tenseur de structure en comparant les deux méthodes présentées, nous avons illustré un exemple sur la figure 3.1.

Pour la méthode par estimation de gradient (première ligne de la figure 3.1), nous avons utilisé les filtres de R. Dériché [42], qui font référence dans ce domaine car ils optimisent le critère de J. Canny sur la détection des bords. Nous avons fixé le paramètre d'échelle « α » [42] de façon à ce que les filtres aient une réponse maximum à la pulsation $8 \times \frac{2\pi}{N}$ (où N est le nombre de lignes/colonnes de l'image)³. Nous désignons ces filtres par h_{x_1} et h_{x_2} , dont le module des spectres $|\hat{h}_{x_1}(\omega)|$ et $|\hat{h}_{x_2}(\omega)|$ est illustré sur la troisième ligne de la figure 3.1. La fenêtre de lissage du tenseur est une gaussienne de taille $\sigma = 1$.

3. Cette valeur de fréquence centrale correspond bien à l'échelle des structures de notre image de test

Pour la méthode par filtres en quadrature séparables en coordonnées polaires, nous avons utilisé une réponse radiale définie par le spectre d'une différence de noyaux de Poisson ajustée en octaves :

$$r(\|\omega\|) = e^{-\frac{N \log(2) \|\omega\|}{2\pi 2^k}} - e^{-\frac{N \log(2) \|\omega\|}{2\pi 2^{k-1}}} \quad (3.18)$$

de façon à ce que les filtres aient une réponse maximum à la pulsation $8 \times \frac{2\pi}{N}$, comme précédemment, en fixant le numéro d'octave $k = 3$. Les modules des spectres des filtres g_0 , $g_{\frac{\pi}{4}}$, $g_{\frac{\pi}{2}}$ et $g_{\frac{3\pi}{4}}$ (voir equation (3.12)) sont illustrés sur la troisième ligne de la figure 3.1.

Pour chaque méthode, nous illustrons la plus grande valeur propre λ_+ , qui indique la force de variation, la mesure de cohérence (équation (3.11)), qui discrimine les structures « i2D » (plus complexes que le modèle linéaire), et la direction θ_+ du vecteurs propre associé à λ_+ , qui indique l'orientation principale. L'affichage de θ_+ se fait par l'encodage de cette valeur dans la *teinte* (« hue ») de l'illustration générée par le système de couleur HSV. L'intensité est fixée à 1 et la saturation est encodée par λ_+ . Ainsi, la valeurs de θ_+ non significatives, c'est-à-dire pour lesquelles λ_+ est faible, sont « blanches » et donc moins visibles. Cela permet de n'observer que les données pertinentes.

Observations comparatives

L'observation de la figure 3.1 montre que la détection de contours effectuée par λ_+ varie significativement entre les deux méthodes, ainsi que la mesure de cohérence.

Une première raison est celle de la différence des gabarits fréquentiels des filtres utilisés. Bien que dans notre cas, nous ayons pris soin de minimiser cette différence en « synchronisant » nos filtres, des différences persistent nécessairement entre la réponse radiale r des filtres en quadrature et le gain des filtres de Dérivée. De plus, l'utilisation d'un lissage du tenseur dans la méthode par gradient « élargit » la réponse λ_+ en atténuant les oscillations autour des bords.

Une deuxième raison est celle du modèle de contour sous-jacent. La première méthode est optimisée pour détecter les variations de type *bord*. En conséquence une *ligne* sera détectée comme deux bords alignés côte-à-côte, et traduite par une « double-réponse », visible ici au niveau de la bouche ou des cheveux du personnage. La seconde méthode est construite sur l'amplitude de la sortie de filtres en quadrature, qui est théoriquement invariante par changement de *phase* du signal (ceci est toutefois limité aux structures *alignées* avec le filtre utilisé). La contrepartie de cette propriété intéressante est qu'un nombre plus important de filtres est nécessaire pour construire le tenseur (au moins 3 filtres complexes, donc 6 filtres réels [61]).

On constate que les lignes que nous venons d'évoquer sont mieux détectées avec la deuxième méthode. Il s'agit de la bouche et des cheveux, qui produisent une réponse *simple*, et correctement centrée sur ces lignes. Les fortes valeurs de *cohérence* χ au niveau de ces lignes confirment que le signal concorde bien au modèle. Notons que cela n'est plus vrai aux extrémités de ces lignes (cheveux et bouche), ainsi qu'aux jonctions (racines des cheveux), et sur les structures isotropes (yeux). Dans ces cas particuliers, la cohérence χ indique des valeurs faibles traduisant une faible concordance au modèle orienté.

Globalement, l'orientation estimée sur l'ensemble des structures est cohérente avec notre perception. Il faut toutefois noter que les cheveux ne présentent pas d'orientation claire avec la méthode gradient, ce qui est dû à un passage par zéro du gradient au centre de la ligne, entraînant une instabilité du calcul de l'orientation locale. Avec la méthode par filtres en quadrature, l'orientation des cheveux est légèrement plus cohérente, en particulier sur la « crête » de la ligne.

Notons que malgré l'utilisation de filtres en quadrature, cette méthode ne permet pas d'extraire explicitement la phase locale de s . En effet, il faudrait utiliser un filtre en quadrature spécifique en chaque point, dont la direction de filtrage serait celle de l'orientation estimée, ce qui n'est pas optimal car la complexité de l'algorithme serait en N^2 . Voyons maintenant comment surpasser ce problème avec une autre méthode de filtrage pour l'analyse de l'orientation.

3.1.3 Filtres en quadrature orientables

Les filtres *orientables* (« steerable ») [54] ont été définis pour améliorer la résolution angulaire de la décomposition de Gabor tout en réduisant le nombre de filtres à utiliser. L'idée est d'utiliser un nombre limité K de rotations d'un filtre de base, dont une simple combinaison linéaire permettra de construire une rotation *quelconque*. Cette orientabilité se répercute sur les K sorties des filtres, qui donneront donc accès *analytiquement* au filtrage du signal dans une direction quelconque. La référence [54] s'attache à répondre aux questions suivantes :

- Comment savoir si un filtre est orientable ?
- Combien de filtres de base seront nécessaires pour l'orienter ?
- Comment combiner les filtres de base pour générer une orientation quelconque ?

Prenons l'exemple très simple de la dérivée partielle. Nous avons déjà vu à propos du gradient que les deux opérateurs $\frac{\partial}{\partial x_1}$ et $\frac{\partial}{\partial x_2}$, forment une base pour connaître la dérivée partielle d'une fonction dans n'importe quelle direction (équations (3.5) et (3.6)). Le gradient constitue donc un cas particulier de filtres orientables, où les filtres de bases sont dirigés vers 0 et $\pi/2$. Dans [54], le filtre permettant d'estimer la dérivée du signal est la dérivée première de gaussienne $g_a(\mathbf{x}) = -2x_1 e^{-\|\mathbf{x}\|^2}$, qui associée à sa rotation de $\pi/2$, $g_b(\mathbf{x}) = -2x_2 e^{-\|\mathbf{x}\|^2}$, forme une base de filtres orientables. Dans ce cas simple, il vient que la sortie du filtrage atteint un maximum dans la direction $\theta_g = \arg\{(s * g_a) + j(s * g_b)\}$, conformément à un modèle de structure sous forme d'un *plan tangent*, comme déjà discuté à propos du gradient.

L'intérêt des filtres orientables est qu'ils se généralisent pour les dérivées d'ordre supérieur. Dans ce cas, les filtres sont plus sélectifs en direction, de plus nombreuses orientations de base sont nécessaires, et l'équation de combinaison est plus complexe. Le modèle sous-jacent de structure locale est alors plus riche, et peut permettre par exemple de mesurer des orientations multiples, comme cela peut être le cas autour de jonctions ou de coins.

De plus, le fait de pouvoir connaître la sortie du filtrage dans toute direction donne une solution à la mesure de la *phase locale*, qui doit se faire par un filtrage en quadrature orienté *localement* dans la direction principale du signal. Pour cela, les auteurs proposent d'utiliser *deux* filtres orientables, l'un étant la transformée de Hilbert partielle de l'autre. Cependant, la contrainte d'orientabilité entre en conflit avec la notion de quadrature. Les auteurs proposent donc d'*approximer* la transformée de Hilbert par minimisation d'erreur quadratique sur un prototype de filtre *polynomial*. Cela signifie que les valeurs de phase extraites devront être considérées avec précaution car inexactes. L'exemple donné dans [54] est basé sur une dérivée de gaussienne d'ordre 2, qui associée à l'approximation de sa transformée de Hilbert nécessitera un total de 7 filtres, pour analyser conjointement l'orientation et la phase locale du signal.

Afin de réduire le nombre de filtres, nous avons choisi de présenter ici un exemple de cette analyse dans le cas de la dérivée de gaussienne d'ordre 1.

Calcul des filtres

Nous allons ici calculer les filtres nécessaires pour réaliser une analyse orientable qui nous permettra d'extraire l'orientation et la phase locales du signal. Les filtres sont définis par leur réponse impulsionnelle continue, et seront discrétisés ensuite.

Comme nous l'avons déjà indiqué, la première base de filtres (g_a, g_b) et sa loi de combinaison pour les orienter est définie par :

$$g_a(\mathbf{x}) = -2x_1 e^{-\|\mathbf{x}\|^2} \quad (3.19)$$

$$g_b(\mathbf{x}) = -2x_2 e^{-\|\mathbf{x}\|^2} \quad (3.20)$$

$$g_\theta(\mathbf{x}) = \mathcal{N}_g \cos(\theta - \theta_g) \quad \text{avec} \quad \mathcal{N}_g \begin{bmatrix} \cos(\theta_g) \\ \sin(\theta_g) \end{bmatrix} = \begin{bmatrix} g_a \\ g_b \end{bmatrix} \quad (3.21)$$

En ce qui concerne la transformée de Hilbert, la restriction à un polynôme d'ordre 2 aboutit à

l'approximation suivante :

$$\mathcal{H}_0^{\text{part}} g_a \approx h_a(\mathbf{x}) = (-1.29x_1^2 + 0.95)e^{-\|\mathbf{x}\|^2} \quad (3.22)$$

qui nécessite 3 orientations de base au minimum⁴. Le choix classique est de prendre les orientations $(0, \pi/3, 2\pi/3)$, ce qui ajoute les deux filtres :

$$h_b(\mathbf{x}) = (-1.29(\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}x_2)^2 + 0.95)e^{-\|\mathbf{x}\|^2} \quad (3.23)$$

$$h_c(\mathbf{x}) = (-1.29(-\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}x_2)^2 + 0.95)e^{-\|\mathbf{x}\|^2} \quad (3.24)$$

La loi de combinaison des trois filtres pour obtenir une orientation quelconque est la suivante [54] :

$$h_\theta(\mathbf{x}) = \frac{h_a + h_b + h_c}{3} + \frac{2\mathcal{N}_h}{3} \cos(2\theta - 2\theta_h) \quad \text{avec} \quad \mathcal{N}_h \begin{bmatrix} \cos(2\theta_h) \\ \sin(2\theta_h) \end{bmatrix} = \begin{bmatrix} h_a - (h_b + h_c)/2 \\ h_b - h_c \end{bmatrix} \quad (3.25)$$

On peut ensuite calculer l'amplitude du filtre en quadrature :

$$|q_\theta|^2 = g_\theta^2 + h_\theta^2 = c + \mathcal{N}_q \cos(2\theta - 2\theta_q) + \frac{2}{3} \mathcal{N}_h^2 \cos(4\theta - 4\theta_h) \quad (3.26)$$

$$\text{avec} \quad \mathcal{N}_q \begin{bmatrix} \cos(2\theta_q) \\ \sin(2\theta_q) \end{bmatrix} = \begin{bmatrix} \frac{g_a^2 - g_b^2}{2} + \frac{4}{3}(h_a + h_b + h_c)(h_a - \frac{h_b + h_c}{2}) \\ g_a g_b + \frac{4}{3}(h_a + h_b + h_c)(h_b - h_c) \end{bmatrix} \quad (3.27)$$

$$\text{et} \quad c = \left(\frac{a + b + c}{3} \right)^2 + \frac{\mathcal{N}_g^2}{2} + 2 \left(\frac{\mathcal{N}_h}{3} \right)^2 \quad (3.28)$$

qui décrit une variation angulaire plus riche que les précédentes. Elle contient une composante constante c qui traduira l'isotropie de la structure locale, une composante en $\cos(2\theta)$ qui décrit l'orientation *principale*, et une composante en $\cos(4\theta)$, que nous négligerons ici, mais qui pourrait être utilisée pour détecter des orientations multiples. L'intérêt pour nous est simplement d'extraire $\theta_q(\mathbf{x})$ de façon à mesurer l'amplitude et la phase de la sortie du filtre en quadrature dans cette direction précise, en tout point. On peut finalement identifier l'amplitude et la phase locales de s en calculant analytiquement $q_{\theta_q} = g_{\theta_q} + jh_{\theta_q}$.

Exemple d'analyse

Au final, la procédure d'analyse par filtres en quadrature orientables « d'ordre 1 » est la suivante :

- Effectuer les 5 filtrages par $(g_a, g_b, h_a, h_b, h_c)$, en échantillonnant leur réponse impulsionnelle par une opération de la forme « $g[k] = g(kT_e)$ » ;
- Calculer l'orientation locale $\theta_q(\mathbf{x})$ à partir de l'équation (3.28) ;
- Utiliser les deux formules de combinaison linéaire de (g_a, g_b) et (h_a, h_b, h_c) pour obtenir analytiquement $g_{\theta_q}(\mathbf{x})$ et $h_{\theta_q}(\mathbf{x})$;
- Extraire l'amplitude $A_q = \sqrt{g_{\theta_q}^2 + h_{\theta_q}^2}$ et la phase $\varphi_q = \arg\{g_{\theta_q} + jh_{\theta_q}\}$.

Un exemple de cette analyse est donné à la figure 3.2. Notons que pour faire correspondre le gabarit fréquentiel à la figure 3.1, nous avons échantillonné les filtres pour $T_e = 0.5$.

On constate que l'amplitude A_q est effectivement invariante par changement de phase, puisque les lignes et les bords produisent une réponse semblable et correctement positionnée. Aussi, l'orientation θ_q est effectivement cohérente avec la perception que nous avons des contours de l'image analysée. Cette méthode fournit donc une extension satisfaisante du tenseur de structure.

4. En pratique, nous avons également retiré la composante continue de ce filtre pour éviter d'introduire un biais dans l'amplitude.

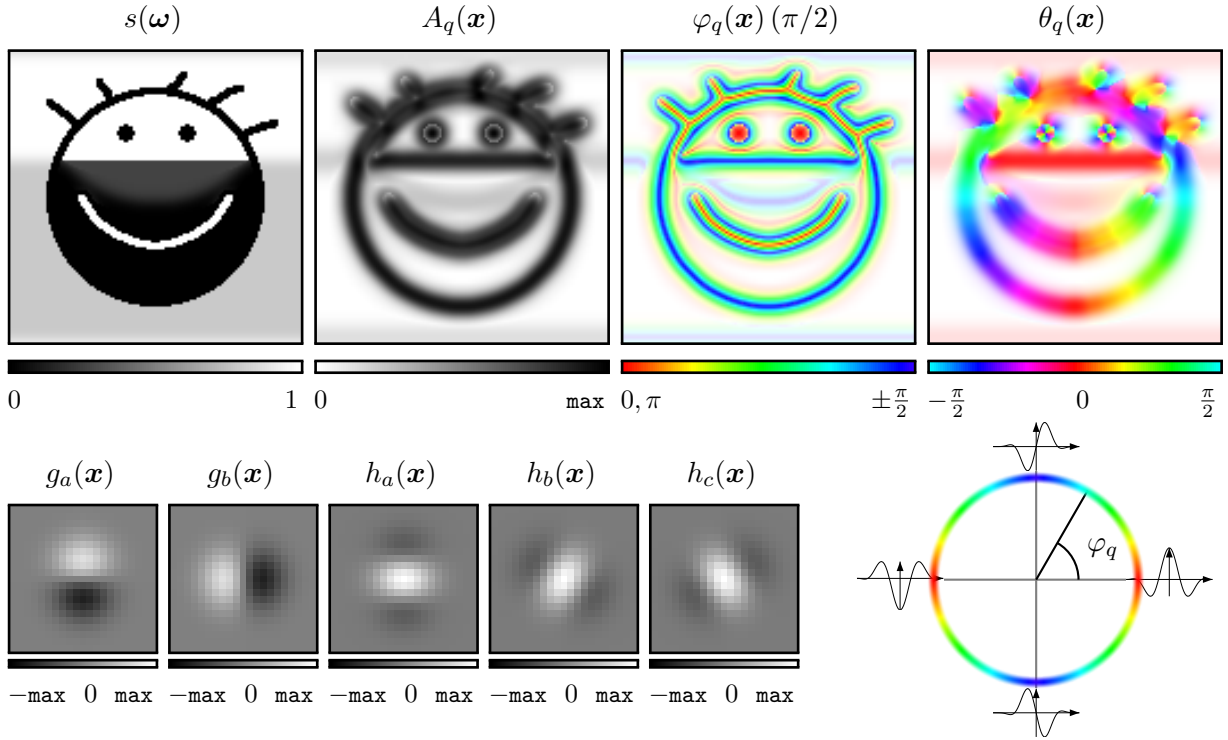


FIGURE 3.2 – Analyse par filtres en quadrature orientables (première ligne), réalisée par 5 filtrages, dont nous illustrons la réponse impulsionnelle (deuxième ligne, gauche). L’affichage en couleurs de la phase φ_q suit la correspondance du schéma en bas à droite.

L’avantage supplémentaire est l’accès à la phase locale. Nous affichons cette phase de façon simplifiée dans $[0; \pi/2]$, selon la correspondance illustrée (fig. 3.2). Cela permet de représenter par la même couleur les lignes foncées et claires, dont la phase ne diffère que par un offset de π , ainsi que les bords de directions opposées, dont la phase ne diffère que par le signe. Cette restriction n’empêche pas de capter toute la continuité du modèle de contour, entre le bord et la ligne. On constate que φ_q indique bien des valeurs proches de 0 ou π autour des lignes (cheveux, bouche) et des valeurs proches de $\pm\pi/2$ autour des bords (tour du visage). La phase est parfaitement stable par changement d’orientation, grâce au formalisme utilisé qui considère une estimation précise de l’orientation principale.

Nous pourrions également exploiter les 2 degrés de liberté restants (car nous n’avons que 3 informations pour 5 filtrages) pour mesurer l’isotropie et détecter les orientations multiples, ceci est exposé dans [54].

Limites

Voyons maintenant les limites de cet outil. D’abord, il est théoriquement impossible de définir un filtre orientable qui soit l’exacte transformée de Hilbert partielle d’un autre. Cela s’explique par le fait qu’un filtre est orientable par K versions de lui-même si et seulement si sa réponse angulaire d est égale aux K premiers coefficients de sa série de Fourier. La réponse angulaire d’un filtre exactement en quadrature sera par définition toujours nulle sur la moitié de sa période. Cette structure « en créneau » correspond à une série de Fourier *infinie*, ce qui est rédhibitoire en pratique car il faudrait une infinité de filtres. C’est pourquoi les filtres de H. Knutsson étudiés plus haut ne sont pas orientables, et pourquoi nous avons du ici définir une approximation de $\mathcal{H}_0^{part} g_a$. Notre approximation est assez pauvre car elle a peu de degrés de liberté, puisque nous avons choisi de minimiser le nombre de filtres. On constate expérimentalement que l’analyse donnée est satisfaisante, mais l’utilisation de filtres approximativement en quadrature n’assure pas théoriquement une mesure exacte de l’amplitude et de la phase. Plus généralement, le nombre

de filtres de base et leurs orientations sont flexibles, mais conditionnent les équations de combinaison ainsi que la qualité d'analyse, ce qui rend la procédure très paramétrique, et donc soulève une question d'optimisation. De plus, le *minimum* de 5 filtres nécessaires pour pouvoir extraire la phase du signal peut être considéré comme contraignant, par rapport aux seuls 2 filtrages d'un simple gradient. Notons que la dimension numérique consiste en un simple échantillonnage des réponses impulsionnelles, ce qui limite cet outil à l'analyse. Enfin, les méthodes présentées dans ce début de chapitre sont implicitement basées sur la transformée de Hilbert partielle, à partir de laquelle aucune notion de signal analytique satisfaisante n'a pu être construite, comme nous l'avons évoqué à la section 1.2.3. Il manque à ces outils un formalisme global qui permettrait de définir des *filtres en quadrature isotropes*, à travers une nouvelle définition de signal analytique 2D. C'est pourquoi nous proposons maintenant d'étudier le signal monogène.

3.2 UN SIGNAL ANALYTIQUE ISOTROPE

Cette section reprend la construction et l'interprétation du *signal monogène* [47], généralisation isotrope du signal analytique, ainsi que son intégration dans un formalisme global de traitement du signal, utilisant les algèbres de Clifford.

3.2.1 Signal monogène

La généralisation mathématique 2D du signal analytique est astucieusement conduite dans le contexte des *fonctions analytiques*. Ceci aboutit à l'apparition de la transformée de Riesz, comme nouvel opérateur de déphasage généralisant la transformée de Hilbert 1D. Le signal monogène permet enfin de définir des filtres en quadrature 2D isotrope, constitués de *trois* composantes.

Equations de Cauchy-Riemann généralisées

Nous reprenons ici les grandes lignes de la généralisation 2D du signal analytique dans le contexte de l'*analyse complexe*. Nous réferrons le lecteur à [47, 140, 109] pour plus de détails. Le signal analytique que nous avons présenté à la section 1.1.2 peut être défini de façon traditionnelle comme la restriction à l'axe réel d'une *fonction holomorphe*⁵. Une fonction holomorphe est une fonction complexe d'une variable complexe $f(x + jy) = f_R(x + jy) + j f_I(x + jy)$ qui est dérivable, la notion de dérivabilité étant définie dans ce cas par les équations de Cauchy-Riemann :

$$\left\{ \frac{\partial f_R}{\partial x} = \frac{\partial f_I}{\partial y} \quad \text{et} \quad \frac{\partial f_R}{\partial y} = -\frac{\partial f_I}{\partial x} \right\} \Leftrightarrow \left\{ \nabla_{\mathbb{C}} f = 0 \quad \text{avec} \quad \nabla_{\mathbb{C}} = \left(\frac{\partial}{\partial x} + j \frac{\partial}{\partial y} \right) \right\} \quad (3.29)$$

L'équation de droite est appelée « équation de Dirac », et met en jeu un « gradient complexe » $\nabla_{\mathbb{C}}$. Voici maintenant un lien avec les fonctions harmoniques. Une fonction harmonique p est définie par :

$$p : \mathbb{R}^2 \mapsto \mathbb{R} \quad \text{et} \quad \Delta p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = 0 \quad (3.30)$$

Dans ce cas, on peut vérifier que le gradient (conjugué) de p est nécessairement une fonction holomorphe :

$$\Delta p = 0 \quad \text{et} \quad f = \frac{\partial p}{\partial x} - j \frac{\partial p}{\partial y} \Rightarrow \nabla_{\mathbb{C}} f = 0 \quad (3.31)$$

On dit dans la terminologie de la théorie du potentiel que f « dérive d'un potentiel harmonique ».

Étant donné une fonction 1D réelle $s(x)$, il existe une unique fonction holomorphe f telle que $\Re\{f(x + j0)\} = s(x)$ (Rigoureusement, le domaine de définition de f est limité à $y > 0$, et les

5. Ces fonctions sont équivalentes aux fonctions « analytiques » complexes, à l'origine du nom de « signal analytique ».

valeurs pour $y = 0$ sont obtenues par continuité [78, 140]). La transformée de Hilbert apparait comme la relation entre la partie réelle et la partie imaginaire de f pour une valeur fixée de y :

$$\Im\{f(x + jy_0)\} = \mathcal{H}\{\Re\{f(\cdot + jy_0)\}\}(x) \quad (3.32)$$

Pour étendre la transformée de Hilbert et donc la notion de signal analytique en 2D, on peut donc tenter d'étendre le concept de fonctions holomorphes en 3D. Alors que les équations de Cauchy-Riemann sont « enfermées » dans l'algèbre des complexes, leur ré-écriture en termes de gradient et de Laplacien à travers la théorie du potentiel est plus intuitive pour des dimensions supérieures. Il suffit de définir f comme le gradient 3D d'une fonction harmonique 3D, en utilisant une écriture vectorielle :

$$f(x, y, z) = \left[\frac{\partial p}{\partial x} \quad \frac{\partial p}{\partial y} \quad \frac{\partial p}{\partial z} \right] \quad \text{et} \quad \Delta p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} = 0 \quad (3.33)$$

On peut en déduire les équations de Cauchy-Riemann généralisées correspondantes :

$$\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 0 \quad \text{et} \quad \frac{\partial f_3}{\partial y} = \frac{\partial f_2}{\partial z} \quad \text{et} \quad \frac{\partial f_1}{\partial z} = -\frac{\partial f_3}{\partial x} \quad \text{et} \quad \frac{\partial f_2}{\partial x} = -\frac{\partial f_1}{\partial y} \quad (3.34)$$

Les fonctions de trois variables à valeurs dans $\mathbb{R}^3$ vérifiant ces équations sont appelées *fonctions monogènes* [47] (notamment en analyse de Clifford). On obtient que pour une valeur fixée de z , les trois composantes d'une fonction monogène sont liées par une transformée de Riesz, qui apparait alors comme une généralisation 2D de la transformée de Hilbert. Cette généralisation des équations de Cauchy-Riemann a été proposée notamment en 1984 dans [109] pour l'étude du champ magnétique de la terre par des champs de vecteurs « potentiels » 3D. Nous allons voir que la transformée de Riesz se justifie parfaitement d'un point de vue traitement du signal, comme généralisation isotrope de la transformée de Hilbert.

Transformée de Riesz

La transformée de Riesz, ainsi nommée d'après le mathématicien Marcel Riesz, est depuis longtemps l'extension multi-dimensionnelle de la transformée de Hilbert dans la communauté mathématique [140, 17, 121]. Appliquée à une fonction réelle de deux variables, elle se définit par deux transformées indépendantes liées aux axes x et y , définies par :

$$\mathcal{R}_0 s(\mathbf{x}) = s_{r_1}(\mathbf{x}) = p.v. \int \frac{\tau_1}{2\pi \|\boldsymbol{\tau}\|^3} s(\mathbf{x} - \boldsymbol{\tau}) d\boldsymbol{\tau} \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{\omega_1}{\|\boldsymbol{\omega}\|} \hat{s}(\boldsymbol{\omega}) \quad (3.35)$$

$$\mathcal{R}_{\frac{\pi}{2}} s(\mathbf{x}) = s_{r_2}(\mathbf{x}) = p.v. \int \frac{\tau_2}{2\pi \|\boldsymbol{\tau}\|^3} s(\mathbf{x} - \boldsymbol{\tau}) d\boldsymbol{\tau} \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{\omega_2}{\|\boldsymbol{\omega}\|} \hat{s}(\boldsymbol{\omega}) \quad (3.36)$$

Cette définition qui ressemble dans la forme à la transformée de Hilbert est elle aussi une valeur principale de Cauchy. Notons que la transformée de Riesz n'est singulière qu'en un *point* $\boldsymbol{\omega} = 0$, contrairement à la transformée de Hilbert partielle qui est singulière sur toute la ligne d'orientation θ (voir section 1.3.1). Dans cette thèse, on utilisera une écriture combinée des deux composantes dans un nombre complexe :

$$\mathcal{R} s(\mathbf{x}) = \mathcal{R}_0 s(\mathbf{x}) + j \mathcal{R}_{\frac{\pi}{2}} s(\mathbf{x}) \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{\omega_1 + j\omega_2}{\|\boldsymbol{\omega}\|} \hat{s}(\boldsymbol{\omega}) \quad (3.37)$$

Propriétés

Cette écriture permet de voir $\mathcal{R}$ comme un opérateur, dont on peut montrer qu'il est unitaire [144], c'est-à-dire que son inverse est égal à son adjoint et vaut :

$$\mathcal{R}^{-1} s \quad \xleftrightarrow{\mathcal{F}} \quad -j \frac{-\omega_1 + j\omega_2}{\|\boldsymbol{\omega}\|} \hat{s}(\boldsymbol{\omega}) \quad (3.38)$$

On a également les propriétés d'invariance par translation et changement d'échelle, c'est-à-dire que $\mathcal{R}\{s(\mathbf{a} \cdot + \mathbf{b})\}(\mathbf{x}) = \mathcal{R}\{s(\cdot)\}(\mathbf{a}\mathbf{x} + \mathbf{b})$, ainsi que l'invariance par rotation à un facteur constant près :

$$\mathcal{R}\{s(\mathbf{R}_\theta \cdot)\}(\mathbf{x}) = e^{j\theta} \mathcal{R}\{s(\cdot)\}(\mathbf{R}_\theta \mathbf{x}) \quad \text{avec} \quad \mathbf{R}_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (3.39)$$

Notons que le facteur est d'amplitude 1, ce qui implique l'invariance stricte du module : $|\mathcal{R}\{s(\mathbf{R}_\theta \cdot)\}(\mathbf{x})| = |\mathcal{R}\{s(\cdot)\}(\mathbf{R}_\theta \mathbf{x})|$. Comme il existe un lien entre $\mathcal{H}$ et la dérivation (section 1.1.2, équation (1.8)), il existe aussi un lien fort instructif entre $\mathcal{R}$ et le gradient complexe $\nabla_{\mathbb{C}}$. Cela nécessite l'introduction de l'opérateur *Laplacien fractionnel* [144, 57], qui se définit dans le domaine fréquentiel par :

$$(-\Delta)^\alpha \xleftrightarrow{\mathcal{F}} \|\omega\|^{2\alpha} \quad (3.40)$$

Alors que le Laplacien classique ($\alpha = 1$) a plutôt un comportement de filtre passe-haut, le Laplacien fractionnel peut aussi bien avoir un comportement passe-bas (pour $\alpha < 0$), et toujours isotrope. Notons dans ce cas la singularité pour $\omega = 0$. Par exemple, on peut interpréter $(-\Delta)^{-\frac{1}{2}}$ comme un opérateur de lissage isotrope, de gain fréquentiel « en $1/\omega$ ». Cet opérateur permet d'établir le lien suivant :

$$\mathcal{R}s = \nabla_{\mathbb{C}} (-\Delta)^{-\frac{1}{2}} s \quad (3.41)$$

où $\nabla_{\mathbb{C}}$ est le gradient complexe (voir l'équation (3.29)). Ce lien, qui s'établit facilement dans le domaine fréquentiel, montre que la transformée de Riesz effectue la même opération que le gradient sur une version *lissée* de l'image : $(-\Delta)^{-\frac{1}{2}} s(\mathbf{x})$. Comme le Laplacien fractionnel est isotrope, on peut considérer qu'il n'altère pas l'orientation des structures locales de l'image. C'est pourquoi on considère que le module et l'argument de $\mathcal{R}s$ portent une information analogue à un gradient, c'est-à-dire :

$$|\mathcal{R}s| = \mathcal{N}_{\mathcal{R}} \equiv \mathcal{N}_{\nabla} \quad \arg\{\mathcal{R}s\} = \theta_{\mathcal{R}} \equiv \theta_{\nabla} \quad (3.42)$$

où $\mathcal{N}_{\nabla}$ et θ_{∇} ont été définies à l'équation (3.1). Notons également que $\mathcal{R}_0$ et $\mathcal{R}_{\frac{\pi}{2}}$ ont une réponse spectrale séparable en coordonnées polaires, et forment une base de filtres orientables, ce qui les place dans la continuité des outils présentés à la section 3.1.

Voyons maintenant comment identifier une analyse type Hilbert dans la transformée de Riesz. Pour cela, considérons une onde plane d'amplitude A et de fréquence ξ , orientée vers α :

$$s(\mathbf{x}) = A \cos(\boldsymbol{\xi}^T \mathbf{x}) \quad \text{avec} \quad \boldsymbol{\xi}^T = \xi [\cos(\alpha) \quad \sin(\alpha)] \quad (3.43)$$

On montre facilement que la transformée de Riesz correspond à :

$$\mathcal{N}_{\mathcal{R}}(\mathbf{x}) = A \left| \sin(\boldsymbol{\xi}^T \mathbf{x}) \right| \quad \text{et} \quad \theta_{\mathcal{R}} = \alpha \quad (3.44)$$

On voit que, tout comme la transformée de Hilbert, $\mathcal{R}$ transforme les cosinus en sinus, ce qui correspond à un « déphasage pur » de $-\frac{\pi}{2}$, c'est-à-dire sans amplification (filtre « passe-tout »). Ce déphasage s'opère de façon invariante à l'orientation du signal α , ce qui fait de $\mathcal{R}$ une généralisation *isotrope* de la transformée de Hilbert d'un point de vue traitement du signal.

Signal monogène

Un signal analytique 2D à trois composantes peut donc être défini. Il s'agit du signal monogène, qui consiste à juxtaposer algébriquement au signal d'origine, sa transformée de Riesz, et à passer en coordonnées sphériques 3D de la façon suivante [47] :

$$s_M = \begin{bmatrix} s \\ \mathcal{R}_0 s \\ \mathcal{R}_{\frac{\pi}{2}} s \end{bmatrix} = A_{\mathcal{R}} \begin{bmatrix} \cos(\varphi_{\mathcal{R}}) \\ \sin(\varphi_{\mathcal{R}}) \cos(\theta_{\mathcal{R}}) \\ \sin(\varphi_{\mathcal{R}}) \sin(\theta_{\mathcal{R}}) \end{bmatrix} \quad \begin{aligned} A_{\mathcal{R}} &= \sqrt{s^2 + \mathcal{N}_{\mathcal{R}}^2} \text{ (amplitude)} \\ \theta_{\mathcal{R}} &= \arg\{\mathcal{R}s\} \in [-\pi; \pi] \text{ (orientation)} \\ \varphi_{\mathcal{R}} &= \arg\{s + j\mathcal{N}_{\mathcal{R}}\} \in [0; \pi] \text{ (phase)} \end{aligned} \quad (3.45)$$

Pour que cette analyse soit interprétable, il faut que ses paramètres soient *localement* constants, c'est-à-dire qu'ils doivent varier lentement devant la fréquence du signal. Le modèle AM/FM sous-jacent $s(\mathbf{x}) = A(\mathbf{x}) \cos(\varphi(\mathbf{x}))$ est alors localement celui d'une onde plane. Cette restriction correspond à l'hypothèse de signal « localement simple » (une seule orientation à la fois) ainsi qu'à la notion de *bande limitée*, car la fréquence doit bien sûr varier lentement devant $\mathbf{x}$. C'est pourquoi le signal monogène est restreint à un modèle i1D d'une part, et toujours utilisé conjointement à un filtrage passe-bande d'autre part.

Cette analyse a été utilisée dans la communauté optique, implicitement (sous le nom de transformée de Hilbert radiale) dans [35], pour la détection de contours, et explicitement à travers le filtre « spiral phase quadrature » de [90, 89], pour la démodulation d'interferogrammes. Voyons maintenant son utilisation à travers de nouveaux filtres en quadrature.

Filtres en quadrature sphériques

Tout comme le signal analytique, le signal monogène ne réalise aucune sélection fréquentielle et doit donc être utilisé dans une procédure de décomposition en sous-bandes de support fréquentiel étroit. Contrairement aux décompositions classiques utilisant un découpage fréquentiel directionnel, nous allons cette fois utiliser un découpage *isotrope*. Les sous-bandes contiendront des structures de toutes orientations, qui seront analysées spatialement par la transformée de Riesz. Un filtre en quadrature « sphérique » sera donc construit par extension monogène d'un filtre passe-bande isotrope.

Par exemple, M. Felsberg suggère d'utiliser des différences de noyau de Poisson (DoP) [47] :

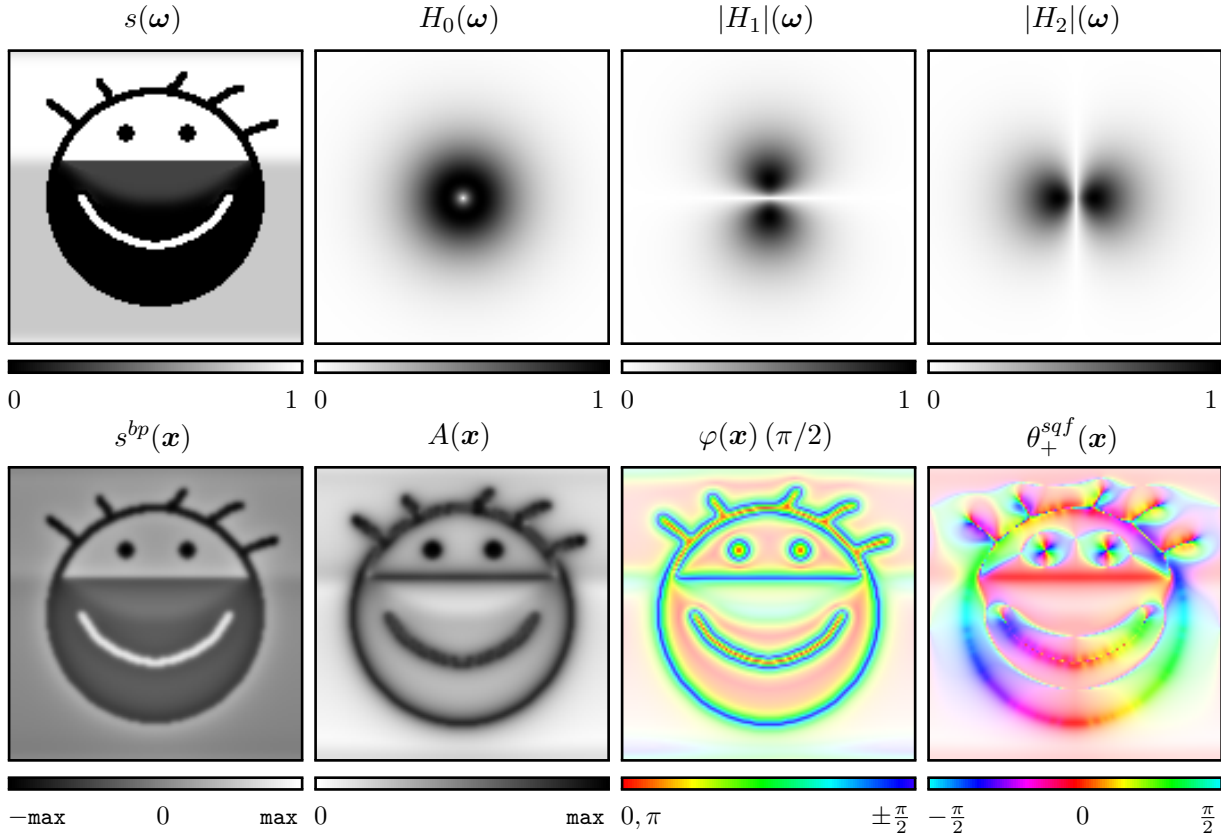
$$h(\mathbf{x}) = p_{s_1} - p_{s_2} \quad \text{avec} \quad p_s(\mathbf{x}) = \frac{s}{2\pi(s^2 + |\mathbf{x}|^2)^{3/2}} \quad \xleftrightarrow{\mathcal{F}} \quad e^{-2\pi s \|\omega\|} \quad (3.46)$$

arragés en octaves autour de la fréquence « $2^k \frac{2\pi}{N}$ », ce qui correspond aux échelles $s_1 = N \log(2)/(2^k 2\pi)$ et $s_2 = N \log(2)/(2^{k-1} 2\pi)$. La variable k est alors l'indice de l'octave analysé. On forme le filtre en quadrature sphérique $[h \ h_{R_1} \ h_{R_2}]$, avec $h_{R_1} + \mathbf{j}h_{R_2} = \mathcal{R}h$. Nous illustrons à la figure 3.3 cette analyse monogène pour $k = 3$ ⁶. L'implantation des filtrages (continus et à support infini) est faite par FFT (avec $\omega \equiv n \frac{2\pi}{N}$). Le résultat est très semblable aux filtres en quadrature orientables de la figure 3.2, mais ne nécessite cette fois que trois filtres (au lieu de 5 minimum), et correspond à un concept satisfaisant de signal analytique 2D. Contrairement aux atomes de Gabor, les filtres DoP ont une moyenne nulle et forment un triplet de Riesz exact, de façon analogue aux différences de Gaussiennes (« DoG »). En revanche ils n'optimisent pas la localisation espace-fréquence. Notons que l'exactitude de la transformée de Riesz n'est pas en conflit avec l'orientabilité de l'analyse, alors que cela posait problème avec la transformée de Hilbert partielle.

En faisant varier les échelles, on peut obtenir une décomposition en ondelettes monogène continue discrétisée [112, 113], comme celle illustrée à la figure 3.4. Cela permet de voir par exemple qu'une ligne comme celle de la flèche dans le panneau est bien reconnue à une certaine échelle ($4 \frac{2\pi}{N}$), qui affiche clairement une phase proche de 0 (rouge). C'est à cette même échelle que la silhouette de l'arbre central est détectée comme un bord continu. En revanche, aux échelles plus fines, le bord de l'arbre se noie dans les détails des autres feuilles, alors que les troncs des arbres en arrière-plan deviennent significatifs. La décomposition est très redondante, et ne permet pas de reconstruire le signal. En revanche, la quantité de données est intéressante car elle est équivalente à celle d'une analyse de Gabor 2D sur seulement 3 orientations, alors que l'information d'orientation est ici bien plus riche. Notons qu'une extension du signal monogène a été proposée dans [52], prenant en compte des propriétés d'ordre supérieur comme la courbure.

Le signal monogène est pour l'instant clairement limitée à l'*analyse* en termes applicatifs. Il a été utilisé en segmentation [7, 74], en classification de texture [156], en estimation de mouvement [103] et en estimation de carte de disparité stéréoscopique [96]. Notons également une extension

6. Cet octave correspond bien à la largeur des structures de notre image de test

FIGURE 3.3 – *Filtres en quadrature sphériques « DoP » de M. Felsberg (Voir aussi figure 3.2).*

3D, pour l'analyse d'images biomédicales [29]. Dans notre travail, nous souhaitons inclure les applications de *traitement*, qui nécessitent un schéma de reconstruction de l'image à partir des coefficients d'analyse. Pour cela, à la lumière du formalisme quaternionique de T. Bülow qui a permis de définir des ondelettes quaternioniques inversibles, un système complet d'outils « signal » doit être construit autour du signal monogène. La seule proposition est celle de M. Felsberg [47], que nous décrivons maintenant.

3.2.2 Dépasser les nombres complexes

Même si les problématiques de traitement de signal considèrent la plupart du temps des signaux à valeurs réelles, le formalisme de Fourier a imposé l'algèbre des nombres complexes, puisque les valeurs propres des systèmes linéaires invariants sont les sinusoides complexes. Au delà de cette apparente contrainte, l'algèbre des nombres complexes apporte une écriture mathématique simple pour caractériser une sinusoïde (amplitude et phase encodées dans un nombre), et pour appliquer simultanément une pondération et un déphasage (multiplication par une exponentielle complexe).

Nous avons vu au chapitre 2, avec le formalisme quaternionique de T. Bülow, que l'étude des signaux multi-dimensionnels se voit limitée par cette algèbre, et ce en dépit d'une définition bien établie de la transformée de Fourier nD . Pour tenter de surpasser le problème de variance par rotation du système quaternionique, M. Felsberg propose [47] d'utiliser les *algèbres géométriques*.

Algèbres géométriques

Les algèbres géométriques forment une structure mathématique reposant sur l'algèbre de Clifford et qui a pour but de créer un langage géométrique de haut niveau, manipulant de façon unifiée le produit scalaire, le produit vectoriel, les rotations *etc.* sur des objets de dimension quelconque comme des « segments de droite orientés », ou des « quarts de plan orientés »,

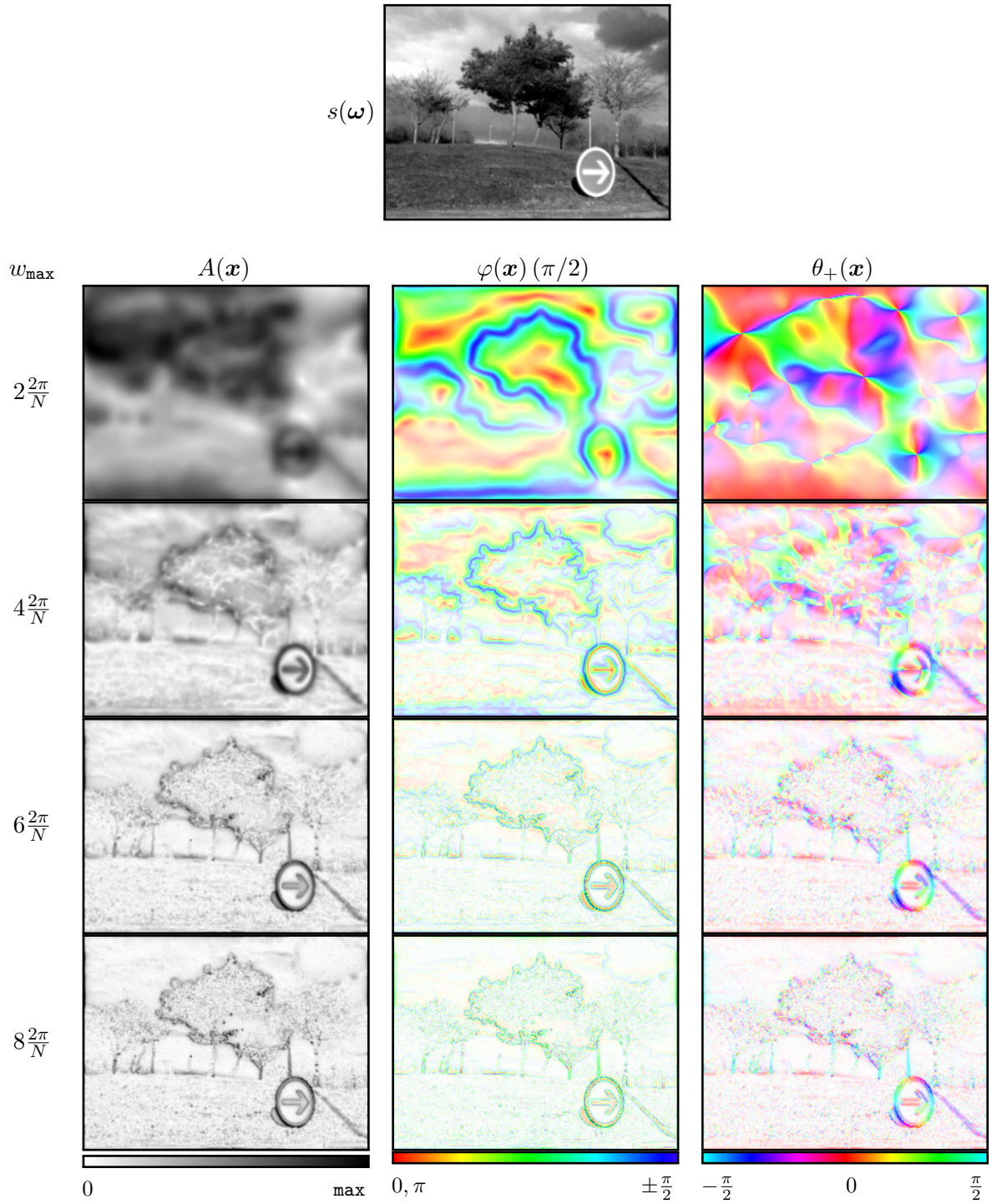


FIGURE 3.4 – Analyse monogène multi-échelle par ondelettes DoP.

plus généraux que de simples vecteurs. La popularité de cet outil fondamental grandit dans de nombreux domaines de la physique et des sciences de l'ingénieur pour sa capacité à ré-écrire simplement des opérations géométriques sophistiquées dans des espaces de dimension élevée, par des expressions algébriques concises.

Nous allons tenter de décrire de façon concise la construction d'une telle algèbre, en nous focalisant sur les éléments qui nous serviront pour la suite de ce chapitre. Le lecteur trouvera plus de précisions dans [25, 116].

On part d'un espace vectoriel E sur $\mathbb{R}$ de dimension n , et d'une forme quadratique $Q : E \mapsto \mathbb{R}$. La détermination de l'algèbre dépend de la définition du produit (loi de composition interne) que l'on appellera *produit géométrique*. On souhaite que ce produit vérifie en particulier $uu = Q(u)$, $u \in E$. En considérant ensuite la forme bilinéaire symétrique B qui est associée à Q , on a donc nécessairement : $2B(u, v) = Q(u+v) - Q(u) - Q(v) = uv + vu$, ce qui constitue une première partie de définition de « uv ». Pour compléter cette définition, on introduit l'opération *anti-symétrique* $u \wedge v = (1/2)(uv - vu)$, qui permet de définir le *produit géométrique* entre u et v par :

$$uv = B(u, v) + u \wedge v \quad (3.47)$$

(il doit être également associatif et distributif). En conséquence de cette construction, de nouveaux objets sont introduits de fait. En effet, en prenant l'exemple où Q est une norme euclidienne sur $\mathbb{R}^n$ et B est un produit scalaire, le produit uv de deux vecteurs *orthogonaux* entraîne :

$$\begin{aligned} B(u, v) = 0 &\Rightarrow uv = -vu \\ &\Rightarrow (uv)^2 = -(uv)(vu) = -u(vv)u = -Q(u)Q(v) \in \mathbb{R}_- \end{aligned} \quad (3.48)$$

L'objet construit par ce produit a donc un carré scalaire et *négatif*, il ne peut donc être ni un scalaire, ni un vecteur. Le *produit géométrique* introduit donc une notion de *multivecteur*, qui englobe les scalaires et les vecteurs comme des cas particuliers. Alors qu'un vecteur symbolise un axe et direction dans l'espace, un *bivecteur* - produit de deux vecteurs - symbolise un *plan orienté*.

Dans notre travail, on se restreindra à la forme quadratique $Q(u) = \sum_{i=1}^n u_i^2$. Une *base* de l'algèbre géométrique peut se construire par extension d'une base $(e_1, \dots, e_n)$ de E .

$$e_{ij} \stackrel{\text{def}}{=} e_i \wedge e_j \quad (3.49)$$

qui définit les « bivecteurs » e_{ij} , vus comme des bases de sous-espaces 2D de $\mathbb{R}^n$. Les « trivecteurs » sont ensuite construits par $e_{ijk} = e_i \wedge e_j \wedge e_k$, et ainsi de suite jusqu'à obtenir l'unique n -vecteur : $I_n = e_1 \wedge \dots \wedge e_n$ qu'on appelle *pseudoscalaire*. En ajoutant l'élément scalaire $\mathbf{1}$, on forme la base de $\mathcal{G}_n$. Pour $n = 3$, on obtient trois *bivecteurs* distincts e_{12} , e_{23} et e_{31} , et le pseudoscalaire $e_{123} = I_3$. Tout multivecteur s'écrit donc :

$$M = a + be_1 + ce_2 + de_3 + ee_{12} + fe_{23} + ge_{31} + hI_3 \in \mathcal{G}_3 \quad (a, b, c, d, e, f, g \in \mathbb{R}) \quad (3.50)$$

Pour $n = 3$, les différents produits géométriques sont récapitulés dans la table 3.1, et permettent de calculer le produit entre deux multivecteurs quelconques par distribution.

Les *spineurs* de $\mathcal{G}_3$ sont les multivecteurs de la forme :

$$s = a + be_{23} + ce_{31} + de_{12} \quad (3.52)$$

Ils forment une sous-algèbre isomorphe aux quaternions et permettent donc de représenter efficacement les rotations 3D. L'action de rotation et amplification par un spineur sur un vecteur s'écrit par un produit « en sandwich » :

$$(y_1 e_1 + y_2 e_2 + y_3 e_3) = (a - be_{23} - ce_{31} - de_{12})(x_1 e_1 + x_2 e_2 + x_3 e_3)(a + be_{23} + ce_{31} + de_{12}) \quad (3.53)$$

et correspond à une rotation autour de l'axe 3D $(be_1 + ce_2 + de_3)$, avec un facteur d'amplification de $\sqrt{a^2 + b^2 + c^2 + d^2}$. Voyons maintenant comment reformuler les outils classiques de traitement du signal dans cette algèbre.

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$-\mathbf{e}_{13}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_{23}$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$\mathbf{e}_{23}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{31}$
$\mathbf{e}_3$	$\mathbf{e}_3$	$\mathbf{e}_{31}$	$-\mathbf{e}_{23}$	1	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$\mathbf{e}_{12}$
$\mathbf{e}_{23}$	$\mathbf{e}_{23}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	-1	$-\mathbf{e}_{12}$	$\mathbf{e}_{31}$	$-\mathbf{e}_1$
$\mathbf{e}_{31}$	$\mathbf{e}_{31}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{12}$	-1	$-\mathbf{e}_{23}$	$-\mathbf{e}_2$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$-\mathbf{e}_{31}$	$\mathbf{e}_{23}$	-1	$-\mathbf{e}_3$
I_3	I_3	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	$-\mathbf{e}_1$	$-\mathbf{e}_2$	$-\mathbf{e}_3$	-1

(3.51)

TABLE 3.1 – Table de multiplication des éléments de base de $\mathcal{G}_3$, par le produit géométrique. Le terme de gauche est à reporter dans la première colonne, et le terme de droite dans la première ligne (ce produit n'est pas commutatif).

Le formalisme de M. Felsberg

En dehors du choix de l'algèbre $\mathcal{G}_3$ pour la modélisation des outils de traitement du signal 2D, la première principale nouveauté du formalisme de M. Felsberg est la distinction algébrique entre les *signaux* et les *filtres*.

Les signaux sont à valeurs vectorielles :

$$s(\mathbf{x}) = s_1(\mathbf{x})\mathbf{e}_1 + s_2(\mathbf{x})\mathbf{e}_2 + s_3(\mathbf{x})\mathbf{e}_3 \quad (3.54)$$

Les signaux « réels » habituels sont encodés dans l'axe $\mathbf{e}_3$ ($s_1 = s_2 = 0$), et des signaux « plus complexes » peuvent être construit avec $s_1, s_2 \in \mathbb{R}$, formant une sorte de « partie imaginaire » sur les axes $\mathbf{e}_1$ et $\mathbf{e}_2$.

Les filtres sont des *spineurs*, et la notion de « partie imaginaire » est cette fois analogue à celle des quaternions. Un filtre « réel » sera modélisé par un spineur scalaire $h = h_1$, qui appliquera une simple amplification sans rotation des vecteurs du signal, et donc en particulier aura une sortie réelle pour une entrée réelle.

L'application d'un filtre sur un signal est définie par la convolution suivante :

$$\underbrace{(h * s)(\mathbf{x})}_{\text{vecteur}} = \int_{\mathbb{R}^2} \underbrace{h(\mathbf{x}')}_{\text{spineur}} \underbrace{s(\mathbf{x} - \mathbf{x}')}_{\text{vecteur}} d\mathbf{x}' \quad (3.55)$$

Le produit dans l'intégrale est un *produit géométrique*, ce qui implique une distribution particulière des différents termes. Nous souhaitons que le résultat du filtrage d'un signal soit un autre signal, en d'autres termes, le produit de h par s doit aboutir à un objet de type *vecteur*. Notons que l'application de la rotation d'un spineur sur un vecteur se fait normalement par un produit « en sandwich », comme expliqué à la section précédente, ce qui n'est pas toujours équivalent au produit « à gauche » utilisé ici par M. Felsberg dans cette définition. En conséquence, l'interprétation de cette convolution comme une rotation des coefficients n'est pas toujours valable (le résultat peut contenir un terme pseudoscalaire). M. Felsberg propose de se restreindre au cas où le signal n'a de valeur que sur $\mathbf{e}_3$ et où le filtre est nul sur le plan $\mathbf{e}_{12}$, afin de conserver cette définition de la convolution plus proche de la définition classique, tout en gardant la notion d'action d'un spineur sur un vecteur. Dans le cas des filtres réels, l'opération se réduit à trois convolutions réelles classiques indépendantes sur les trois axes.

La transformée de Fourier proposée est une reformulation de la version classique, associée à une base d'ondes planes orientées :

$$\hat{s}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} s(\mathbf{x}) e^{-I_3 \boldsymbol{\omega}^\top \mathbf{x}} d\mathbf{x} \quad (3.56)$$

L'imaginaire complexe est donc remplacé par le pseudoscalaire I_3 . Les propriétés classiques sont bien sûr conservées, en particulier le théorème du décalage :

$$s(\mathbf{x} - \mathbf{x}') \xleftrightarrow{\mathcal{F}} \hat{s}(\boldsymbol{\omega}) e^{-I_3 \boldsymbol{\omega}^\top \mathbf{x}'} \quad (3.57)$$

Cette transformée est valable pour une fonction s à valeurs quelconques dans $\mathcal{G}_3$. Un « signal » (donc à valeurs vectorielles) aura une transformée de Fourier particulière de la forme :

$$\hat{s} = \Re\{\hat{s}_1\}\mathbf{e}_1 + \Im\{\hat{s}_1\}\mathbf{e}_{23} + \Re\{\hat{s}_2\}\mathbf{e}_2 + \Im\{\hat{s}_2\}\mathbf{e}_{31} + \Re\{\hat{s}_3\}\mathbf{e}_3 + \Im\{\hat{s}_3\}\mathbf{e}_{12} \quad (3.58)$$

directement exprimable par les parties réelles et imaginaires des transformées de Fourier complexes $\hat{s}_i$.

Le spectre d'un « filtre » (spineur) verra ses composantes de Fourier réparties sur des axes différents :

$$\begin{aligned} \hat{h} = \Re\{\hat{h}_0\} + \Im\{\hat{h}_0\}I_3 + \Re\{\hat{h}_{23}\}\mathbf{e}_{23} + \Im\{\hat{h}_{23}\}\mathbf{e}_1 \\ + \Re\{\hat{h}_{31}\}\mathbf{e}_{31} + \Im\{\hat{h}_{31}\}\mathbf{e}_2 + \Re\{\hat{h}_{12}\}\mathbf{e}_{12} + \Im\{\hat{h}_{12}\}\mathbf{e}_3 \end{aligned} \quad (3.59)$$

Le spectre d'un signal filtré est égal au produit géométrique du spectre du filtre par celui du signal « $\hat{h}\hat{s}$ » (dans cet ordre, le produit n'est pas commutatif). Dans le cas restreint d'un signal sur $\mathbf{e}_3$ et d'un filtre nul sur $\mathbf{e}_{12}$, on retrouve les trois spectres d'un signal réel filtré par trois filtres indépendamment :

$$\begin{aligned} (h_1 + h_{23}\mathbf{e}_{23} + h_{31}\mathbf{e}_{31}) * (s_3\mathbf{e}_3) \xleftrightarrow{\mathcal{F}} \Re\{\hat{h}_{31}\hat{s}_3\}\mathbf{e}_1 + \Im\{\hat{h}_{31}\hat{s}_3\}\mathbf{e}_{23} + \Re\{\hat{h}_{23}\hat{s}_3\}\mathbf{e}_2 \\ + \Im\{\hat{h}_{23}\hat{s}_3\}\mathbf{e}_{31} + \Re\{\hat{h}_1\hat{s}_3\}\mathbf{e}_3 + \Im\{\hat{h}_1\hat{s}_3\}\mathbf{e}_{12} \end{aligned} \quad (3.60)$$

Notons que contrairement au formalisme des quaternions, le concept de phase 2D développé ne se retrouve pas directement dans les atomes de Fourier. En conséquence, nous n'avons pas pu à ce stade étudier la forme des atomes monogènes, comme nous l'avons fait pour les atomes quaternioniques (section 2.2.1). En outre, la transformée de Fourier inverse n'est pas étudiée dans la thèse de M. Felsberg⁷, ce qui est principalement dû au fait que la notion centrale du formalisme est la définition d'un signal analytique 2D isotrope, dédié à l'analyse.

Le *signal monogène* s_M associé à un signal $s = s_3\mathbf{e}_3$ est défini dans $\mathcal{G}_3$ par :

$$s_M = s_1\mathbf{e}_1 + s_2\mathbf{e}_2 + s_3\mathbf{e}_3 \quad \text{avec} \quad s_1 + \mathbf{j}s_2 = \mathcal{R}\{s_3\} \quad (3.61)$$

Cette formulation est parfaitement analogue au signal analytique classique qui consiste à juxtaposer une version déphasée du signal de départ dans une sorte de « partie imaginaire ».

Le spectre de s_M est

$$\hat{s}_M = \Re\{\hat{s}_3\} \left(\mathbf{e}_3 + \frac{-\omega_1}{\|\boldsymbol{\omega}\|}\mathbf{e}_{23} + \frac{\omega_2}{\|\boldsymbol{\omega}\|}\mathbf{e}_2 \right) + \Im\{\hat{s}_3\} \left(\mathbf{e}_{12} + \frac{\omega_1}{\|\boldsymbol{\omega}\|}\mathbf{e}_1 + \frac{\omega_2}{\|\boldsymbol{\omega}\|}\mathbf{e}_{31} \right) \quad (3.62)$$

L'amplitude est définie par une norme Euclidienne analogue au module complexe $A = \sqrt{s_3^2 + s_1^2 + s_2^2}$, et la phase est définie de façon à pouvoir écrire le signal monogène en fonction de l'exponentielle de cette phase : $s_M = Ae^r$. On obtient par la définition du logarithme des spineurs ([47] page 30) la phase vectorielle suivante :

$$r = \frac{s_2\mathbf{e}_1 - s_1\mathbf{e}_2}{\sqrt{s_1^2 + s_2^2}} \arg\left\{s + \mathbf{j}\sqrt{s_1^2 + s_2^2}\right\} = \varphi_{\mathcal{R}}(\sin(\theta_{\mathcal{R}})\mathbf{e}_1 - \cos(\theta_{\mathcal{R}})\mathbf{e}_2) \quad (3.63)$$

Ce vecteur 2D a pour direction dans l'espace la direction de la transformée de Riesz, et pour amplitude la phase entre la transformée de Riesz et le signal de départ.

7. Elle est étudiée dans le contexte de l'analyse Cliffordienne [13], mais sans lien *a priori* avec le signal monogène

3.2.3 Discussion

Le formalisme du signal monogène exprimé dans l'algèbre géométrique $\mathcal{G}_3$ permet de généraliser la notion de signal analytique de façon isotrope, et de ré-écrire les outils classiques de façon à prendre en compte des signaux à trois composantes comme le signal monogène [47]. Une analyse multi-bande monogène peut être réalisée par un triplet de filtres liés par la transformée de Riesz, et permet de mesurer efficacement l'orientation et la phase des structures locales.

Dans la suite de notre travail, la formulation du signal monogène dans $\mathcal{G}_3$ ne nous a pas apporté par rapport à l'écriture dans $\mathbb{C}$, c'est pourquoi par simplicité nous n'utiliserons plus cette écriture dans le reste de cette thèse, sauf à la section 3.4.1.

La définition d'ondelettes monogènes est un sujet ouvert, qui semble ne pas encore avoir trouvé de solution aussi aboutie numériquement que le Dual-Tree 1D ou la QWT. En effet, ces derniers sont associés à des formalismes dans lesquels les atomes de Fourier *sont* eux-mêmes des signaux « analytiques ». On peut donc d'une part construire des ondelettes analytiques par un mécanisme de fenêtrage, et d'autre part définir une méthode de reconstruction en répliquant la transformée de Fourier inverse, c'est-à-dire par une somme de parties paires/impaires, liées aux composantes cartésiennes des coefficients (voir aussi la discussion page 44). Dans le formalisme monogène, les atomes de Fourier ont *deux* composantes alors que les signaux monogènes en ont *trois*. En outre, une ondelette monogène (triplet de Riesz à partir d'une ondelette isotrope) ne correspond pas dans la forme à un atome de Fourier fenêtré, puisque la première est isotrope, alors que le second est orienté. Notons que la transformée de Fourier inverse n'est pas discutée dans la thèse de M. Felsberg [47], ce qui confirme que ce travail a été pour l'instant strictement réservé à l'*analyse* d'image.

Devant ces difficultés à définir une représentation « complète », c'est-à-dire permettant aussi des applications en traitement, une piste a été ouverte par M. Unser en 2009 à travers la proposition d'une transformée en ondelettes monogène avec reconstruction parfaite [144]. Nous allons maintenant analyser en détail ce travail complexe et fondamental pour la suite de notre proposition.

3.3 ONDELETTES MONOGÈNES PEU REDONDANTES

La définition de décompositions monogènes peu redondantes constitue un défi important car l'isotropie, qui constitue l'avantage principal de cet outil, n'est pas réalisable avec les méthodes séparables qui sont à l'origine des ondelettes 2D complexes et quaternioniques.

La première proposition de la littérature définissant un banc de filtres à reconstruction parfaite effectuant une analyse monogène a été publiée lorsque nous commençons ce travail de recherche, il s'agit de l'article de M. Unser *et al.* de 2009 [144]. Le banc de filtre proposé est le fruit d'une série de contributions ayant pour auteurs principaux M. Unser et D. Van De Ville. La méthode est celle des bancs de filtres à reconstruction parfaite liés à une base d'ondelettes, dans l'esprit du travail fondateur de S. Mallat [101] que nous avons présenté à la section 1.2.2. Cependant, l'idée est ici de tirer profit de plusieurs avancées techniques disponibles aujourd'hui dans ce domaine, à savoir :

- Considérer des ondelettes 2D *non-séparables*, peu utilisées mais dont les règles sont bien établies depuis le début des années 1990 [76, 85, 30, 4, 46] ;
- Utiliser éventuellement des filtres à *valeurs complexes*, puisque le formalisme théorique le permet ;
- Considérer des filtres à *réponse impulsionnelle infinie*, à travers la méthode d'implantation des bancs de filtres dans le domaine fréquentiel [111] ;
- Ajouter une certaine *redondance* par un schéma d'échantillonnage moins critique, tout en contrôlant les propriétés des *trames* d'ondelettes associées [145].

Il faut noter que le corpus autour de l'article qui nous intéresse [144] est dense. En effet, la transformée en ondelettes monogène de [144] est un *exemple* d'utilisation d'un formalisme général

de décompositions en ondelettes 2D de type spline, liées à des opérateurs continus possédant des propriétés d'invariance. Nous allons donc tenter ici de synthétiser ce travail, en nous orientant vers la compréhension du banc de filtres monogène. Nous referons le lecteur à [147, 145, 148, 144] pour les détails.

3.3.1 Ondelettes isotropes

Les pyramides orientables, construites par extension des filtres orientables présentés en début de chapitre, offrent une bonne analyse de l'orientation locale à travers un banc de filtres inversible. Cependant, cet outil est indépendant du formalisme des ondelettes et nécessite une optimisation numérique pour la reconstruction. Du côté des bases d'ondelettes bi-orthogonales, une proposition approximativement isotrope est celle de [147]. Il s'agit d'ondelettes non-séparables de type *splines polyharmoniques*, associées à des filtres infinis et un échantillonnage quinconce. Nous proposons de décrire ici la construction du banc de filtres associé.

Splines polyharmoniques

Les splines polyharmoniques sont des extensions multidimensionnelles des B-splines symétriques d'ordre impair. Ces B-splines sont définies dans le domaine fréquentiel par un sinus cardinal élevé à une certaine puissance liée à l'ordre de la spline $\hat{\beta}_{2m-1} = \text{sinc}^{2m}(\omega/2)$, que l'on peut ré-écrire de la façon suivante :

$$\hat{\beta}_{2m-1} = \left(\frac{4 \sin^2(\frac{\omega}{2})}{\omega^2} \right)^m \quad (3.64)$$

Cette formule fait intervenir le spectre du filtre *discret* $[-1 \ 2 \ -1]$ qui approxime le Laplacien, et la transformée de Fourier du Laplacien *continu* $(-\Delta) \xleftrightarrow{\mathcal{F}} \|\omega\|^2$. De ce mélange entre le continu et le discret naît l'idée d'interpolation optimale de données discrètes par des B-splines. Les splines *polyharmoniques* sont définies en remplaçant le numérateur du spectre par une approximation discrète du Laplacien 2D. Les splines polyharmoniques élémentaires sont basées sur le masque classique M (rappelé ci-dessous) mais les auteurs de [147] préfèrent utiliser une version plus isotrope M^{iso} :

$$M = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad M^{iso} = \frac{1}{6} \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix} \quad (3.65)$$

aboutissant à la définition suivante :

$$\hat{\beta}_\gamma = \left(\frac{4(\sin^2 \frac{\omega_1}{2} + \sin^2 \frac{\omega_2}{2}) - \frac{8}{3} \sin^2 \frac{\omega_1}{2} \sin^2 \frac{\omega_2}{2}}{\|\omega\|^2} \right)^{\frac{\gamma}{2}} \quad (3.66)$$

Dans [147], cette fonction est prise comme fonction d'échelle pour la définition d'une analyse multirésolution 2D non-séparable. A partir d'un schéma basé sur une matrice de dilatation en quinconce, les auteurs associent à cette fonction d'échelle une seule fonction d'ondelette approximativement isotrope, qui a la forme d'une somme de 4 atomes de Gabor dans les directions $k\pi/2$ (on a l'égalité lorsque $\gamma \leftarrow \infty$). En outre, lorsque l'échelle devient grossière, il est montré que cette décomposition en ondelettes tend à se comporter comme un opérateur de Laplacien fractionnel $(-\Delta)^{\gamma/2} \xleftrightarrow{\mathcal{F}} \|\omega\|^\gamma$ (Cette propriété sera utilisée par la suite).

L'implantation se fait dans le domaine de Fourier numérique, impliquant un traitement des bords de type « periodisation ». Il faut donc générer les réponses fréquentielles des filtres à la taille de l'image [111]. Notons qu'un préfiltrage est également utilisé pour faire correspondre la séquence d'échantillons de départ à la fonction d'échelle utilisée. Cette transformée est *a priori* trop rigide pour obtenir de bonnes propriétés d'invariance, à cause de l'échantillonnage

critique qui ne donne pas assez de degrés de liberté, et donc une ondelette pas suffisamment isotrope. C'est la raison pour laquelle les auteurs se sont ensuite tournés vers des bancs de filtres légèrement plus redondants. Notons que l'isotropie obtenue est néanmoins acceptable au regard des contraintes de constructions et pourrait être intéressante pour son orthogonalité en compression sans pertes d'image par exemple.

Ajouter une redondance contrôlée

L'évolution décrite dans [145] propose de construire des décompositions de type *pyramide*, déduites à partir d'une base particulière d'ondelettes *orthogonales*. Ce cadre permet de définir une procédure de reconstruction qui prend en compte l'ensemble (redondant) des coefficients, de façon optimale selon un critère quadratique.

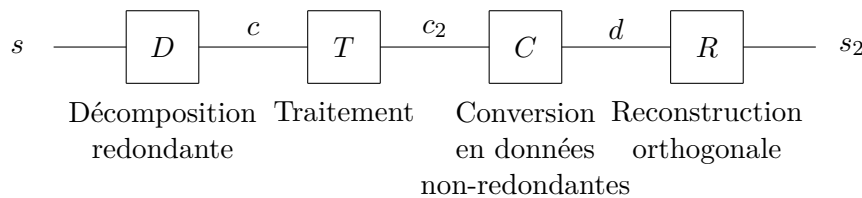
La partie synthèse du banc de filtres est définie comme suit. On considère une base orthogonale *dyadique* dont les trois ondelettes mères sont *égales*, à un décalage spatial près. Elles sont centrées sur les trois « cosets » associés à un échantillonnage dyadique : $[0 \ 1]$, $[1 \ 0]$ et $[1 \ 1]$. Les trois sous-bandes haute-fréquences obtenues avec ces 3 décompositions ont donc le même contenu fréquentiel, mais contiennent séparément les coefficients d'indices pairs et impairs. On peut envisager de retrouver l'ensemble des coefficients non-décimés, par un *entrelacement* de ces sous-bandes, sauf qu'il manquera un échantillon dans chaque bloc dyadique. L'idée de [145] est d'ajouter une quatrième ondelette, toujours identique, positionnée sur le quatrième coset $[0 \ 0]$, afin de compléter la sous-bande entrelacée. On reconstitue donc à chaque échelle une unique sous-bande, qui peut s'obtenir par l'utilisation d'une ondelette unique en supprimant l'étape de sous-échantillonnage du canal passe-haut. En revanche, comme le canal passe-bas est toujours sous-échantillonné, la décomposition est différente d'un schéma non-décimé : c'est une décomposition en pyramide.

Cette procédure très particulière d'extension d'une base orthogonale permet d'une part grâce à la redondance introduite de gagner des propriétés d'invariance sur les coefficients pour améliorer l'analyse, et d'autre part de conserver une méthode de reconstruction efficace par une procédure d'optimisation.

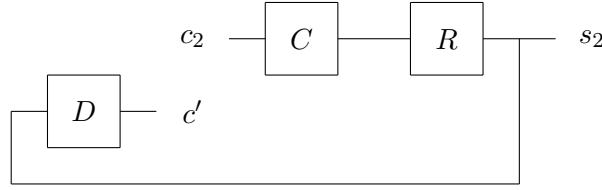
Pour définir la reconstruction, les coefficients redondants vont être « convertis » en un ensemble non-redondants de coefficients, qui seront injectés dans la partie synthèse du banc de filtre *orthogonal*. Nous allons décrire de façon formelle la définition de cette conversion.

On définit d'abord les coefficients d'ondelettes redondants c obtenus par décomposition en pyramide D du signal s ($c = Ds$). On considère ensuite un traitement quelconque (par exemple un seuillage) qui donne les coefficients redondants *modifiés* c_2 . Enfin, on applique une conversion C qui fournit un ensemble non redondant de coefficients $d = Cc_2$, à injecter dans le banc de filtres de reconstruction orthogonal R pour obtenir le signal reconstruit :

$$s_2 = Rd = RCc_2 \quad (3.67)$$



Le but est de définir la conversion C , quel que soit le traitement appliqué, c'est-à-dire, quel que soit c_2 . Une solution triviale est de *sélectionner* les échantillons, en ignorant ceux correspondant à l'ondelette de coset $[0 \ 0]$. Mais il serait plus intéressant que tous les échantillons soient influents sur la reconstruction. L'idée est d'ajouter une contrainte de *cohérence* entre le signal reconstruit s_2 et les coefficients traités c_2 . Si s_2 est cohérent avec c_2 , cela signifie que les coefficients $c' = Ds_2$ obtenus par *re-décomposition* de s_2 doivent ressembler à c_2 .



On peut mesurer cette ressemblance par la distance quadratique :

$$\|c_2 - c'\| = \|c_2 - DRCc_2\| \quad (3.68)$$

Puisque l'on connaît D et R , la minimisation de cette distance fournira C de façon unique.

Les détails de calcul ainsi que les résultats de cette minimisation sont données dans [145]. La méthode résultante de *conversion* C consiste en une combinaison des deux canaux (basses-fréquences, et hautes fréquences) pour reconstruire trois sous-bandes hautes-fréquences non-redondantes. Dans le cadre de cette thèse, nous soulignerons juste que des filtrages supplémentaires sont requis, ainsi que des étapes de ré-échantillonnage, mais que la complexité de l'algorithme reste linéaire, puisqu'il n'y a pas d'itération, et que ces filtres sont définis une fois pour toutes en fonctions des filtres d'analyse et de synthèse de la transformée orthogonale.

Pour exploiter complètement ce schéma, les auteurs proposent d'utiliser le degré de liberté introduit par la redondance pour ajuster la forme de l'ondelette, typiquement pour la rendre plus isotrope. Plus précisément, on peut multiplier la réponse fréquentielle de l'ondelette par un filtre *inversible* (un filtre dont la réponse fréquentielle est non nulle en tout point), et propager cet ajustement dans les formules de reconstruction, sans altérer son optimalité. Les auteurs proposent donc d'améliorer l'isotropie de l'ondelette spline polyharmonique quinquante de [147]. On obtient alors une décomposition encore plus proche d'un comportement Laplacien, comme évoqué plus haut.

Des expérimentation en débruitage par seuillage doux montrent que cette famille de transformées est presque aussi efficace que les ondelettes non-décimées, qui constituent la référence en termes de débruitage, mais souffrant d'une grande redondance. Ici, la redondance est grandement réduite ($\times \frac{4}{3}$), sans avoir les défauts des schémas strictement orthogonaux qui introduisent beaucoup plus d'artéfacts.

Prendre des valeurs complexes

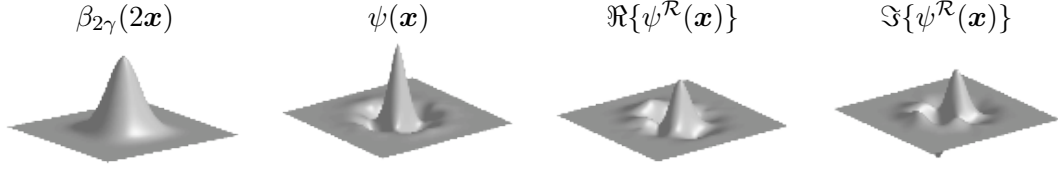
Motivés par l'essor des représentations complexes comme le Dual-Tree [124] (voir section 1.2.3), les auteurs proposent dans [148] une version complexe de leur pyramide isotrope, dont le comportement peut au final s'interpréter comme un gradient.

Suite à l'observation que les ondelettes splines polyharmoniques se comportent comme un Laplacien, les auteurs proposent de considérer un opérateur différentiel plus général combinant la notion de gradient complexe avec celle de Laplacien fractionnel. En imposant les propriétés d'invariance par translation et changement d'échelle, et de co-variance par rotation (une rotation du signal entraîne la multiplication par un facteur constant), on obtient une famille d'opérateurs :

$$L_{\gamma,N} = (-\Delta)^{\frac{\gamma-N}{2}} \left(-j \frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} \right)^N \xleftrightarrow{\mathcal{F}} \hat{L}_{\gamma,N}(\omega) = \|\omega\|^{\gamma-N} (\omega_1 - j\omega_2)^N \quad (3.69)$$

qui combine, à des ordres variables, le gradient complexe et le Laplacien fractionnel (voir équation (3.40), page 81).

Les auteurs construisent à partir de cet opérateur une analyse multirésolution pour laquelle l'ondelette ψ est définie comme l'application de l'adjoint de L sur une fonction « lissante » ϕ . La décomposition en ondelettes associée se comporte donc comme une version multi-échelle de l'opérateur L . On retrouve par exemple la transformée décrite à la section 3.3.1 pour $N = 0$, avec le Laplacien fractionnel $(-\Delta)^{\frac{\gamma}{2}}$.

FIGURE 3.5 – *Spline polyharmonique, ondelette isotrope et sa transformée de Riesz [144]*

Une extension complexe des B-splines polyharmoniques est définie en fonction de ce nouvel opérateur :

$$\hat{\beta}_{\gamma,N}(\boldsymbol{\omega}) = \frac{(4(\sin^2 \frac{\omega_1}{2} + \sin^2 \frac{\omega_2}{2}) - \frac{8}{3} \sin^2 \frac{\omega_1}{2} \sin^2 \frac{\omega_2}{2})^{\gamma/2} \exp[j \arg((\omega_1 - j\omega_2)^N)]}{\|\boldsymbol{\omega}\|^{\gamma-N} (\omega_1 - j\omega_2)^N} \quad (3.70)$$

Notons qu'elle ne diffère pas du cas réel pour $\boldsymbol{\omega} \in [-\pi; \pi]^2$, puisque dans ce cas $\hat{\beta}_{\gamma,N} = \hat{\beta}_{\gamma,0}$.

L'analyse par ondelettes sous-jacente applique intrinsèquement cet opérateur invariant sur des sous-bandes du signal à analyser, cela permet ainsi de donner un cadre théorique lors de l'analyse des coefficients issus de cette décomposition. En particulier, les coefficients sont approximativement invariants par rotation.

3.3.2 Ondelettes monogènes de M. Unser *et al.*

Nous allons maintenant présenter les différents blocs de la décomposition monogène proposée dans [144]. Elle consiste à appliquer en parallèle deux transformées approximativement isotropes réalisées par les pyramides décrites précédemment :

- une décomposition dite « primaire », à coefficients réels ($N = 0$), qui effectue un découpage fréquentiel de base, de façon approximativement isotrope ;
- une décomposition dite « partie Riesz », à coefficients complexes ($N = 1$), qui construit la transformée de Riesz de chaque sous-bande primaire.

pour une redondance totale de $4 \times$.

Cette construction est justifiée par le fait que la transformée de Riesz est un cas particulier de l'opérateur invariant décrit plus haut : $\mathcal{R}^{-1} = -jL_{0,1}$. Plus généralement, on a $\mathcal{R}^{-1}L_{\gamma,0} = jL_{\gamma,1}$. Etant donnée la construction proposée, les deux analyses multirésolution sont alors liées à l'application des opérateurs $L_{\gamma,0}$ et $L_{\gamma,1}$.

On obtient une ondelette « primaire » ψ qui ressemble à un chapeau mexicain, et une ondelette complexe $\psi^{\mathcal{R}}$ définie comme transformée de Riesz (inverse) de ψ :

$$\psi(\mathbf{x}) = (-\Delta)^{\frac{\gamma}{2}} \beta_{2\gamma}(2\mathbf{x}) \quad \psi^{\mathcal{R}} \xleftrightarrow{\mathcal{F}} \frac{j\omega_1 + \omega_2}{\|\boldsymbol{\omega}\|} \hat{\psi}(\boldsymbol{\omega}) \quad (3.71)$$

où le Laplacien fractionnel est défini par l'équation (3.40). Ces fonctions sont illustrées à la figure 3.5.

Les versions dilatées et translatées sont définies par $\psi_{i,\mathbf{k}}(\mathbf{x}) = 2^i \psi(2^i \mathbf{x} - \mathbf{k}/2)$ (*idem* pour $\psi^{\mathcal{R}}$). Les coefficients d'ondelettes vont alors s'interpréter à travers les relations suivantes :

$$c_{i,\mathbf{k}} = \langle s, \psi_{i,\mathbf{k}} \rangle = (\psi_i * s)(2^{-(i+1)} \mathbf{k}) \quad (3.72)$$

$$d_{i,\mathbf{k}} = \langle s, \psi_{i,\mathbf{k}}^{\mathcal{R}} \rangle = \mathcal{R}(\psi_i * s)(2^{-(i+1)} \mathbf{k}) \quad (3.73)$$

Cela signifie que les coefficients constituent bien un signal *monogène* à chaque échelle. $c_{i,\mathbf{k}}$ et $d_{i,\mathbf{k}}$ sont donc fusionnés dans un 3-vecteur et convertis en coordonnées polaires conformément au formalisme de Felsberg (équations (3.45)).

Un code Java est disponible sur le site des auteurs (<http://bigwww.epfl.ch/demo/monogenicj/>) pour l'implantation de la transformée, et nous donnons un exemple de décomposition sur la figure 3.6. La méthode d'affichage est la même que pour les figures 3.3 et 3.4,

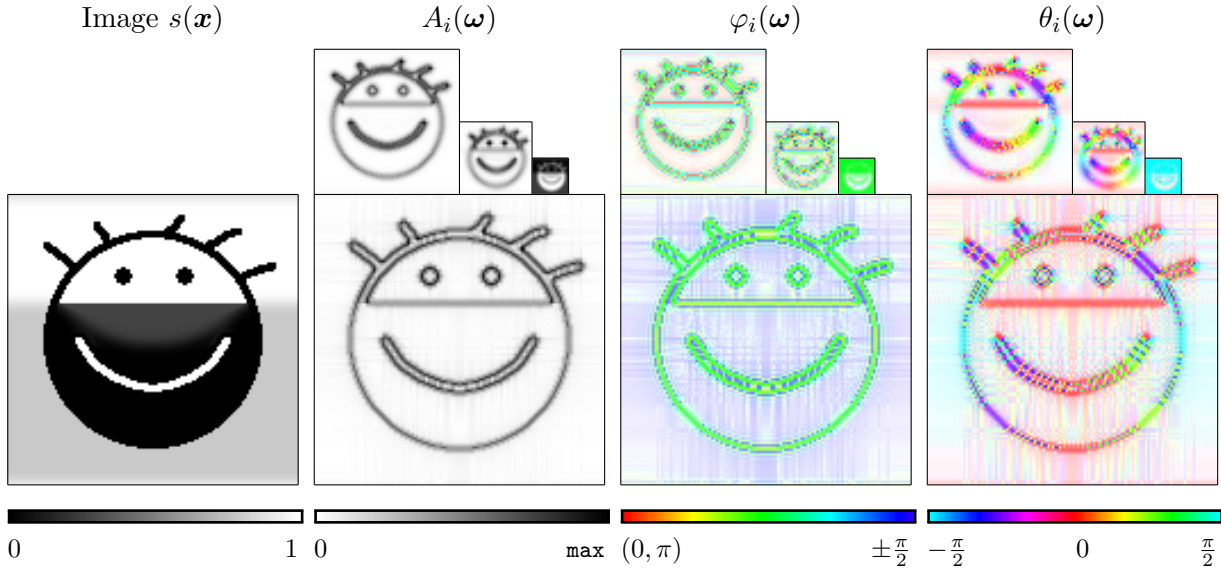


FIGURE 3.6 – Transformée en ondelettes monogène (MWT) [144] sur une image synthétique ($\gamma = 3$, $i \in \{-1, -2, -3\}$).

c'est-à-dire que la phase et l'orientation sont exprimées par la *teinte* (le « H » de « HSV »), alors que l'amplitude pondère la *saturation* (« S »). On a donc des pixels pâles - sans teinte franche - lorsque les coefficients sont trop faibles pour avoir une phase et une orientation pertinentes. La progression d'une échelle à l'autre est dyadique, et on juxtapose les différentes sous-bandes, en partant de la première échelle, qui est de la même taille que l'image. Le choix du paramètre γ , qui doit être supérieur à 1, est un problème ouvert. Pour des valeurs proches de 1, l'isotropie est mieux approximée, mais pour de plus grandes valeurs, les ondelettes sont plus régulières ($\lceil \gamma \rceil$ moments nuls). On constate expérimentalement qu'au delà de $\gamma \approx 10$, l'algorithme devient numériquement instable.

L'interprétation des coefficients est la même que pour les ondelettes monogènes continues, mais est moins aisée *visuellement* à cause du sous-échantillonnage. Un exemple d'utilisation en démodulation AM/FM est présenté dans [144], où l'estimation de l'orientation est très satisfaisante. Il est également proposé d'améliorer l'estimation de l'orientation dans un formalisme de tenseur de structure, qui consiste à appliquer un lissage gaussien sur $T_{11} = (\Re d)^2$, $T_{22} = (\Im d)^2$ et $T_{12} = \Re d \Im d$, puis à extraire l'orientation par l'équation (3.9), plutôt que directement de $\arg\{d\}$. L'orientation est beaucoup plus lisse et semble correspondre mieux à notre perception [144]. Une application de cette transformée en vision stéréoscopique a également été proposée dans [97].

L'algorithme nécessite de calculer les filtres avant la décomposition en fonction de la taille de l'image, puisque le calcul est fait dans le domaine de Fourier, par FFT. La complexité de l'algorithme reste tout de même linéaire par rapport à la dimension des données.

Les contraintes de définition liées à la structure d'analyse multi-résolution et de reconstruction parfaite sont à l'origine d'une invariance par rotation seulement approximative, et à une quasi-isotropie des ondelettes. De plus, l'ensemble des degrés de libertés ont été utilisés pour que l'ondelette d'analyse soit la plus isotrope possible. En contrepartie, les filtres de reconstruction sont loin d'être isotropes, ce qui peut affaiblir la qualité de la transformée dans le cadre de problématiques en débruitage ou en codage par exemple. Ce genre de problème n'est pas présent dans les pyramides orientables par exemples, pour lesquelles la reconstruction n'est pas parfaite et les sous-bandes ne sont pas nécessairement orthogonales.

L'intérêt principal de cette proposition est l'utilisation du formalisme des bases d'ondelettes, qui permet d'appliquer sur des données discrètes des opérateurs différentiels continus à plusieurs échelles. On a de plus la reconstruction parfaite dans chacune des deux transformées et l'orthogonalité entre les échelles (« semi-orthogonalité », qui n'assure pas l'orthogonalité entre les coefficients d'une même sous-bande). L'idée d'utiliser deux transformées en parallèle est analogue

à celle du Dual-Tree, à cela près que la reconstruction n'est pas explicitement définie comme l'addition des reconstructions indépendantes. En fait, outre les reconstructions indépendantes des deux bancs de filtres qui sont bien établies, la question d'une reconstruction monogène unifiée n'est pas vraiment traitée, et d'ailleurs notons que seules des applications d'*analyse* d'image sont proposées. Notons que l'article plus récent [69] propose une définition alternative d'ondelettes monogènes, mais ne traite pas non plus cette problématique de reconstruction.

Par rapport aux différentes ondelettes géométriques existantes (curvelets, bandlets, ridgelets *etc.*), la transformée monogène apporte une notion de phase avec une interprétation physique, ainsi que des propriétés d'invariance obtenues pour une redondance minimale. Elle se positionne dans la littérature comme une représentation peu redondante à l'analyse enrichie, conformément au formalisme monogène qui constitue à ce jour la généralisation la plus satisfaisante du signal analytique. D'une manière plus générale, l'utilisation de techniques avancées en termes de définitions de transformées en ondelettes, la place clairement au dessus de la transformée séparable utilisée le plus souvent, car elle ne contient qu'une seule sous-bande par échelle, qui ne privilégie donc pas les contours horizontaux, verticaux et diagonaux.

La décomposition monogène et ses propriétés sont valables pour des signaux 2D réels. En revanche, l'utilisation de ce banc de filtres pour analyser une image couleur n'est pas définie. Ceci est lié en partie au fait que la notion de phase par exemple n'est pas définie pour un signal vectoriel. Plus généralement, les représentations monogènes ne suffisent pas pour appréhender les images couleur. C'est pourquoi nous allons maintenant proposer une extension couleur non-triviale de cette représentation.

3.4 EXTENSION COULEUR PAR LES ALGÈBRES GÉOMÉTRIQUES

Nous proposons dans cette section une transformée en ondelettes monogène, spécialement définie pour les images couleur. La difficulté réside dans la contrainte que nous imposons de ne pas définir un schéma *marginal*, c'est-à-dire consistant à appliquer un outil scalaire indépendamment sur chaque composante couleur [11], sans réelle unification en termes « d'entité couleur ». En effet, cette méthode trop simple aboutit à des données présentant plusieurs amplitudes, phases et orientations pour décrire une structure couleur, ce qui n'est pas satisfaisant en termes de description géométrique. Le but recherché est une décomposition de l'image en coefficients dont les valeurs peuvent s'interpréter *physiquement*. Pour cela, une extension rigoureuse du formalisme de M. Felsberg doit être définie pour les images couleur. Nous n'avons trouvé dans la littérature qu'une seule contribution à ce sujet [38], proposant un signal monogène couleur par une généralisation des équations de Cauchy-Riemann. La première étape dans notre travail est donc de définir une décomposition en ondelettes couleur correspondant à ce modèle.

D'une manière plus générale, il n'existe pas à notre connaissance d'outil *signal* dédié aux images couleur, si l'on excepte les tentatives quaternioniques décrites au chapitre 2, et quelques représentations en ondelettes (multi-ondelettes [94], lifting vectoriel [8]) n'offrant pas d'interprétation physique des données car plutôt dédiées à la compression. Nous pensons que le formalisme monogène, qui résout un problème de dimension spatiale des données en dépassant les limites algébriques des nombres complexes, est un cadre propice à une extension pour des signaux à valeurs vectorielles - qui posent également un problème de dimension.

3.4.1 Le signal monogène couleur de G. Demarcq

La définition de signal monogène couleur proposée dans [37, 38] prend la suite de ce que nous avons présenté aux sections 3.2.1 et 3.2.2, sur la construction du signal monogène par la généralisation des équations de Cauchy-Riemann. Comme dans le travail de M. Felsberg, les algèbres géométriques que nous avons introduites à la section 3.2.2 sont utilisées, nous allons donc les employer à nouveau ici.

Dans un premier temps, nous devons introduire l'équation de Dirac, qui généralise le système de Cauchy-Riemann. L'opérateur de Dirac, appliqué à une fonction $s : \mathbb{R}^n \mapsto \mathcal{G}_n$ est défini dans l'algèbre géométrique $\mathcal{G}_n$ par :

$$D_n s = \sum_{k=1}^n \mathbf{e}_k \frac{\partial s}{\partial x_k} \quad (3.74)$$

Cet opérateur généralise le gradient complexe que nous avons introduit à l'équation (3.29). L'équation de Dirac pour s est définie par :

$$D_n s = 0 \quad (3.75)$$

Dans le cas $n = 2$, et en posant $f(x_1, x_2) = \mathbf{e}_1 f_R(x_1 + \mathbf{j}x_2) + \mathbf{e}_2 f_I(x_1 + \mathbf{j}x_2)$, l'équation de Dirac est équivalente au système de Cauchy-Riemann (équation (3.29)). La fonction f est alors appelée *holomorphe*. Dans le cas $n = 3$, on retrouve les fonctions *monogènes* étudiées précédemment, qui ont servi à construire le signal monogène de M. Felsberg. Plus généralement, pour $n > 2$, les fonctions de $\mathbb{R}^n$ à valeurs vectorielles dans $\mathcal{G}_n$ sont dites *monogènes* si elles vérifient l'équation de Dirac, ce qui est un résultat connu en analyse de Clifford.

La proposition de [38] pour la définition d'un signal monogène couleur est de considérer les fonctions monogènes pour $n = 5$. Une image couleur est habituellement définie par trois fonctions de deux variables spatiales, $[s^R(x_1, x_2) \ s^G(x_1, x_2) \ s^B(x_1, x_2)]$, dont on peut considérer qu'il s'agisse des canaux rouge, vert et bleu du système colorimétrique RGB, mais que nous traiterons dans notre travail comme un vecteur de dimension 3. Afin de l'intégrer dans l'algèbre géométrique, une image couleur sera codée comme un vecteur de $\mathcal{G}_5$:

$$s(x_1, x_2) = \mathbf{e}_3 s^R(x_1, x_2) + \mathbf{e}_4 s^G(x_1, x_2) + \mathbf{e}_5 s^B(x_1, x_2) \quad (3.76)$$

La suite de la construction est analogue à la méthode de Felsberg. Il s'agit de construire une fonction monogène f satisfaisant certaines conditions aux bords :

$$f : \mathbb{R}^5 \mapsto \mathcal{G}_5 \quad (3.77)$$

$$D_5 f = 0 \quad \text{pour } x_1, x_2 \in \mathbb{R} \quad x_3, x_4, x_5 \in \mathbb{R}^+ \quad (3.78)$$

$$f(x_1, x_2, 0, 0, 0) = \mathbf{e}_1 s_1(x_1, x_2) + \mathbf{e}_2 s_2(x_1, x_2) + \mathbf{e}_3 s^R(x_1, x_2) + \mathbf{e}_4 s^G(x_1, x_2) + \mathbf{e}_5 s^B(x_1, x_2) \quad (3.79)$$

où s_1 et s_2 sont pour l'instant inconnues. Comme nous l'avons vu pour $n = 2$, cette formulation est équivalente à trouver le *potentiel harmonique* $p : \mathbb{R}^5 \mapsto \mathbb{R}$ (vérifiant donc $\Delta p = 0$) dont le gradient $D_5 p$ est égal à f . On obtient donc le système d'équations suivant :

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \frac{\partial^2 p}{\partial x_2^2} + \frac{\partial^2 p}{\partial x_3^2} + \frac{\partial^2 p}{\partial x_4^2} + \frac{\partial^2 p}{\partial x_5^2} = 0 \quad \text{pour } x_1, x_2 \in \mathbb{R} \quad x_3, x_4, x_5 \in \mathbb{R}^+ \quad (3.80)$$

$$\mathbf{e}_3 \frac{\partial p}{\partial x_3} + \mathbf{e}_4 \frac{\partial p}{\partial x_4} + \mathbf{e}_5 \frac{\partial p}{\partial x_5} = s(x_1, x_2) \quad \text{pour } x_3 = x_4 = x_5 = 0 \quad (3.81)$$

Il vient que ce système peut être décomposé en trois sous-systèmes, permettant d'obtenir la solution monogène 5D $f_c(x_1, x_2, x_3, x_4, x_5)$. La seconde étape pour définir le signal monogène couleur est de restreindre cette fonction au sous-domaine $x_3 = x_4 = x_5 = 0$, c'est-à-dire :

$$s_M(x_1, x_2) = f(x_1, x_2, 0, 0, 0) \quad (3.82)$$

L'expression de la solution fait apparaître une transformée de Riesz entre la « partie réelle », représentée par les axes $\mathbf{e}_3$, $\mathbf{e}_4$ et $\mathbf{e}_5$, et la partie imaginaire ($\mathbf{e}_1$, $\mathbf{e}_2$), pour des valeurs fixées de x_3 , x_4 , et x_5 . Plus précisément, la définition explicite du signal monogène couleur s_M est la suivante :

$$s_M(\mathbf{x}) = \mathbf{e}_1 s_1(\mathbf{x}) + \mathbf{e}_2 s_2(\mathbf{x}) + \mathbf{e}_3 s^R(\mathbf{x}) + \mathbf{e}_4 s^G(\mathbf{x}) + \mathbf{e}_5 s^B(\mathbf{x}) \quad \text{avec } s_1 + \mathbf{j} s_2 = \mathcal{R}\{s^R + s^G + s^B\} \quad (3.83)$$

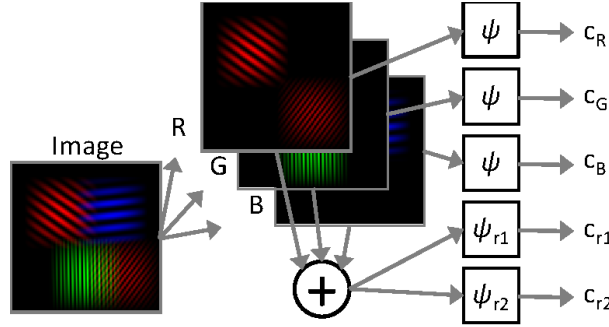


FIGURE 3.7 – Schéma de décomposition en ondelettes monogènes couleur. Chaque canal couleur est décomposé par la transformée primaire symbolisée par un « bloc ψ », et la somme « $R + G + B$ » est analysée par la « partie Riesz » (blocs ψ_{r1} et ψ_{r2}).

Pour compléter la représentation, une notion de représentation polaire doit être définie. Pour cela, une donnée de « phase » est proposée, et s'obtient selon un paramètre constant V caractérisant à la fois une direction spatiale et une direction dans l'espace couleur, pour toute l'image. Cette phase est construite par le *produit géométrique* suivant :

$$\begin{aligned} s_A V &= (s^R \mathbf{e}_1 + s^G \mathbf{e}_2 + s^B \mathbf{e}_3 + s^C \mathbf{e}_4 + s^D \mathbf{e}_5) (u \mathbf{e}_1 + v \mathbf{e}_2 + a \mathbf{e}_3 + b \mathbf{e}_4 + c \mathbf{e}_5) \\ &= p_0 + p_{12} \mathbf{e}_{12} + p_{13} \mathbf{e}_{13} + p_{14} \mathbf{e}_{14} + p_{15} \mathbf{e}_{15} + p_{23} \mathbf{e}_{23} + p_{24} \mathbf{e}_{24} + p_{25} \mathbf{e}_{25} + p_{34} \mathbf{e}_{34} + p_{35} \mathbf{e}_{35} + p_{45} \mathbf{e}_{45} \end{aligned} \quad (3.84)$$

Cette opération projette le signal monogène sur l'axe 5D déterminé par V . Le produit $s_A V$ est un « spineur » - l'élément des algèbres géométriques qui caractérise les rotations - dont l'angle de rotation définit la « phase » proposée $\varphi_V = \arg\{p_0 + \mathbf{j} \sqrt{\sum p_{mn}^2}\}$. φ_V est identifié comme une phase couleur liée au paramètre V . Cet angle de phase représente une distance angulaire entre le signal monogène et la structure couleur fixée par V . Des applications sont proposées par les auteurs en détection de contours et flot optique. Notons également une application en stéréoscopie dans [95]. Plus de détails sur ce développement pourront être trouvés dans la thèse de G. Demarcq [37].

Nous proposons maintenant une transformée en ondelettes monogène couleur construite à partir de ce formalisme, afin de définir la première représentation monogène multi-échelle des images couleur, permettant une reconstruction parfaite et à redondance raisonnable.

3.4.2 Décomposition en ondelettes

Le but est de réaliser une décomposition en ondelettes telle que les sous-bandes forment des signaux monogènes couleur. Comme pour les représentations analytiques étudiées jusqu'ici, la définition du signal monogène couleur consiste à juxtaposer les trois composantes du signal d'origine à une version déphasée, ici construite par transformée de Riesz sur la somme des trois canaux. La réalisation la plus évidente d'une décomposition multi-échelles va être d'appliquer une décomposition isotrope sur $[s^R \ s^G \ s^B]$ d'une part, et une décomposition « déphasée » sur $[s^R + s^G + s^B]$ d'autre part. En effet, si l'on considère la partie « primaire » et la « partie Riesz » de la décomposition monogène décrite à la section 3.3.2, symbolisées formellement par les opérateurs Ψ et $\Psi^{\mathcal{R}}$, on peut construire la décomposition en ondelettes monogènes couleur suivante :

$$\begin{bmatrix} s^R & s^G & s^B \end{bmatrix} \begin{bmatrix} \Psi & 0 & 0 & \Psi^{\mathcal{R}} \\ 0 & \Psi & 0 & \Psi^{\mathcal{R}} \\ 0 & 0 & \Psi & \Psi^{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} c_{i,k}^R & c_{i,k}^G & c_{i,k}^B & c_{i,k}^{\mathcal{R}} \end{bmatrix} \quad (3.85)$$

Les transformées « en niveaux de gris » présentées plus haut peuvent donc directement être utilisées, pour définir une transformée en ondelettes monogènes couleur. Le schéma de décomposition est illustré à la figure 3.7. On peut vérifier avec les équations (3.72) et (3.73) que les

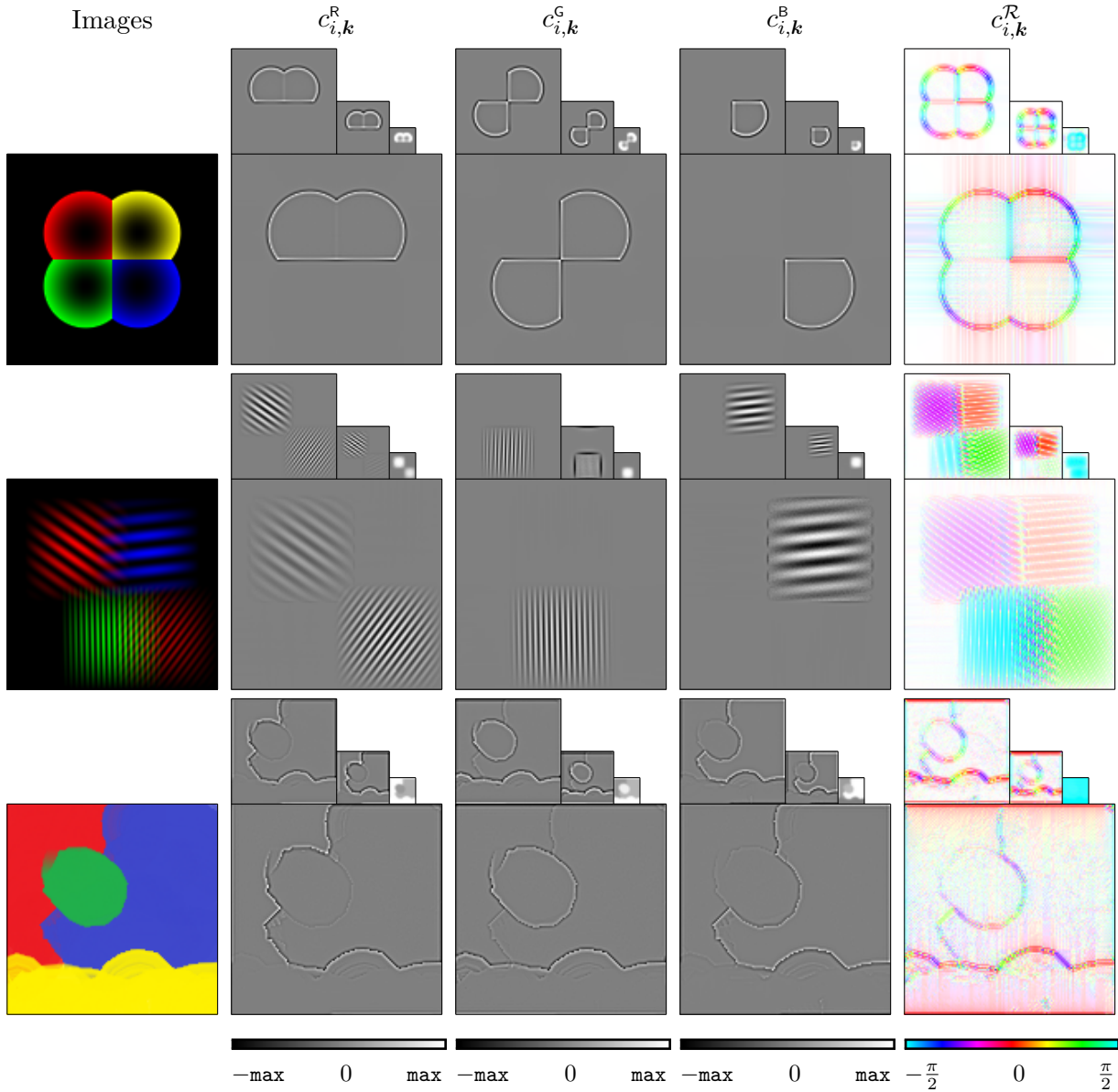


FIGURE 3.8 – *MWT couleur d'images synthétiques.* La partie Riesz est représentée dans le système couleur HSV avec son module dans l'intensité et l'argument (modulo π) dans la teinte. La valeur de γ est fixée à 3 expérimentalement.

coefficients forment effectivement un signal monogène couleur, c'est-à-dire :

$$c_{i,k}^c = \langle s^c, \psi_{i,k} \rangle = (\psi_i * s^c)(2^{-(i+1)}\mathbf{k}) \quad \text{avec } c \in \{R, G, B\} \quad (3.86)$$

$$c_{i,k}^R = \langle s, \psi_{i,k}^R \rangle = \mathcal{R}\{(\psi_i * s^R) + (\psi_i * s^G) + (\psi_i * s^B)\}(2^{-(i+1)}\mathbf{k}) \quad (3.87)$$

On obtient donc une analyse multi-échelle monogène couleur sous forme d'une pyramide dont les coefficients sont des vecteurs de dimension 5, et dont des exemples sont illustrés à la figure 3.8. Il s'agit d'une nouvelle transformée spécialement définie pour les images couleur. Elle possède des propriétés intéressantes d'invariance, en particulier par rapport aux rotations, grâce à l'utilisation d'ondelettes approximativement isotropes tout en gardant une redondance raisonnable de $20:9 \approx 2.2$. L'utilisation d'un banc de filtre non-séparable est particulièrement appréciable car les structures ne sont pas classifiées arbitrairement (contrairement aux schémas séparables « ligne/colonne »). Grâce à une nouvelle notion de signal analytique construite pour les images couleur, les coefficients de cette représentation sont obtenus de façon non-marginale. En effet, l'intensité $R+G+B$ est considérée en plus des canaux RGB, ce qui implique en quelque

sorte deux espaces couleur différents. Les coefficients comprennent donc plusieurs composantes. Les trois premières sont associées à des informations d'énergie couleur, que l'on peut définir par :

$$(c_{i,k}^R)^2 \quad (c_{i,k}^G)^2 \quad (c_{i,k}^B)^2 \quad (3.88)$$

Les deux dernières composantes sont associées à une notion de forme géométrique locale. Elle donnent accès à une norme et une orientation définies par :

$$\mathcal{N} = |c_{i,k}^{\mathcal{R}}| \quad \theta = \arg\{c_{i,k}^{\mathcal{R}}\} \quad (3.89)$$

Cette analyse géométrique est obtenue de façon unique dans l'esprit du formalisme monogène, mais cette fois à partir d'une image couleur, et complète bien l'absence de classification des structures due à l'utilisation d'une sous-bande unique par échelle. Le fait de n'avoir qu'une seule orientation pour les trois canaux couleur n'aurait pas pu être obtenu par un schéma marginal, qui aurait produit trois orientations différentes pour une même structure.

Observons à la figure 3.8 les trois premières composantes des coefficients, c^R , c^G et c^B , pour lesquelles le blanc (resp. noir) représente les valeurs fortement positives (resp. negative). On observe une classification de l'information couleur selon les canaux. Par exemple, le contour de l'élément bleu de la première image n'est présent que dans la troisième composante. L'aspect multi-échelle est visible sur la décomposition de la deuxième image. Ainsi, l'oscillation verte, qui est plus rapide que les autres, produit une plus forte énergie dans les premières échelles de la deuxième composante. On voit que les contours ronds réguliers sont représentés par des coefficients dont l'amplitude ne varie pas selon l'orientation, ce qui confirme l'invariance par rotation de la transformée, à savoir qu'aucune direction particulière n'est privilégiée. Finalement, la partie primaire de la transformée donne une représentation multi-échelle compacte de l'énergie couleur de l'image, indépendante de la structure géométrique.

Observons maintenant la partie Riesz $c_{i,k}^{\mathcal{R}}$ de la transformée. On retrouve l'analyse de l'orientation que l'on avait avec les ondelettes monogènes en niveaux de gris. L'amplitude $|c_{i,k}^{\mathcal{R}}|$ permet de mesurer la pertinence de cette information d'orientation, nous l'avons donc encodée dans la luminance de l'illustration. Cela permet de détecter rapidement les structures 1D de l'image couleur - qui correspondent en général aux contours. On obtient donc une donnée d'orientation *unique* en chaque point et chaque échelle, conformément à notre perception, et contrairement à ce qu'aurait produit un schéma marginal. Comme c'était déjà le cas en niveaux de gris, les valeurs obtenues pour l'orientation sont cohérentes avec les orientations perçues des structures. Notons que les zones d'orientation constante dans la deuxième image produisent une orientation également constante dans la partie Riesz (même « teinte » dans l'illustration). Enfin, les contours *isoluminants* - correspondant à une variation de couleur sans variation d'intensité - ne sont pas détectés par la transformée de Riesz, qui est appliquée, justement, sur l'intensité. C'est le cas des contours rouge/vert et vert/bleu des première et troisième images.

3.4.3 La reconstruction

Les traitements comme le débruitage nécessitent la partie synthèse du banc de filtres pour reconstruire l'image traitée. Dans le cas des représentations redondantes, il y a en général plusieurs façon de reconstruire parfaitement. Cependant, lorsque les coefficients sont modifiés, la méthode de reconstruction doit être telle que les modifications obtenues soient cohérentes avec le traitement dans le domaine des ondelettes. Ce problème existe déjà dans le cas scalaire, puisque les pyramides que nous utilisons sont redondantes d'un facteur $\frac{4}{3} \times$. L'algorithme de reconstruction associé est bien défini par les auteurs dans [145] et consiste à utiliser cette redondance dans une procédure optimale. Dans notre cas, une autre sorte de redondance - celle du formalisme monogène - nous suggère les trois méthodes globales suivantes :

- Reconstruire l'image à partir des seules parties « primaires » c^R , c^G et c^B , et donc réduire la partie Riesz à une seule fonction d'*analyse*.

Échelle s	-1	-2	-3	-4
Écart-type σ_s	1.339	1.502	1.512	1.560

TABLE 3.2 – *Ecart-types des sous-bandes de décomposition d'un bruit gaussien de variance 1 (avec $\gamma = 3$).*

- Reconstruire l'intensité de l'image $R + G + B$ à partir de la partie Riesz seulement.
- Combiner les deux reconstructions par une méthode spécifique à l'application.

Dans tous les cas, la reconstruction est parfaite. Nous proposons maintenant d'étudier cette transformée d'un point de vue expérimental.

3.5 MISE À L'ÉPREUVE EN DÉBRUITAGE COULEUR

Nous proposons dans cette section de mettre à l'épreuve la représentation multi-échelle monogène couleur que nous avons proposée. Même si le formalisme monogène est *a priori* plutôt destiné à l'*analyse* des images, nous avons choisi une application de *traitement* - le débruitage - c'est-à-dire que la problématique de *reconstruction* sera à considérer. Les méthodes de débruitage d'image par ondelettes sont connues pour être efficaces contre différents bruits tout en offrant des algorithmes très rapides - contrairement à des méthodes itératives comme les approches variationnelles qui peuvent donner de meilleurs résultats mais pour un coût algorithmique important dans certaines applications. Les méthodes classiques consistent à effectuer un seuillage des coefficients d'ondelette. Cependant dans notre cas, l'information portée par les coefficients est plus riche que les ondelettes orthogonales « habituelles ». Notre décomposition non-orthogonale doit donc faire l'objet d'une étude sur la modélisation des coefficients.

De plus, nous avons vu que la transformée proposée est composée de deux sortes de données :

- La partie « primaire » : un ensemble de coefficients associés à une information colorimétrique, formant trois pyramides liées aux canaux couleur.
- La partie « Riesz » : Une mesure géométrique faite d'une amplitude et d'un angle en tout point, portant une information de forme/structure locale.

Pour construire l'algorithme de débruitage, une sélection des coefficients doit être définie.

Pour cela, on caractérisera expérimentalement les coefficients liés au bruit dans les différentes sous-bandes pour identifier leurs distributions et leur corrélation. Ensuite, les informations colorimétriques et géométriques seront fusionnées dans une procédure de seuillage unique sur l'amplitude des coefficients, réalisant un débruitage par ondelettes monogènes couleur.

3.5.1 Étude des coefficients liés au bruit

Les méthodes classiques de débruitage par ondelettes reposent sur l'hypothèse que la distribution des coefficients liés au bruit peut être modélisée efficacement par une loi gaussienne centrée. Cela implique en général un seuil constant sur l'ensemble des coefficients [44]. Ici, nous avons une transformée non-orthogonale, nécessitant d'étudier le comportement d'un bruit à travers la décomposition monogène couleur.

Coefficients primaires

La figure 3.9 montre que la décomposition d'un bruit gaussien centré de variance $\sigma^2 = 1$ produit des sous-bandes au comportement également gaussien - centrées, mais avec des variances différentes. Ceci indique que la transformée n'est pas parfaitement normalisée, et donc ce facteur d'amplification doit être injecté dans la procédure de seuillage. Des valeurs expérimentales des écart-types sont données dans la table 3.2. Les distributions continues théoriques ont été tracées sur les histogrammes (côté gauche de la figure 3.9) et montrent que le modèle gaussien

$$f(x) = \frac{1}{\sigma_s \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_s^2}} \quad (3.90)$$

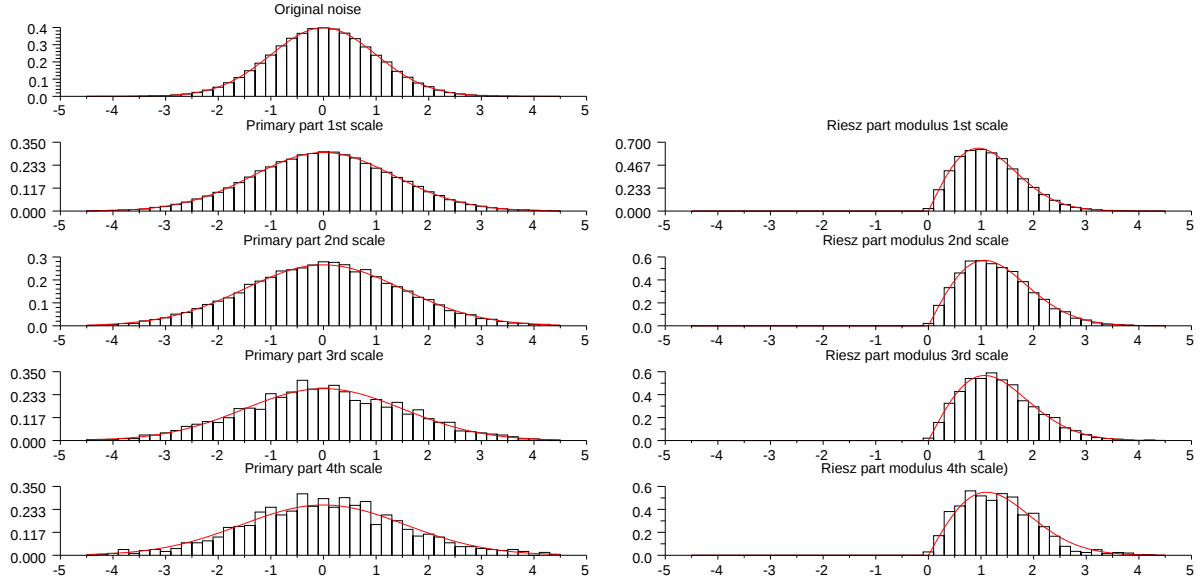


FIGURE 3.9 – Histogrammes (barres) des sous-bandes primaires et du module des sous-bandes de la partie Riesz; modèles de densités de probabilités (lignes).

où σ_s est l'estimation de variance mesurée sur les coefficients, est bien adapté. La décomposition est appliquée à travers un ensemble de filtres passe-bande, avec une corrélation très limitée entre les fonctions de base, ce qui explique que les sous-bandes soient gaussiennes et apparemment indépendantes.

Les valeurs de σ_s pourraient également être calculées analytiquement à partir de la définition des filtres. En effet, le filtrage linéaire y d'un signal aléatoire stationnaire x par un filtre H peut être étudié par densités spectrales de puissances $\Psi_x = |\mathcal{F}[x]|^2$ et autocorrélations $R_x(\tau) = \mathcal{F}^{-1}\Psi_x$. On a en particulier $\Psi_y = \Psi_x |H|^2$. La variance de sortie σ_y^2 est égale à $R_y(0)$, ce qui se réduit à $\sigma_y^2 = \frac{\sigma_x^2}{4\pi^2} \iint_{(0,0)}^{(2\pi,2\pi)} |H(\omega)|^2 d\omega$.

Par exemple la première échelle du banc de filtres est directement liée au filtre passe-haut :

$$\sigma_1^2 = \frac{\sigma^2}{4\pi^2} \iint_{(0,0)}^{(2\pi,2\pi)} \left| \frac{\left(4(\sin^2 \frac{\omega_1}{2} + \sin^2 \frac{\omega_2}{2}) - \frac{8}{3} \sin^2 \frac{\omega_1}{2} \sin^2 \frac{\omega_2}{2}\right)^\gamma}{2\|\omega\|^\gamma} \right|^2 d\omega \quad (3.91)$$

Les autres coefficients sont liés à des filtres équivalents à chaque étape du banc de filtres.

Des données plus réalistes sont introduites avec l'image **peppers** dégradée par un bruit blanc gaussien additif (Figure 3.10a). La figure 3.10b illustre l'histogramme de c^R (partie primaire, canal rouge) à la première échelle. Comme dans la méthode classique, nous supposons que la première échelle de la décomposition en ondelettes d'une image *naturelle* contient en très grande majorité du bruit. On retrouve à nouveau expérimentalement le modèle gaussien.

Pour confirmer le modèle gaussien, nous avons calculé un test de Kolmogorov-Smirnov qui compare les coefficients avec une vraie distribution gaussienne. Le test est positif pour l'image **peppers** avec une « p -value » de 0.9.

Voyons maintenant les coefficients de la partie Riesz.

Coefficients de la partie Riesz

Rappelons que l'information structurale est portée par le module et l'argument d'un nombre complexe. La manipulation et le seuillage de données circulaires comme un angle n'est pas aisée. Dans ce travail exploratoire, et comme une première étape naturelle, nous nous concentrons sur le module. De plus, cette donnée est liée à une notion d'*amplitude* liée aux structures géométriques, et pour lesquelles un seuillage peut paraître approprié. Seuiller un angle est moins intuitif.

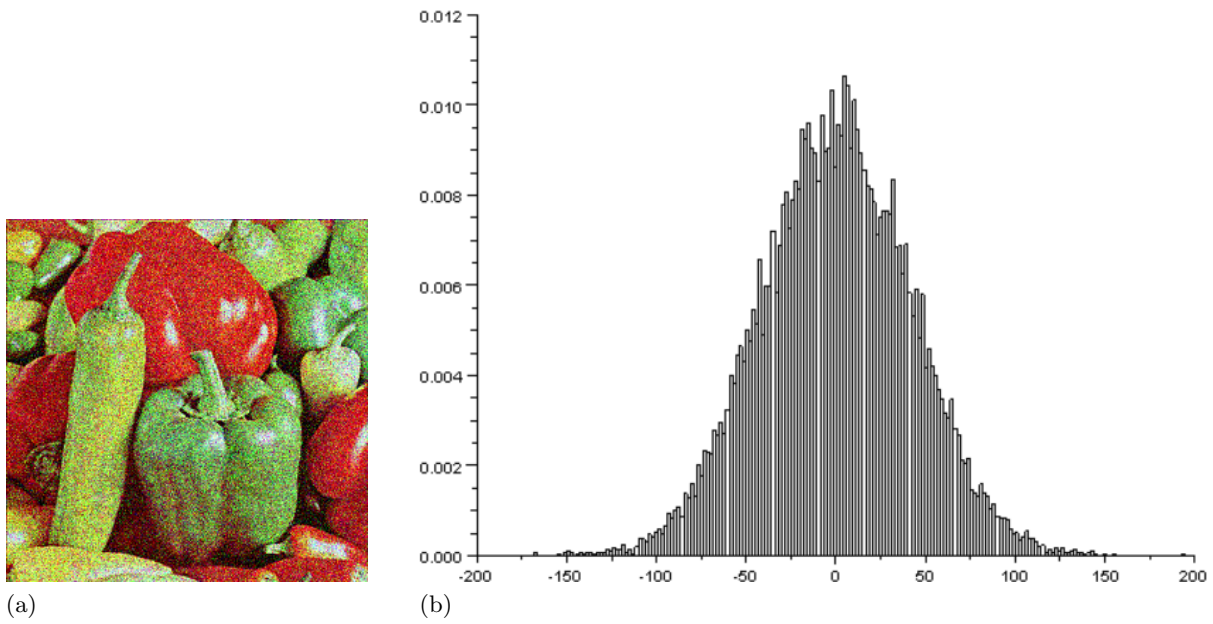


FIGURE 3.10 – Image **peppers** dégradée par un bruit blanc gaussien additif ($SNR=83$ dB) et distribution des coefficients de la première échelle du canal rouge c^R .

Les coefficients de la partie Riesz peuvent eux aussi être étudiés en tant que sorties de filtrages linéaires passe-bande, c'est pourquoi nous proposons à nouveau de faire une hypothèse gaussienne pour leur distribution. Les parties réelles et imaginaires des sous-bandes suivent donc une loi gaussienne centrée et de variance σ_s^2 .

L'étude du module implique de caractériser la loi d'une variable aléatoire $Z = (X^2 + Y^2)^{\frac{1}{2}}$ où (X, Y) sont deux variables aléatoires gaussiennes indépendantes de moyenne nulle et d'écart-type σ_s . Dans ce cas, il est bien connu que Z suit une loi de Rayleigh $f_Z(m) = \frac{m}{\sigma_s^2} e^{-m^2/2\sigma_s^2} \mathbf{1}_{m>0}$, où $\mathbf{1}_{m>0}$ est l'échelon de Heaviside. Les moments de Z sont donc $E(Z) = \sigma_s \sqrt{\pi/2}$ et $V(Z) = \sigma_s^2(2 - \frac{\pi}{2})$. Voici la procédure que nous employons pour décrire la distribution des coefficients de la partie Riesz, illustrée à la figure 3.9 :

- Générer une image-test couleur faite de bruit gaussien seulement ;
- La décomposer à travers la transformée en ondelettes monogènes couleur ;
- Calculer les histogrammes des sous-bandes primaires et du module des sous-bandes de la partie Riesz ;
- Mesurer les écarts-types des sous-bandes primaires pour avoir une valeur expérimentale de σ_s ;
- Calculer et représenter les distributions f et f_Z avec les σ_s mesurés ;

Comme pour la partie primaire, la modélisation du module de la partie Riesz (côté droit de la figure 3.9) correspond bien aux histogrammes mesurés. Notons que la distribution de Rayleigh serait complètement justifiée si X et Y étaient complètement indépendants.

Pour l'image naturelle bruitée introduite plus haut (**peppers**), on observe l'histogramme de la figure 3.11 pour la première échelle du module de la partie Riesz. On retrouve la forme d'une distribution de Rayleigh, conformément à l'hypothèse selon laquelle la première échelle ne contient que du bruit.

Une illustration plus complète est donnée figure 3.12 avec les histogrammes sur plusieurs échelles. On voit que la dispersion augmente lorsque l'information utile devient, avec l'échelle, plus importante que le bruit. Il semble donc être une bonne idée d'estimer le seuil sur la première échelle, ce qui éliminera tous les coefficients de bruit tout en préservant l'information structurale dans les autres échelles.

Finalement, le modèle gaussien sera choisi pour la modélisation des sous-bandes primaires,

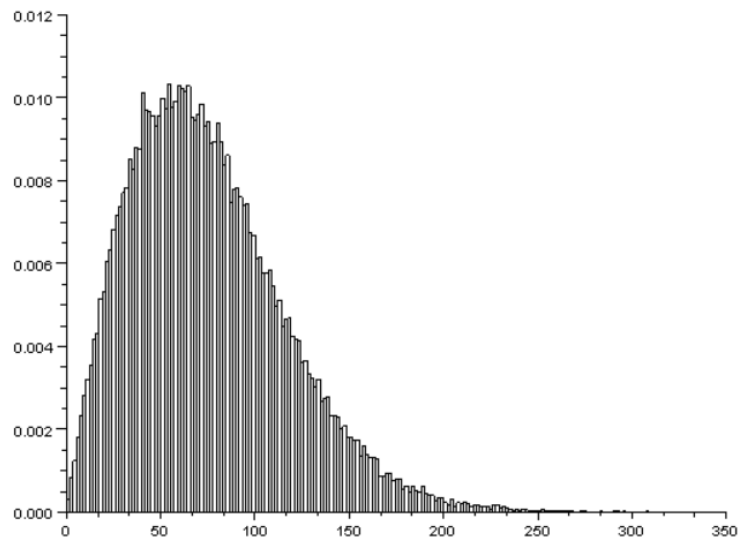


FIGURE 3.11 – *Histogramme du module de la partie Riesz à la première échelle pour l'image `peppers` bruitée.*

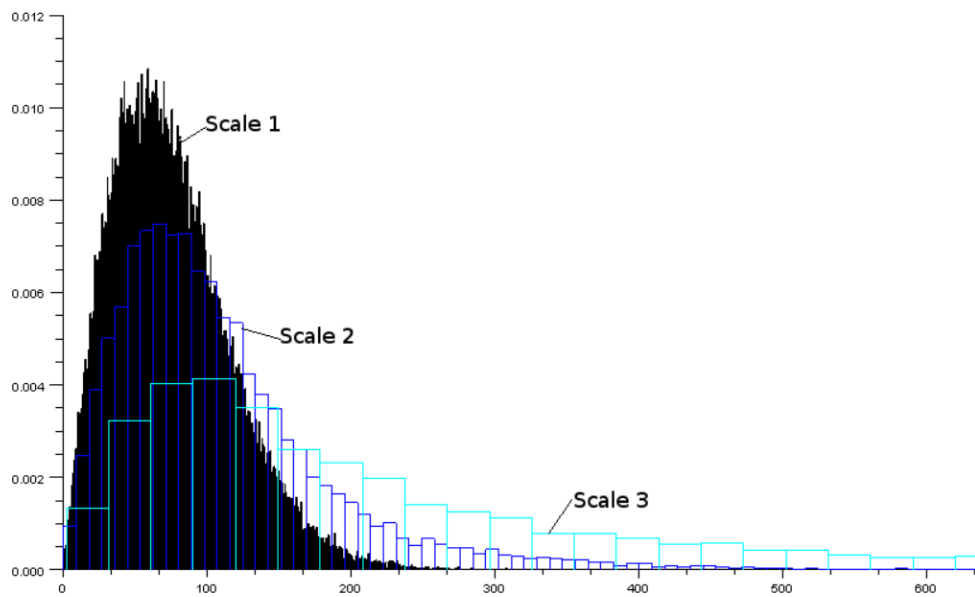


FIGURE 3.12 – *Histogrammes du module de la partie Riesz sur trois échelles.*

ainsi qu'une distribution de Rayleigh pour le module de la partie Riesz. Nous pouvons maintenant appliquer cette méthode de seuillage pour débruiter des images couleur.

3.5.2 Seuillage des coefficients

Nous proposons de nous concentrer d'abord sur le seuillage de la partie primaire - qui est plus intuitif. La procédure est la suivante :

- Estimation du niveau de bruit à partir de la première échelle,
- Seuillage « doux » de toutes les échelles avec un seuil égal à trois fois l'écart-type du bruit mesuré.

L'utilisation de la première échelle pour l'estimation du bruit est classique [102], et permet d'utiliser efficacement la valeur *médiane* du module pour estimer l'écart-type d'une gaussienne : $\hat{\sigma} = \text{median}(|c_1^R|)/0.6745$ (avec c_1^R les coefficients de la première échelle).

Le seuillage doux est défini comme suit :

$$z_t = \text{sgn}(z) [|z| - g]_+ \quad (3.92)$$

avec z le coefficient, g le seuil et $[\cdot]_+$ vaut $\max(\cdot, 0)$.

Comme il a déjà été discuté à la section 3.4.3 au sujet de la reconstruction de l'image à partir des coefficients, plusieurs méthodes sont possibles pour obtenir l'image débruitée. Dans notre travail, il semble important de faire en sorte que la partie Riesz soit influente dans la reconstruction, et donc dans la procédure de débruitage, pour éviter de définir une méthode *marginale*. La partie Riesz apporte une analyse géométrique par rapport à la partie primaire. Mais la combinaison des deux parties dans une reconstruction unifiée est un sujet ouvert. Dans ce travail, nous proposons d'utiliser la partie Riesz dans la procédure de seuillage mais pas dans la reconstruction. La reconstruction de l'image débruitée se fera donc à partir de la partie primaire seulement, mais le seuillage de la partie primaire sera *conditionné* par les informations contenues dans la partie Riesz, afin d'utiliser complètement la représentation monogène couleur proposée.

Le seuillage à partir du module des coefficients de la partie Riesz n'est pas évident à cause de leur distribution de Rayleigh, pour laquelle les méthodes classiques ne tiennent plus. Un problème similaire a été rencontré dans [86] avec des ondelettes de Gabor. Cette référence suggère le seuil suivant :

$$E(Z) + k * \sqrt{V(Z)} \quad (3.93)$$

où $E(Z)$ et $V(Z)$ sont les moments de la distribution de Rayleigh qui ont déjà été données plus haut, et qui sont estimés à partir du module des coefficients de la partie Riesz à la première échelle. Nous proposons d'utiliser $k = 3$ qui a expérimentalement donné des résultats satisfaisants. Dans le contexte des histogrammes affichés à la figure 3.12, cela correspond au seuil 193. Ce graphique montre qu'un tel seuil élimine la plus grande partie du bruit de la première échelle, tout en préservant les forts coefficients des autres échelles.

La combinaison des deux seuillages (partie primaire et partie Riesz) se fait par une opération logique type « ET ». A une position, on garde le coefficient primaire seulement si les deux parties sont au dessus de leurs seuils respectifs. Cette étape pourrait bien sûr être facilement améliorée à travers une étude approfondie de la combinaison de l'information colorimétrique et de l'information géométrique de ce genre de représentation monogène couleur, mais ce n'est pas l'objectif premier.

3.5.3 Résultats et analyse

Nous proposons ici nos différents résultats de débruitage couleur sur les figures 3.13, 3.14, 3.15 et 3.16.

La première remarque est que la dégradation introduite est assez forte, ce qui va nous permettre de bien discerner les résultats de différentes méthodes. Avec de plus faibles dégradations, la



FIGURE 3.13 – Débruitage de l'image *house*. Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.



FIGURE 3.14 – Débruitage de l'image *peppers*. Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.



FIGURE 3.15 – Débruitage de l'image *parrot*. Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.

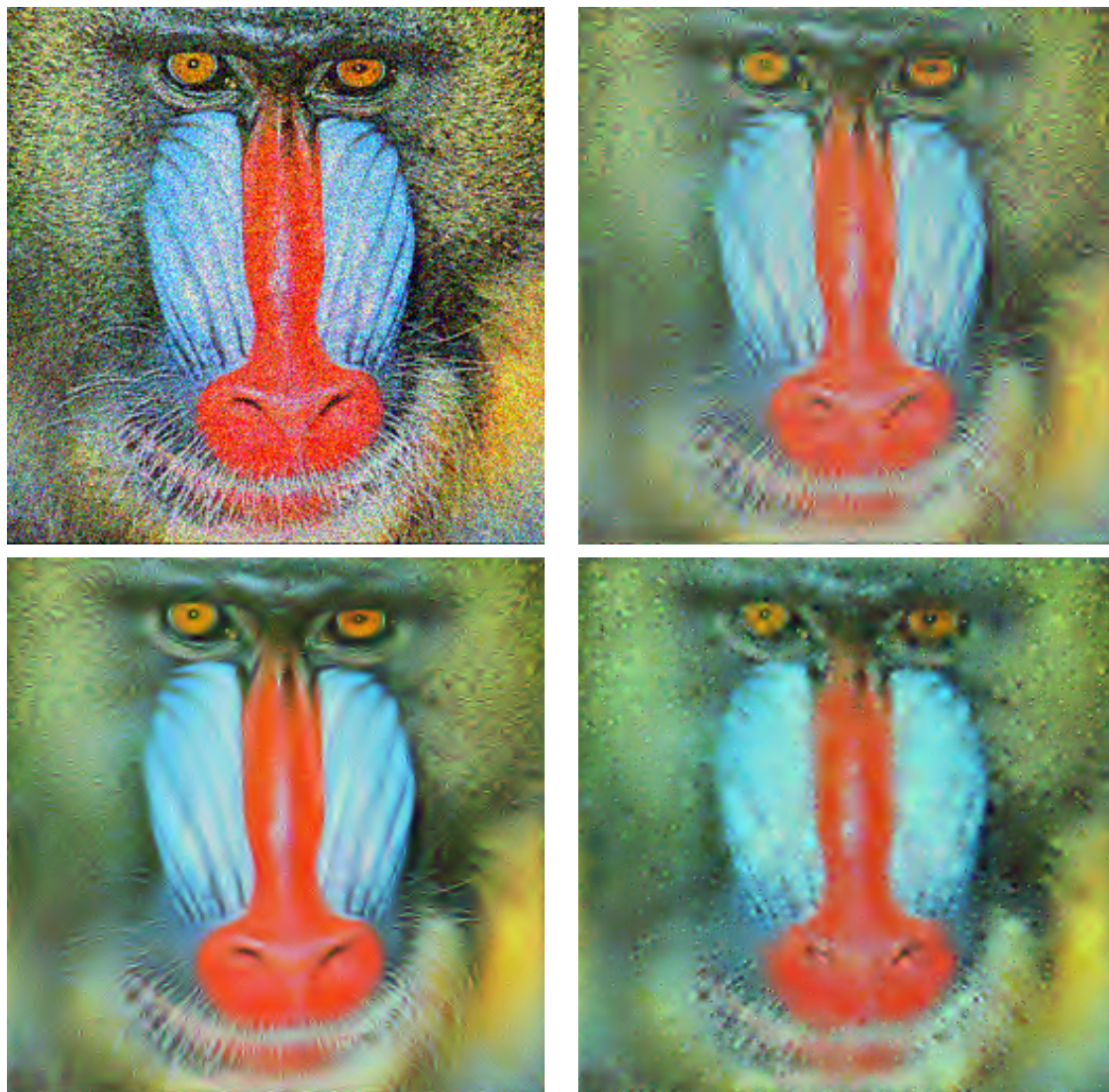


FIGURE 3.16 – Débruitage de l'image *mandrill*. Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.

comparaison serait plus difficile visuellement. Notre approche est comparée aux deux techniques classiques en débruitage par ondelettes que sont :

- Le seuillage doux sur les coefficients d'une transformée en ondelettes orthogonale (filtres Daubechies 4)
- Le seuillage dur sur les coefficients d'une transformée non décimée (Daubechies 4)

Les seuils sont déterminés à partir de l'estimation de la variance du bruit décrite plus haut, et la sous-bande basses-fréquences n'est pas modifiée.

Expérimentalement, et de manière générale, la performance des ondelettes monogènes couleur se place *entre* celles des ondelettes orthogonales et celles des ondelettes non décimées. Ces dernières ont une forte redondance (L échelles sur une image de N pixels donnera $(3 * L + 1) * N$ coefficients) tout en gardant une certaine orthogonalité entre les décompositions entrelacées, ce qui en fait un des meilleurs outils de débruitage connus à ce jour. Au contraire, les ondelettes orthogonales sont connues pour introduire des oscillations indésirables autour des contours - c'est le phénomène pseudo-Gibbs - constituant la contrepartie pour leur rapidité d'implantation.

On remarque que l'approche monogène couleur préserve l'information couleur principale des images pour une redondance limitée par rapport à la transformée non-décimée. La couleur au niveau des contours en particulier, est bien respectée. Les textures sont considérablement lissées, par exemple sur l'image **mandrill**, figure 3.16, ce qui est classique en débruitage quelle que soit la méthode. Aucune fausse couleur n'est introduite, contrairement aux artéfacts colorés introduits par les ondelettes orthogonales à la figure 3.15 : des artéfacts verts autour de la zone noire du bec du perroquet rouge, des artéfacts jaunes sur la droite du bec, des artéfacts verts sur la grande zone jaune du perroquet de droite.

Des artéfacts sont néanmoins introduits par la méthode proposée, mais en comparaison aux ondelettes orthogonales, ils sont d'une part plus *lisses*, ce qui est dû à la redondance de la représentation, et d'autre part plus *ronds*, grâce à l'isotropie, et donc *a priori* moins gênants visuellement.

Des améliorations du côté du schéma numérique utilisé sont clairement suggérées ici, car les ondelettes polyharmoniques du banc de filtres de [144] ont été ajustées pour l'analyse, et non pour la synthèse. En conséquence, les ondelettes du banc de filtres de reconstruction pourraient être améliorées pour diminuer les artéfacts.

Finalement, avec cette transformée monogène couleur, les résultats sont de même nature que l'expérience comparative que nous avons menée pour tester la QWT en codage d'image (section 2.4). L'utilisation du formalisme du signal analytique place ces nouvelles transformées dans la catégorie « légèrement redondantes », en améliorant la puissance d'analyse des coefficients, et permettant ainsi de surpasser les ondelettes décimées dans un contexte très dégradé. Ici, l'utilisation du signal analytique nous a également permis de définir une transformée *couleur* non-triviale.

Voici maintenant les limites de l'outil, que cette expérimentation a mises en avant. La première selon nous est le manque de définition claire de la reconstruction à partir d'une décomposition monogène. Le banc de filtres de [144] ne propose pas un algorithme de reconstruction unifié, mais plutôt deux reconstructions séparées, et ne donne pas de piste pour fusionner les coefficients. La question même de la « forme » d'éventuels atomes monogènes de reconstruction n'est pas traitée dans la littérature. C'est pourquoi les quelques applications que l'on trouve sont uniquement des applications d'analyse d'image, et cela explique sans doute les difficultés que nous avons eues à définir une reconstruction pour la transformée couleur proposée. En effet, le formalisme du signal monogène couleur de [39] n'apporte pas non plus de solution pour la reconstruction. De plus, la notion de phase proposée dans [39] n'est pas assez aboutie pour en avoir une interprétation physique liée aux images couleur. Enfin, bien que cette extension couleur ne soit pas strictement marginale, la notion d'entité couleur et de géométrie couleur est mal prise en compte. Notamment, la transformée de Riesz étant appliquée sur l'*intensité*, l'orientation des contours isoluminants n'est pas accessible. Plus généralement, il n'y a pas de modèle de structure géométrique couleur associé à cette représentation. L'interprétation des coefficients est

donc difficile, et la contribution du formalisme monogène dans les performances de débruitage couleur est difficile à identifier.

Nous pensons que la principale voie d'amélioration autour de ce travail est dans la définition du concept d'analyse monogène couleur, qui devrait être plus dirigé vers l'interprétation physique des coefficients. La méthode proposée ici manque d'unité entre ses différentes composantes, et de lien avec les caractéristiques visuelles des images.

3.6 BILAN ET CONCLUSION DU CHAPITRE

Après avoir récapitulé quelques outils classiques de vision bas-niveau, qui constituent les principaux travaux existant autour du concept de *phase* pour les images, nous avons étudié dans ce chapitre une extension isotrope du signal analytique : le signal monogène. Cette définition est l'évolution des méthodes précédentes du point de vue de l'analyse de l'orientation et de la phase locale, et s'inclut dans un nouveau formalisme de traitement du signal dû à M. Felsberg, où l'analyse multi-échelle d'une image se fait par des filtres en quadrature sphériques, construits par un filtre passe-bande isotrope et les deux composantes de sa transformée de Riesz. Contrairement aux décompositions type Gabor qui fournissent K amplitudes et K phases, selon la résolution angulaire voulue, l'analyse monogène donne une amplitude et une phase uniques, jointes à une donnée d'orientation locale précise. L'orientation n'est donc plus une « coordonnée » comme la position ou la fréquence, mais une « caractéristique » continue qui s'ajoute aux informations d'amplitude et de phase. Ce changement s'accompagne d'une nouvelle propriété d'invariance par rotation du signal, ce qui est très avantageux dans de nombreuses applications d'analyse d'image.

La dimension numérique des représentations monogènes multi-échelles a été ouverte par l'article de M. Unser *et al.* [144], qui propose d'utiliser deux bancs de filtres à reconstruction parfaite. Cette méthode s'appuie sur des techniques avancées de définition d'ondelettes 2D, et permet d'obtenir une invariance par rotation approximative. Au final, il s'avère que la reconstruction du signal à partir des coefficients monogènes n'est pas vraiment étudiée, ce qui restreint cet outil aux seules applications d'analyse d'image.

Etant donné que le formalisme des ondelettes analytiques implique des problématiques liées aux dimensions des données, il nous a paru propice d'étudier l'extension de ces méthodes pour les images couleur - donc les valeurs sont de dimension 3. A partir de l'article [39] qui propose un signal monogène couleur, nous avons donc proposé une nouvelle transformée en ondelettes spécialement définie pour analyser les images couleur. Cette transformée est non-séparable, approximativement invariante par rotation, non-marginale, et fournit des coefficients sous forme de sous-bandes monogènes. Grâce à la prise en compte d'un signal multi-dimensionnel à valeurs vectorielles dès le début de la construction mathématique, nous pensons que ce type de transformée en ondelette peut potentiellement améliorer les représentations existantes, à travers une meilleure prise en compte de la couleur et de la géométrie des structures locales des images. Cette définition est une première étape, que nous avons prise sous un angle d'étude de faisabilité, afin d'avancer dans la compréhension des ondelettes analytiques pour l'image, dans son état actuel.

Afin de situer cette première définition d'un point de vue pratique, nous avons conduit une expérimentation comparative en débruitage couleur, par seuillage d'ondelettes. A partir d'une modélisation des coefficients de la transformée, nous avons proposé une méthode de seuillage, que nous avons comparée aux méthodes classiques par ondelettes séparables orthogonales et ondelettes non-décimées. Notre transformée donne des performances intermédiaires par rapport aux deux autres, ce que l'on peut rapprocher du résultat obtenu en codage d'image par QWT (section 2.4). Il vient que la faible redondance de la représentation suffit à gagner de bonnes propriétés par rapport aux ondelettes orthogonales, sans pour autant surpasser les ondelettes non-décimées qui sont elles, très redondantes. La transformée monogène couleur offre donc un

compromis avantageux, en se plaçant dans la catégorie des représentations légèrement redondantes pour une meilleure capacité d'analyse. Notons toutefois que les performances pourraient certainement être améliorées en remaniant la définition des bancs de filtres de [144], qui ont été optimisés pour l'analyse, au détriment de la qualité de reconstruction.

Cependant, dans ces résultats, la contribution du formalisme monogène couleur est difficile à identifier pour deux raisons. D'une part, les coefficients monogènes couleur n'ont pas d'interprétation aussi univoque que le signal monogène « niveaux de gris ». Nous n'avons pas de notion d'amplitude/phase, excepté une définition de « phase » paramétrique pour laquelle il est difficile de faire le lien avec une donnée physique concrète. D'autre part, cette méthode hérite des définitions en niveaux de gris une reconstruction mal définie.

Malgré l'intérêt certain de cette généralisation isotrope du signal analytique, l'aspect numérique d'une part, et extension couleur d'autre part, nécessitent une réflexion nouvelle. C'est pourquoi nous proposons maintenant de définir une nouvelle représentation monogène couleur.

UNE NOUVELLE REPRÉSENTATION MONOGÈNE COULEUR

DANS le chapitre précédent, nous avons proposé une première généralisation de la représentation en ondelettes monogènes à la couleur à partir de concepts très mathématiques : les équations de Cauchy-Riemann généralisées et les algèbres géométriques. Nous allons ici tenter une deuxième stratégie visant à mieux décrire l'information visuelle, en partant d'idées directement issues de la communauté de traitement d'images couleur.

La problématique est de définir des outils d'analyse et de traitement dont l'entrée est un signal à valeurs *vectorielles*. En effet, contrairement aux niveaux de gris, les couleurs sont encodées la plupart du temps par *plusieurs* nombres, selon un *espace colorimétrique* donné. On a donc un problème pour calculer des *distances* entre couleurs, ainsi que pour les *ordonner*.

La première approche, la plus simple, consiste à ne considérer que l'*intensité* de l'image, qui est analogue à un niveau de gris. Mais il est impossible dans ce cas d'exploiter l'information couleur, qui est pourtant primordiale dans certaines applications. La seconde approche est la méthode dite *marginale*, qui est d'appliquer un outil classique - scalaire - sur chaque composante indépendamment. Un filtrage linéaire par exemple, comme un moyenneur, peut très bien être réalisé de façon marginale (voir le lissage de l'image bruitée à la figure 4.1). Mais plusieurs méthodes de traitement d'image sont non-linéaires. Le filtrage médian par exemple, qui donne de meilleurs résultats pour un bruit impulsif en préservant mieux les contours que les filtres linéaires, produit de « fausses couleurs » lorsqu'on l'applique de façon marginale. Ce phénomène, illustré à la figure 4.1 par la 4ème image, sur laquelle du jaune est apparu, montre les limites

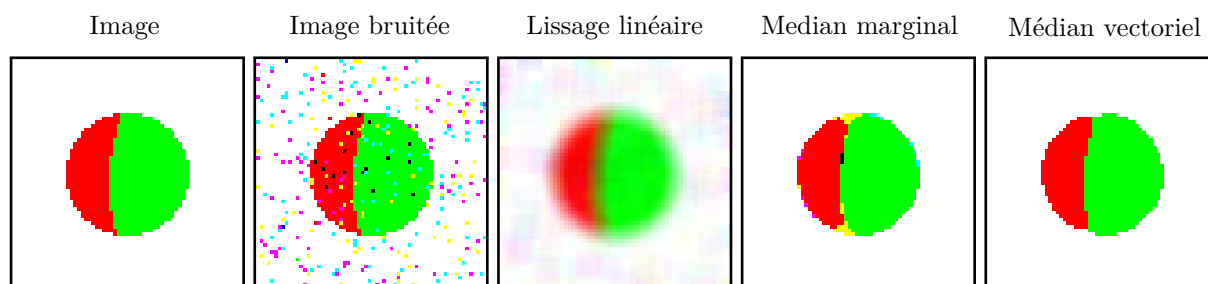


FIGURE 4.1 – Filtrage linéaire (moyennage) et non-linéaire (médian) d'une image couleur. (Fenêtre de taille 7×7).

des méthodes marginales. Ces limites pointent le besoin d'outils vectoriels, prenant en compte les couleurs comme des entités. Par exemple, il existe une méthode classique de filtrage médian vectoriel, impliquant des *distances euclidiennes* entre les couleurs, et permettant d'améliorer considérablement le débruitage couleur de la figure 4.1. Un autre problème des méthodes marginales apparaît en *analyse* d'image. En effet, l'analyse requiert souvent l'extraction de caractéristiques géométriques et colorimétriques, qui doivent prendre en compte toutes les composantes dans un même calcul.

Les représentations en ondelettes - auxquelles nous nous intéressons dans ce travail de thèse - sont presque toujours définies pour des signaux scalaires. Nous pouvons trouver quelques constructions couleur dans la littérature. Un algorithme de lifting vectoriel [8, 75] a été proposé pour le codage d'images multi-spectrales ou stéréoscopiques. Cependant cette méthode est uniquement destinée à la compression sans perte, et ne développe pas de propriétés d'analyse telles que les représentations analytiques que nous étudions. Une transformée en ondelettes couleur définie dans l'algèbre des triplets a été proposée dans [88], mais n'a ni propriété particulière, ni schéma numérique associé. Enfin, les ondelettes quaternioniques couleur [20] dont nous avons parlé à la section 2.1.2, constituent une première piste autour d'une représentation multi-résolution destinée à prendre en compte l'information couleur de façon efficace aussi bien en termes d'analyse que de traitement. Cette représentation quaternionique pose un problème de *paramètre* de direction d'analyse dans l'espace couleur, que l'on ne sait pas fixer.

Ce chapitre présente une définition innovante de transformée en ondelettes couleur, pour laquelle nous avons choisi de suivre la piste des ondelettes monogènes, qui constituent comme nous l'avons vu une évolution importante en termes d'efficacité de représentation des images. Le but est de définir une extension couleur de la représentation monogène offrant d'une part une interprétation des coefficients en termes de caractéristiques visuelles, et d'autre part un algorithme de reconstruction cohérent.

Sur les espaces colorimétriques

Il existe de nombreuses façons de représenter les couleurs, à travers différents « espaces colorimétriques », plus ou moins adaptés à la perception humaine, ou aux appareils de mesure ou de restitution. Dans les cas les plus connus - notamment l'espace RGB, et les espaces intensité-chrominance - l'espace est de dimension 3. La première étape pour définir des outils adaptés à la couleur est selon nous de définir des outils prenant en compte des données *vectérielles* en général. Nous considérerons donc les images couleur comme des fonctions :

$$s : \mathbb{R}^2 \mapsto \mathbb{R}^3 \quad (4.1)$$

$$\mathbf{x} \mapsto [s^R(\mathbf{x}) \ s^G(\mathbf{x}) \ s^B(\mathbf{x})]^T$$

sans détermination concrète de l'espace colorimétrique utilisé. Néanmoins, par souci de simplicité et d'interprétation des résultats, les expérimentations seront faites avec l'espace RGB, pour lequel les trois quantités correspondent à l'intensité des composantes primaires *rouge*, *vert* et *bleu*. De plus, un grand avantage de cet espace est qu'il est euclidien, c'est-à-dire que l'on dispose d'une mesure simple de distance entre deux couleurs : la norme euclidienne. Bien que l'espace RGB soit mal adapté à une considération perceptuelle de la couleur, nous avons choisi de laisser la prise en compte d'espaces plus adaptés dans les perspectives de cette thèse.

Le chapitre est organisé ainsi. Nous commencerons avec l'extension couleur du gradient, pour ensuite définir une transformée de Riesz pour les images couleur. A partir d'une extension de la notion de phase, nous construirons un modèle innovant de signal monogène couleur. Nous finirons avec la définition d'une transformée en ondelettes correspondante.

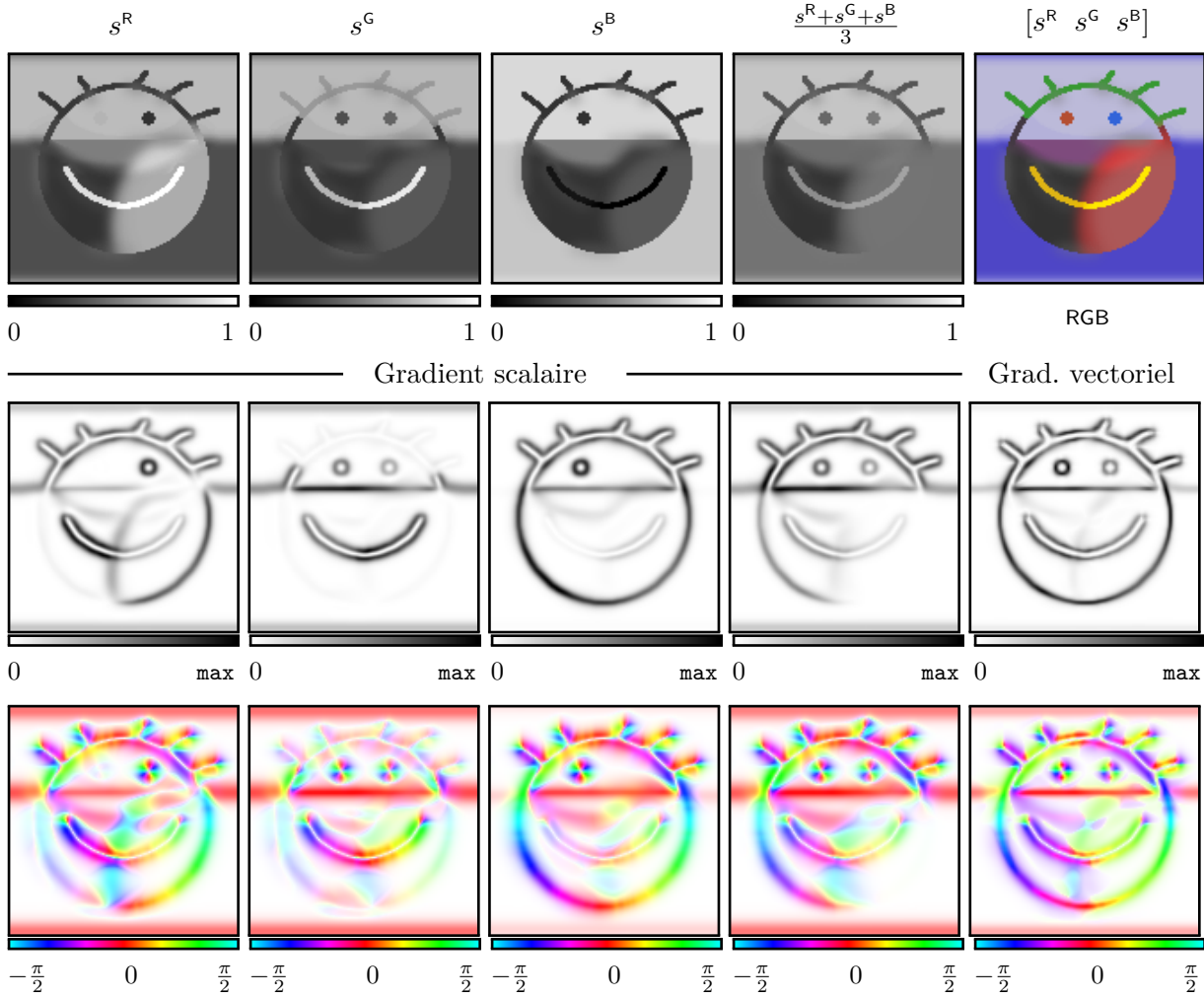


FIGURE 4.2 – Gradient d’une image couleur $s = [s^R \ s^G \ s^B]$. Les quatre premières colonnes montrent l’utilisation d’un gradient classique sur les canaux couleur de s ainsi que sur son intensité. La cinquième colonne illustre le gradient vectoriel. Dans les 5 cas, on affiche la norme (2ème ligne) et la direction (3ème ligne) du gradient.

4.1 VARIATIONS LOCALES D’UNE IMAGE COULEUR

Cette section propose d’étendre la discussion menée dans la section 3.1.1 sur le tenseur de structure basé gradient que nous avons étudié dans le cas d’images *scalaires*.

Nous rappelons que le gradient ∇ , défini comme le vecteur des dérivées partielles selon les axes x et y , donne accès en tout point à une mesure de variation $\mathcal{N}_{\nabla}(\mathbf{x})$, associée à une direction principale $\theta_{\nabla}(\mathbf{x})$ selon un modèle de structure locale de type « bord rectiligne ». Dans le cas où les valeurs prises par la fonction étudiée sont vectorielles, il y a globalement deux utilisations possibles de ∇ .

La première est *marginale*, c’est-à-dire qu’elle consiste à appliquer le gradient indépendamment sur chaque composante de la fonction. Cette méthode, illustrée à la figure 4.2 (sur les 3 premières colonnes), dispatche les informations de contours sur les trois canaux couleur. On voit que *certaines* contours sont donc détectés à partir de *certaines* canaux, et l’orientation mesurée en un point peut être différente d’un canal à l’autre, comme c’est le cas sur le bas de la bouche pour s^R et s^G . Notons qu’une stratégie de « fusion » consiste à utiliser la somme des trois *normes* de gradient, ce qui permet de détecter tous les contours. Cependant, cette méthode ne correspond pas à un modèle précis et il n’y a pas de méthode pour *unifier* les trois orientations. L’information obtenue par une utilisation *marginale* du gradient n’est donc pas suffisante.

La seconde méthode consiste à se ramener à une image scalaire, en étudiant l’intensité de

l'image couleur, ici définie par la moyenne des trois canaux. La 4ème colonne de la figure 4.2 permet de voir que cette méthode n'exploite pas l'information couleur contenue dans l'image, ce qui la rend incapable de détecter les contours dits « isoluminants ». Ces contours séparent deux couleurs différentes mais d'*intensité semblable* (bas-droit du visage), et disparaissent dans une étude basée sur l'intensité. On a donc une perte d'information.

C'est de ce constat qu'est née l'idée de définir une norme et une orientation basées sur un modèle vectoriel, prenant en compte les variations locales de la fonction à travers toutes ses composantes simultanément. La première proposition est le gradient couleur de S. Di Zenzo en 1986 [43], et l'idée a ensuite été développée grâce à l'outillage de la géométrie différentielle, notamment dans [123, 142, 126].

Partant du principe qu'une dérivée partielle mesure une *différence* entre deux valeurs voisines, le cas des valeurs vectorielles entraîne donc une dérivée à valeurs vectorielles elle aussi. Par extension des équations (3.5) et (3.6) (page 71) définissant la dérivée partielle orientée D_θ par rotation de $\partial/\partial x$, on définit donc la dérivée partielle orientée d'un signal vectoriel comme ceci :

$$D_\theta s(\mathbf{x}) = \begin{bmatrix} \cos(\theta) \frac{\partial s^R}{\partial x_1} + \sin(\theta) \frac{\partial s^R}{\partial x_2} \\ \cos(\theta) \frac{\partial s^G}{\partial x_1} + \sin(\theta) \frac{\partial s^G}{\partial x_2} \\ \cos(\theta) \frac{\partial s^B}{\partial x_1} + \sin(\theta) \frac{\partial s^B}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \mathcal{N}_\nabla^R \cos(\theta - \theta_\nabla^R) \\ \mathcal{N}_\nabla^G \cos(\theta - \theta_\nabla^G) \\ \mathcal{N}_\nabla^B \cos(\theta - \theta_\nabla^B) \end{bmatrix} \quad (4.2)$$

La « quantité de variation » entre deux vecteurs voisins peut intuitivement être représentée par la norme euclidienne $\|D_\theta s\|$, qui mesure la *distance* entre eux. On peut alors reprendre l'idée de maximisation de cette variation par rapport à θ , avec le passage au carré $\|D_\theta s\|^2$ qui permet de fusionner les directions opposées. Cette forme quadratique correspond au tenseur de structure suivant :

$$\|D_\theta s\|^2 = \begin{bmatrix} \cos \theta & \sin \theta \end{bmatrix} \underbrace{\begin{bmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{bmatrix}}_T \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad (4.3)$$

avec

$$T_{11} = \left(\frac{\partial s^R}{\partial x_1} \right)^2 + \left(\frac{\partial s^G}{\partial x_1} \right)^2 + \left(\frac{\partial s^B}{\partial x_1} \right)^2 \quad (4.4)$$

$$T_{12} = \frac{\partial s^R}{\partial x_1} \frac{\partial s^R}{\partial x_2} + \frac{\partial s^G}{\partial x_1} \frac{\partial s^G}{\partial x_2} + \frac{\partial s^B}{\partial x_1} \frac{\partial s^B}{\partial x_2} \quad (4.5)$$

$$T_{22} = \left(\frac{\partial s^R}{\partial x_2} \right)^2 + \left(\frac{\partial s^G}{\partial x_2} \right)^2 + \left(\frac{\partial s^B}{\partial x_2} \right)^2 \quad (4.6)$$

La norme $\mathcal{N}$ et la direction θ_+ de la variation maximale sont à nouveau obtenues des valeurs/vecteurs propres de T . Ainsi :

$$\mathcal{N} = \sqrt{\lambda_+ + \lambda_-} = \sqrt{T_{11} + T_{22}} \quad \theta_+ = \frac{1}{2} \arg\{T_{11} - T_{22} + j2T_{12}\} \quad (4.7)$$

Il est intéressant de voir que cette définition de norme et de direction de gradient couleur correspond à une combinaison des analyses marginales, à savoir :

$$(\lambda_+ - \lambda_-) \begin{bmatrix} \cos(2\theta_+) \\ \sin(2\theta_+) \end{bmatrix} = (\mathcal{N}_\nabla^R)^2 \begin{bmatrix} \cos(2\theta_\nabla^R) \\ \sin(2\theta_\nabla^R) \end{bmatrix} + (\mathcal{N}_\nabla^G)^2 \begin{bmatrix} \cos(2\theta_\nabla^G) \\ \sin(2\theta_\nabla^G) \end{bmatrix} + (\mathcal{N}_\nabla^B)^2 \begin{bmatrix} \cos(2\theta_\nabla^B) \\ \sin(2\theta_\nabla^B) \end{bmatrix} \quad (4.8)$$

L'orientation d'un contour couleur est donc calculée comme l'orientation principale « moyenne » des composantes marginales. Voici donc les données qui seront utilisées pour détecter les contours d'une image couleur de façon efficace, comme illustré à la figure 4.2 (5ème colonne).

Rappelons que dans le cas scalaire, il faut introduire un lissage pour que le tenseur de structure soit consistant, en faisant apparaître une notion de *coherence* (équation (3.11)). Si ce

lissage n'est pas appliqué, on a $T_{11}T_{22} = T_{12}^2$, ce qui entraîne que le second vecteur propre de T est nul $\lambda_- = 0$, d'où la cohérence vaut toujours $\chi = 1$. L'analyse donnée par le tenseur est donc dégénérée, et on ne peut pas discriminer les structures ne correspondant pas au modèle de bord rectiligne. Ici, il n'est pas nécessaire de lisser les termes de T , car sa structure vectorielle introduit une *nouvelle* information de cohérence. En effet, bien que nous n'ayons pas utilisé de lissage dans cette définition, on constate que $T_{11}T_{22} \neq T_{12}^2$. On peut montrer que l'égalité est retrouvée si et seulement si les gradients marginaux sont *colinéaires*, c'est-à-dire $\theta_{\nabla}^R = \theta_{\nabla}^G = \theta_{\nabla}^B$. Associé à cela, on peut montrer que la variation $\|D_{\theta}s\|^2$ est une sinusoïde (plus une constante) dont le maximum $\mathcal{N}$ est atteint pour $\theta = \theta_+$, mais dont le minimum en $\theta = \theta_+ + \pi/2$ n'est pas forcément nul. Ce minimum devient nul si les gradients marginaux sont colinéaires. Cela signifie que la notion de bord rectiligne couleur correspond à une rupture « complète », c'est-à-dire que les canaux couleurs doivent être synchronisés dans cette variation. Ce modèle est satisfaisant d'un point de vue perceptuel : une structure perçue comme un bord qui sépare deux régions de couleurs distinctes aura en général une orientation *unique*. La variation entre les deux couleurs se traduit alors par trois variations marginales dans cette même direction. Dans le cas contraire, la notion de bord rectiligne est plus confuse, ce qui se traduit par une valeur de λ_- non négligeable, et qui biaise l'information portée par θ_+ , qui indiquera alors une direction « moyenne » non pertinente. En pratique, les contours que nous percevons sont correctement détectés, comme le montre la figure 4.2, mais nous pourrions dans la suite garder à l'esprit que le modèle sous-jacent au gradient couleur est restreint aux structures vectorielles dont les composantes varient dans la même direction principale.

Le gradient couleur est une méthode fondamentale en image, qui permet de détecter efficacement tous les contours couleur, même ceux dont l'intensité est constante (« isoluminants »). Nous proposons maintenant de définir un signal monogène couleur par extension de cet outil.

4.2 UN NOUVEAU SIGNAL MONOGÈNE COULEUR

Puisque le signal monogène est construit sur la transformée de Riesz, elle-même liée de près au gradient, nous proposons de tirer parti du gradient couleur, qui est un outil bien établi, afin de définir un signal monogène couleur prenant en compte tous les contours et calculant correctement leur orientation principale ainsi que leur phase.

4.2.1 Transformée de Riesz couleur

Nous avons vu que la transformée de Riesz est égale à un gradient appliqué sur une version lissée du signal (section 3.2.1). C'est ce qui lui confère la capacité de détecter l'orientation locale des structures modélisées dans le signal monogène.

Pour les images en niveaux de gris, nous nous sommes intéressés à la norme $\mathcal{N}_{\mathcal{R}}$ et à la direction $\theta_{\mathcal{R}}$ de la transformée de Riesz, qui décrivaient pleinement l'information structurelle de l'image, c'est pourquoi nous proposons maintenant de redéfinir ces mesures à partir du formalisme du gradient couleur.

On définit les deux composantes de la transformée de Riesz « vectorielle » (marginale) de cette façon :

$$\mathcal{R}_0 s = \begin{bmatrix} \mathcal{R}_0 s^R & \mathcal{R}_0 s^G & \mathcal{R}_0 s^B \end{bmatrix}^T \quad (4.9)$$

$$\mathcal{R}_{\frac{\pi}{2}} s = \begin{bmatrix} \mathcal{R}_{\frac{\pi}{2}} s^R & \mathcal{R}_{\frac{\pi}{2}} s^G & \mathcal{R}_{\frac{\pi}{2}} s^B \end{bmatrix}^T \quad (4.10)$$

L'aspect non-marginal de notre méthode réside dans une extraction non-linéaire de la norme et

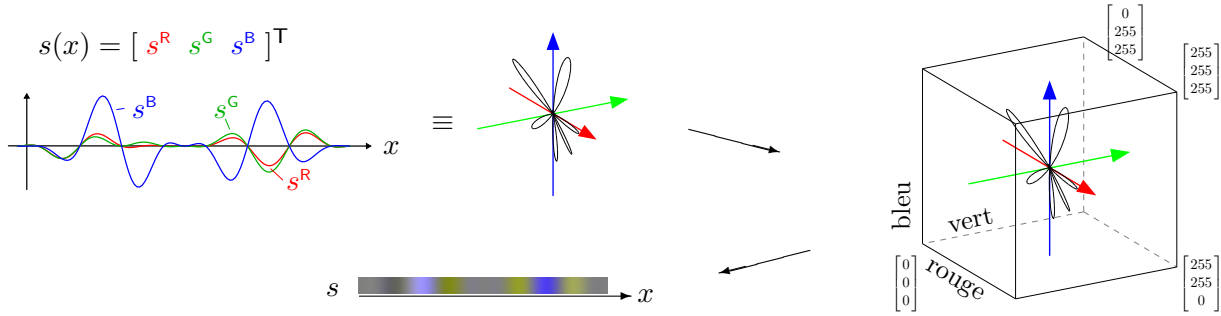


FIGURE 4.3 – Illustration d'un signal à valeurs dans $\mathbb{R}^3$: exemple 1D. On identifie les vecteurs comme des éléments du cube RGB, en plaçant $[0, 0, 0]$ au centre, c'est-à-dire sur le « gris moyen ». Les « couleurs » observées n'en sont donc pas à proprement parler.

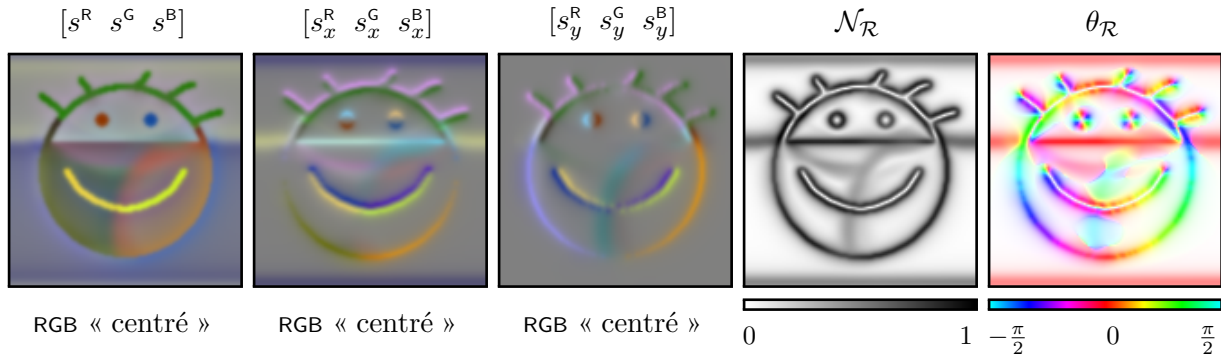


FIGURE 4.4 – Transformée de Riesz couleur proposée.

de l'orientation, grâce au formalisme du gradient couleur, qui nous donne de façon évidente :

$$\mathcal{N}_{\mathcal{R}} = \sqrt{\|\mathcal{R}_0 s\|^2 + \|\mathcal{R}_{\frac{\pi}{2}} s\|^2} \quad (4.11)$$

$$\theta_{\mathcal{R}} = \frac{1}{2} \arg \left\{ \|\mathcal{R}_0 s\|^2 - \|\mathcal{R}_{\frac{\pi}{2}} s\|^2 + 2j (\mathcal{R}_0 s)^T (\mathcal{R}_{\frac{\pi}{2}} s) \right\} \quad (4.12)$$

Rappelons que la transformée de Riesz n'a de sens physique que lorsque le signal étudié est à bande-étroite. Dans la suite, nous utiliserons une méthode multi-échelle qui permettra de décomposer l'analyse en sous-bandes vérifiant cette condition. Pour l'instant, nous devons utiliser un filtrage passe-bande isotrope, comme nous l'avons fait en niveaux de gris.

Afin d'illustrer les données utilisées, nous présentons ici la procédure d'affichage d'une « sous-bande » couleur. Un tel signal correspond à trois matrices dont les valeurs réelles peuvent être négatives. Nous avons choisi d'illustrer ces données par une image couleur, dont les pixels sont définis par normalisation et centrage de la sous-bande couleur, comme schématisé à la figure 4.3. Il est important de noter que les couleurs visibles sur ces illustrations ne sont pas directement liées aux couleurs réelles contenues dans l'image, mais permettent de représenter les données. Cette méthode sera utilisée dans la suite de ce mémoire pour toutes les sous-bandes couleur.

Nous pouvons maintenant illustrer la transformée de Riesz couleur ainsi construite. La figure 4.4 montre une sous-bande de l'image de test que nous avons utilisée dans la figure 4.2. Le filtrage linéaire se calcule sur des signaux vectoriels de façon marginale, c'est donc comme cela que l'image a été filtrée, avec le même filtre que précédemment (DoP, $k = 3$). Les deux composantes de la transformée de Riesz vectorielle sont ensuite affichées, et représentent un *déphasage* couleur dans les directions x et y . On constate enfin que la norme et la direction proposées donnent la même analyse que le gradient couleur, ce qui signifie qu'une partie des informations de la représentation monogène - l'orientation locale - est déjà correctement étendue à la couleur.

Du point de vue de la construction d'un signal analytique couleur, notre extension de la transformée de Riesz consiste au final à considérer la norme euclidienne $\mathcal{N}_{\mathcal{R}}$, qui remplace alors le « module » $|\mathcal{R}s|$ du cas scalaire, qui lui-même remplaçait la transformée de Hilbert $\mathcal{H}s$ du

cas 1D. Ce choix est justifié par le sens physique de cette quantité, à savoir la force de variation locale, qui traduit directement une notion de bord rectiligne et donc d'analyse géométrique des structures couleur de l'image. De plus, du fait du lien théorique avec les transformées de Hilbert et de Riesz, cette grandeur peut également s'interpréter comme la « norme du déphasage du signal », et ainsi contribuer à l'analyse de sa phase locale. Nous proposons d'introduire une écriture de la forme « $s + j\mathcal{N}$ », à partir de laquelle l'amplitude et la phase pourront être identifiées par le module et l'argument complexes. Il y a cependant une difficulté, à savoir que notre signal s est désormais vectoriel, alors que notre « déphasage » $\mathcal{N}_R$ est scalaire. C'est pourquoi nous allons introduire dans la section suivante une nouvelle notion de phase couleur. Nous proposons dans un premier temps de nous restreindre aux signaux 1D.

4.2.2 Phase de signaux vectoriels

Nous proposons ici de construire et d'étudier l'amplitude et la phase instantanées de signaux 1D à valeurs vectorielles. Par simplicité, et puisque le but final concerne les images couleur, nous nous limiterons à des vecteurs de dimension 3.

Axe couleur et modèle oscillant

L'intérêt visé par l'extension du concept de phase à la couleur est une caractérisation des structures locales présentant une forte variation. C'est pourquoi pour commencer, on s'intéresse à un signal $I(x)$ présentant une rupture franche entre deux couleurs particulières I_1 et I_2 :

$$I(x) = I_1 + (I_2 - I_1)\mathbf{1}_{x>0}(x) \quad I_1, I_2 \in \{0, \dots, 255\}^3 \quad (4.13)$$

où $\mathbf{1}_{x>0}$ est l'échelon de Heaviside (Ce genre de signal 1D est analogue à un « bord » dans le monde 2D). Nous en illustrons un exemple à la figure 4.5a, dans le cas où I_1 est un vert et I_2 un bleu.

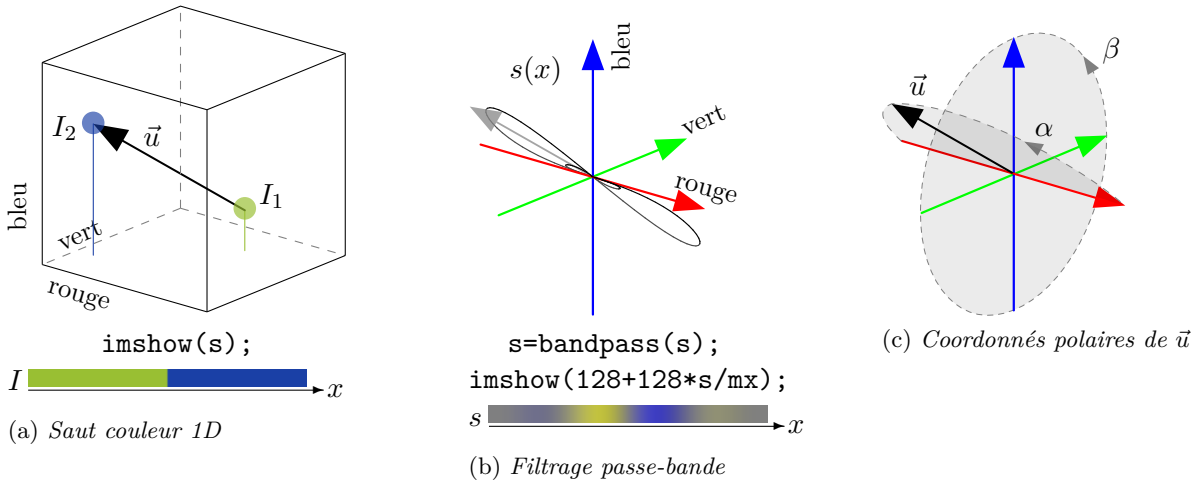


FIGURE 4.5 – Modèle oscillant de signal vectoriel 1D lié à une rupture entre deux couleurs.

Cette rupture intervient dans l'espace RGB le long d'un certain *axe couleur*, que nous appelons $\vec{u}$, défini par :

$$\vec{u} = \frac{I_2 - I_1}{\|I_2 - I_1\|} \quad (4.14)$$

Nous avons vu que l'analyse de la phase doit se faire dans une *sous-bande* du signal étudié. Nous allons donc utiliser un filtre passe-bande g qui n'altère pas la phase du signal, c'est-à-dire symétrique. Le filtrage linéaire d'un signal vectoriel se fait de façon marginale, ce qui produit

dans l'exemple de notre échelon le signal suivant :

$$s(x) = \begin{bmatrix} g * I^R \\ g * I^G \\ g * I^B \end{bmatrix} = \|I_2 - I_1\| (g * \mathbf{1}_{x>0})(x) \vec{u} \quad (4.15)$$

(La caractéristique passe-bande du filtre supprime la composante continue I_1). Ce signal est illustré à la figure 4.5b dans un repère 3D. Notons que la courbe décrite par s devrait théoriquement être confondue dans la droite générée par $\vec{u}$, mais que pour mieux la visualiser, nous avons artificiellement appliqué une légère variation de $\vec{u}$. La technique d'affichage d'une sous-bande que nous avons déjà introduite à la figure 4.3 a également été utilisée pour afficher cette oscillation sous forme d'une *image* (figure 4.5b, en bas).

On constate que le signal obtenu consiste en une oscillation *scalaire* « $g * \mathbf{1}$ », orientée dans l'espace couleur par l'axe $\vec{u}$ que nous avons introduit, et amplifiée par la distance entre les deux couleurs I_1 et I_2 . De plus, du fait de la symétrie du filtre et de la structure d'échelon du signal, on peut montrer que cette oscillation est *impaire*. Remplaçons désormais l'échelon $\mathbf{1}_{x>0}$ par une impulsion de Dirac : l'on obtient une structure paire, correspondant dans le monde 2D à une *ligne* de couleur I_2 sur un fond de couleur I_1 . Dans ce cas l'oscillation résultante, de la forme $g * \delta$, est elle aussi *paire*.

Finalement, la sous-bande couleur obtenue par un filtre symétrique appliqué marginalement sur un signal couleur 1D vérifie les propriétés suivantes :

- elle présente des oscillations locales au niveau des fortes variations de I ;
- les couleurs qui entrent en jeu dans ces variations déterminent :
 - l'amplitude des oscillations, par une mesure de *distance couleur* ;
 - l'axe couleur $\vec{u}$ dans l'espace RGB, le long duquel s évolue ;
- ces oscillations répliquent la *parité* des structures de I .

A partir de ces observations, et afin de définir un concept de phase couleur, nous proposons un *modèle* de sous-bande composé d'une oscillation *scalaire* multipliée par un axe couleur. La partie scalaire de ce signal peut être exprimée par son amplitude et sa phase instantanées grâce au signal analytique classique :

$$s(x) = \begin{bmatrix} s^R(x) \\ s^G(x) \\ s^B(x) \end{bmatrix} = A(x) \cos(\varphi(x)) \vec{u}(x) \quad (4.16)$$

L'amplitude $A(x)$ indique alors la *force de variation* du signal, et peut être identifiée dans l'exemple de notre échelon par $\|I_2 - I_1\|$, qui traduit directement une notion de « différence de couleur ». Une rupture entre des couleurs très différentes produira donc une oscillation ample.

En complément, l'axe couleur $\vec{u}(x)$ contient l'information colorimétrique de la variation. L'amplitude et l'axe étaient *constants* dans nos exemples d'échelon et d'impulsion, et sont désormais des fonctions de x . Néanmoins, afin que les données soient interprétable, il faudra que le signal ainsi modélisé ait une amplitude et un axe couleur dont les variations sont lentes devant la composante oscillatoire $\cos(\varphi(x))$.

La phase $\varphi(x)$, de par le formalisme classique, encode la structure locale de la partie scalaire. Elle contient donc l'information de parité de la rupture couleur. Afin de préciser cette nouvelle notion de phase, et de définir une méthode d'extraction des paramètres de ce modèle, nous proposons maintenant d'étudier la transformée de Hilbert d'une sous-bande couleur.

Transformée de Hilbert couleur

Nous avons vu à la section précédente que la transformée de Riesz d'un signal couleur devait être utilisée de façon marginale, pour extraire ensuite ses caractéristiques de façon non-linéaire. La norme du « signal déphasé » est alors définie par la norme euclidienne de la transformée

de Riesz. Nous proposons donc d'étudier la transformée de Hilbert *marginale* d'une sous-bande couleur 1D conforme au modèle de l'équation (4.16), ainsi que sa norme euclidienne. Notre modèle suppose que A et $\vec{u}$ soient localement *constant*, c'est-à-dire que leurs variations (au sens de chaque composante cartésienne pour $\vec{u}$) sont *lentes* devant $\cos(\varphi(x))$. On peut alors supposer que les spectres fréquentiels de $\cos(\varphi(x))$ et des composantes de $A\vec{u}$ ont un support distinct. Le théorème de Bedrosian (equation (1.7), page 17) assure donc que la transformée de Hilbert marginale de s peut être écrite de cette façon :

$$\mathcal{H}s(x) = \begin{bmatrix} \mathcal{H}s^R(x) \\ \mathcal{H}s^G(x) \\ \mathcal{H}s^B(x) \end{bmatrix} = A(x) \sin(\varphi(x)) \vec{u}(x) \quad (4.17)$$

On remarque que l'application d'une transformée de Hilbert marginale produit bien un *déphasage* de notre sous-bande couleur, puisque cette opération est équivalente à appliquer une transformée de Hilbert classique sur la partie scalaire de notre modèle. La sous-bande obtenue $\mathcal{H}s$ oscille donc selon le même axe couleur $\vec{u}(x)$.

L'objectif du signal analytique est l'*extraction* automatique des caractéristiques locales du modèle sous-jacent : l'amplitude A , la phase φ , et dans notre cas l'axe couleur $\vec{u}$. En ce qui concerne l'amplitude et l'axe, il est facile de remarquer que :

$$\sqrt{\|s(x)\|^2 + \|\mathcal{H}s(x)\|^2} = A(x) \quad (4.18)$$

$$s(x)/\|s(x)\| = \begin{cases} \vec{u}(x) & \text{si } \varphi(x) \in [-\frac{\pi}{2}; \frac{\pi}{2}] \\ -\vec{u}(x) & \text{si } \varphi(x) \in [-\pi; -\frac{\pi}{2}] \cup [\frac{\pi}{2}; \pi] \end{cases} \quad (4.19)$$

Ces deux caractéristiques peuvent donc être reconstruites simplement (au signe près pour $\vec{u}$) à partir du signal de départ et de sa transformée de Hilbert marginale, en prenant les *normes euclidiennes* de ces vecteurs. L'extraction de la phase $\varphi(x)$ n'est pas aussi immédiate, nous allons donc l'étudier maintenant.

Phase couleur

Nous avons vu que la transformée de Hilbert marginale d'une sous-bande couleur réalise bien un *déphasage*, selon le modèle d'oscillation couleur que nous avons défini à l'équation (4.16). Nous avons également vu que l'amplitude et l'axe instantanés du signal peuvent être facilement calculés à partir des normes euclidiennes $\|s(x)\|$ et $\|\mathcal{H}s(x)\|$. Nous proposons donc ici de définir la phase d'un signal couleur 1D à partir de ces normes euclidiennes. Cependant, l'utilisation de telles normes peut s'apparenter à l'utilisation de *valeurs absolues* dans le cas scalaire, ce qui peut suggérer *a priori* à une perte d'information. C'est pourquoi nous allons dans un premier temps étudier les conséquences d'une telle restriction dans le cadre du signal analytique classique.

Cas scalaire Reprenons le modèle classique $s(t) = A(t) \cos(\varphi(t))$, et remplaçons s et $\mathcal{H}s$ par leurs valeurs absolues. On obtient la ré-écriture suivante :

$$\begin{aligned} s &= \underbrace{\sqrt{s^2 + (\mathcal{H}s)^2}}_A \cos(\underbrace{\arg\{s + j\mathcal{H}s\}}_\varphi) \\ &= \underbrace{\sqrt{|s|^2 + |\mathcal{H}s|^2}}_A \cos(\underbrace{\arg\{|s| + j|\mathcal{H}s|\}}_{\varphi_2}) \underbrace{\frac{s}{|s|}}_{\text{signe}} \end{aligned} \quad (4.20)$$

Finalement, l'information perdue est simplement le *signe* du signal d'origine. Ce déplacement d'information se traduit par une simplification de la phase φ . Elle est l'argument d'un nombre complexe dont les parties réelle et imaginaire sont désormais *positives*, de par ces valeurs absolues, elle est donc restreinte à l'intervalle $[0; \frac{\pi}{2}]$. En termes d'interprétation de la phase comme

descripteur de la structure locale, les « montées » et les « descentes » sont donc confondues ($\varphi_2 \approx \pi/2$), ainsi que les « bosses » et les « creux » ($\varphi_2 \approx 0$), comme illustré à la figure 4.6. Cela signifie que la caractérisation du type de structure reste riche, puisque toutes les valeurs

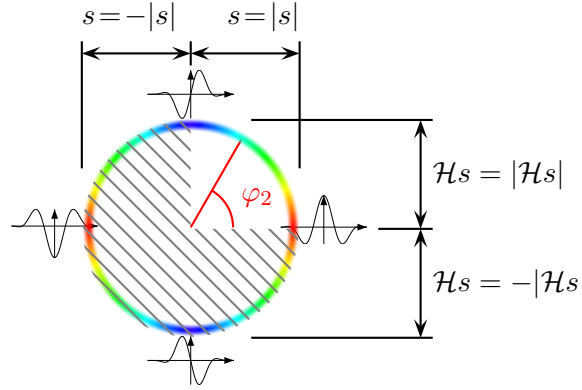


FIGURE 4.6 – Restriction de la phase à $[0; \pi/2[$, et interprétation correspondante.

intermédiaires entre le « profil pair » (bosse, creux) et le « profil impair » (montée, descente) sont conservées - comme nous en avons discuté autour de la figure 3.2, et page 78. Le signe du signal constitue donc une information « bosse/creux » complémentaire à la phase. En se projetant dans le modèle 2D correspondant, ce signe fera la différence entre les *lignes* claires sur fond foncé et les *lignes* foncées sur fond clair. En d'autres termes, il indiquera si la structure encodée dirige ses valeurs « vers le noir » ou bien « vers le blanc », ce qui préfigure tout-à-fait la notion d'*axe couleur* que nous avons introduite pour les signaux vectoriels.

Cas vectoriel L'utilisation de normes euclidiennes pour l'extraction de la phase de notre sous-bande couleur entraîne la formule conditionnelle suivante :

$$\varphi_2(x) = \arg\{\|s(x)\| + \mathbf{j}\|\mathcal{H}s(x)\|\} = \begin{cases} \varphi(x) & \text{si } \varphi(x) \in [0; \frac{\pi}{2}] \\ \pi - \varphi(x) & \text{si } \varphi(x) \in [\frac{\pi}{2}; \pi] \\ \pi + \varphi(x) & \text{si } \varphi(x) \in [-\pi; -\frac{\pi}{2}] \\ -\varphi(x) & \text{si } \varphi(x) \in [-\frac{\pi}{2}; 0] \end{cases} \quad (4.21)$$

En utilisant les équations (4.18) et (4.19), on peut remarquer que ces conditions sur le quadrant de φ sont toutes absorbées dans la ré-écriture suivante :

$$s = \underbrace{\sqrt{\|s\|^2 + \|\mathcal{H}s\|^2}}_A \cos(\underbrace{\arg\{\|s\| + \mathbf{j}\|\mathcal{H}s\|\}}_{\varphi_2}) \underbrace{\frac{s}{\|s\|}}_{\pm \vec{u}} = A \cos(\varphi) \vec{u} \quad (4.22)$$

Cela signifie que la méthode d'extraction des paramètres du modèle, bien que laissant une incertitude sur le quadrant de la phase et sur le signe de l'axe, permet de reconstruire parfaitement la sous-bande s . En d'autres termes, et comme nous venons de le voir dans le cas scalaire, la restriction de la phase imposée par l'emploi de normes euclidiennes n'a pas d'incidence sur l'interprétation des données. Concrètement, cette nouvelle phase sera toujours limitée à $[0; \pi/2]$, mais permettra de discriminer les lignes et les bords tout comme dans le cas scalaire. En compensation de cette simplification de la phase, l'axe couleur $\vec{u}$ sera à prendre en compte « au signe près ». Toutefois, la droite décrite par cette axe dans l'espace couleur sera la même quel que soit le signe, et permettra de caractériser le contexte colorimétrique de la structure locale, de la même façon.

Finalement, nous avons défini une méthode d'extraction des paramètres locaux d'amplitude, de phase et d'axe d'une sous-bande couleur, par les équations (4.18), (4.19) et (4.21), conforme à un modèle de signal oscillant dans une direction de $\mathbb{R}^3$ localement constante. L'évolution par

rapport au cas scalaire est que nous avons une nouvelle information d'axe couleur, qui caractérise l'aspect coloré de la structure analysée, alors que l'amplitude et la phase conservent la même interprétation de « force de variation » et de « parité de la structure ». Ces nouvelles données vont maintenant pouvoir être étendues aux images couleur grâce à un nouveau concept de signal monogène couleur.

4.2.3 Signal monogène couleur

Après avoir défini l'extraction de la norme et de la direction de la transformée de Riesz par le formalisme du gradient couleur, nous venons de voir que sous certaines conditions, l'amplitude et la phase locales d'un signal vectoriel s peuvent être extraites à partir des seules normes euclidiennes $\|s\|$ et $\|\mathcal{H}s\|$. Une information d'axe couleur vient alors compléter l'analyse. La quantité scalaire $\|\mathcal{H}s\|$ correspond à une version déphasée du signal privée de son signe, tout comme $|\mathcal{R}s|$ dans le cas du signal monogène en niveaux de gris. $\|\mathcal{H}s\|$ trouve naturellement son extension 2D dans la norme de la transformée de Riesz $\|\mathcal{R}s\| = \mathcal{N}_{\mathcal{R}}$, que nous avons définie comme norme du déphasage d'une sous-bande couleur. C'est pourquoi nous proposons la définition suivante de signal monogène couleur :

$$\begin{aligned}
 s_M^{\text{color}} &= [s^R \ s^G \ s^B \ \mathcal{N}_{\mathcal{R}}]^T \\
 \text{Amplitude :} \quad A &= \sqrt{\|s\|^2 + \mathcal{N}_{\mathcal{R}}^2} \in [0; +\infty[\\
 \text{Phase :} \quad \varphi_2 &= \arg\{\|s\| + j\mathcal{N}_{\mathcal{R}}\} \in [0; \frac{\pi}{2}[\\
 \text{Axe couleur :} \quad \begin{cases} \alpha = \arg\{s^R + j\sqrt{(s^G)^2 + (s^B)^2}\} \in [0; \pi[\\ \beta = \arg\{s^G + js^B\} \in [-\pi; \pi[\end{cases}
 \end{aligned} \tag{4.23}$$

Il s'agit d'un vecteur de dimension 4 construit par juxtaposition du signal d'origine avec sa version déphasée, constituée par la norme de Riesz couleur $\mathcal{N}_{\mathcal{R}}$ obtenue par l'équation (4.11). Ses coordonnées sphériques correspondent à l'amplitude, la phase et l'axe couleur, qui décrivent les structures locales du signal en termes directement interprétables. Notons que nous utilisons les coordonnées polaires de l'axe couleur α et β , définies par :

$$\vec{u} = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \cos(\beta) \\ \sin(\alpha) \sin(\beta) \end{bmatrix} \tag{4.24}$$

et illustrées sur la figure 4.5c.

Le modèle sous-jacent à cette définition est un signal oscillant dans $\mathbb{R}^3$ le long d'un axe localement constant (au signe près) :

$$s(\mathbf{x}) = A(\mathbf{x}) \cos(\varphi_2(\mathbf{x})) \vec{u}(\mathbf{x}) \tag{4.25}$$

Une illustration est donnée à la figure 4.7, nous proposons de la discuter.

Amplitude, phase et orientation L'interprétation de A et φ_2 est exactement la même que dans le cas scalaire du signal monogène de M. Felsberg, à l'exception du fait que la phase est « simplifiée » dans l'intervalle $[0; \pi/2]$. Notons que cette simplification était déjà utilisée au chapitre précédent dans les illustrations, pour améliorer la visualisation des lignes et des bords (voir notamment la figure 3.2). On constate que l'amplitude répond de manière équitable à tous les contours couleur, et que la phase indique clairement les lignes (en rouge) et les bords (en bleu).

L'orientation, qui n'est pas utilisée dans le modèle de signal mais que l'on peut adjoindre à l'analyse par l'équation (4.12), est identique à celle d'un gradient couleur, et correspond bien à la perception que nous avons des orientations des contours de l'image.

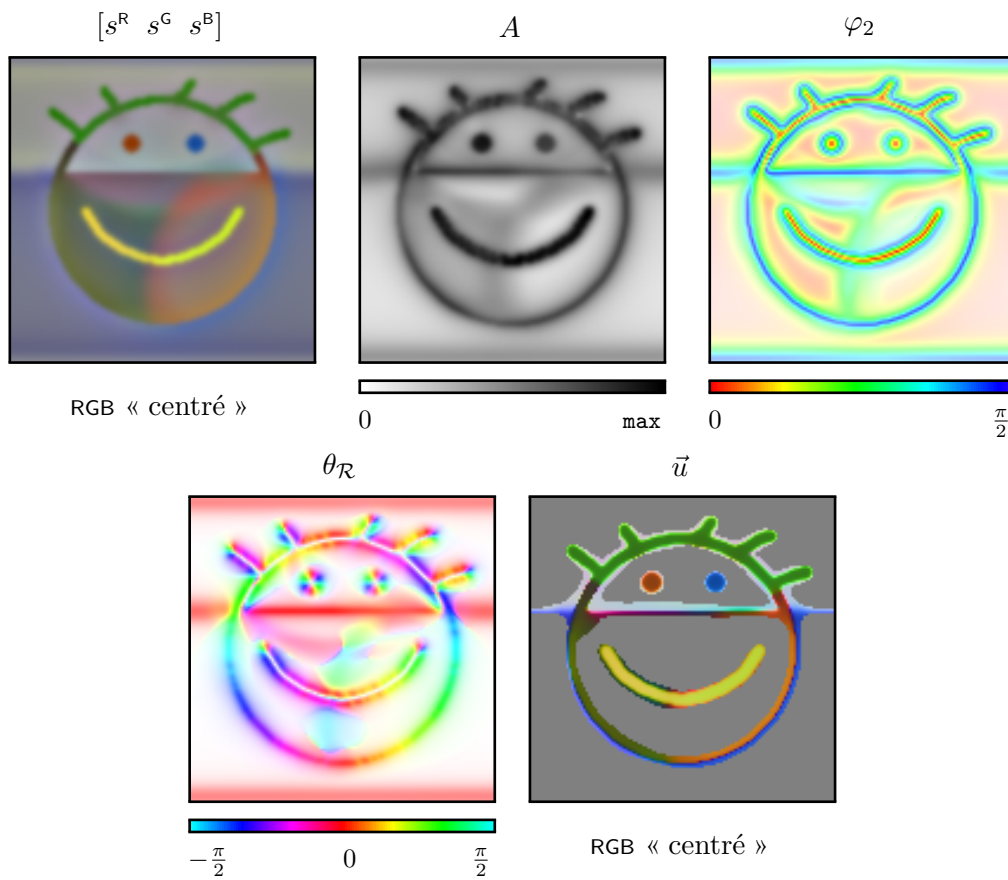
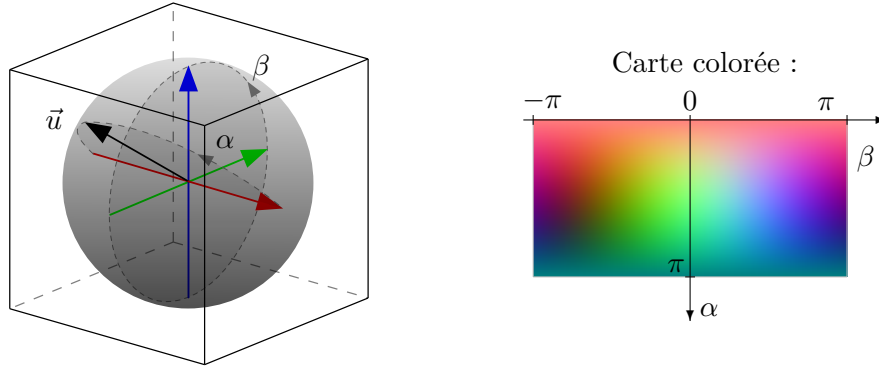


FIGURE 4.7 – *Signal monogène couleur proposé.* La saturation de φ_2 est contrôlée par A , celle de $\theta_{\mathcal{R}}$ par $\mathcal{N}_{\mathcal{R}}$. Les valeurs de $\vec{u}$ sont illustrées par la carte colorée décrite à la figure 4.8. Celles dont l'amplitude correspondante A est faible sont remplacées par un pixel gris.

FIGURE 4.8 – Illustration des valeurs de l'axe $\vec{u}$ par intégration de la sphère dans le cube RGB.

Axe couleur L'axe couleur $\vec{u}$ enfin, pose tout d'abord une difficulté de représentation. En effet, il s'agit d'une donnée « circulaire 2D », et les variables α et β correspondant aux coordonnées surfaciques d'un point d'une sphère. Pour les illustrations, nous avons choisi de faire correspondre une couleur unique à chaque vecteur $\vec{u}$ possible. Pour cela, une solution très simple est d'inclure la sphère décrite par $\vec{u}$ dans le cube RGB comme illustré à la figure 4.8. Cette technique ressemble à la façon dont nous avons affiché les sous-bandes couleur, mais se réduit cette fois à un sous-ensemble 2D du cube RGB, constituant la « carte » des couleurs qui encoderont donc les valeurs de $\vec{u}$ (figure 4.8, droite). Sur l'illustration du signal monogène, on constate que $\vec{u}$ ressemble fortement à s . Cela s'explique par la technique d'affichage qui est semblable. Cependant, les données de l'axe couleur sont localement constantes au signe près, alors que la sous-bande s varie rapidement. L'axe couleur indique une direction dans l'espace couleur 3D, le long de laquelle l'oscillation de s a lieu. On constate qu'au niveau des lignes, cet axe est univoque et constant, ce qui confirme l'hypothèse de « variation lente » nécessaire à la cohérence du modèle. En revanche, le comportement de l'axe couleur au niveau des bords est moins intuitif. La raison est que le filtrage passe-bande de l'image est isotrope et pair, ce qui produit au niveau des bords des *passages par zero*, qui rendent l'extraction de l'axe $s/\|s\|$ instable¹.

Images en niveaux de gris Il faut noter que cette nouvelle définition inclut le formalisme monogène de M. Felsberg comme un cas particulier. En effet, considérons un signal scalaire $f : \mathbb{R}^2 \mapsto \mathbb{R}$. On appelle alors A_f , φ_f et θ_f les paramètres monogènes classiques obtenus par les équations (3.45) à partir de f . En définissant le signal couleur $s = [f \ f \ f]^T$, on peut montrer que l'analyse monogène couleur proposée aboutit à la correspondance suivante :

$$A = \sqrt{3}A_f \quad (4.26)$$

$$\varphi_2 = \begin{cases} \varphi_f & \text{si } \varphi_f \in [0; \pi/2] \\ \pi - \varphi_f & \text{si } \varphi_f \in [\pi/2; \pi] \end{cases} \quad (4.27)$$

$$\theta = \theta_f \quad (4.28)$$

$$\vec{u} = \text{sgn}(f) \frac{1}{\sqrt{3}} [1 \ 1 \ 1]^T \quad (4.29)$$

L'analyse produite est donc bien cohérente avec le signal monogène classique. En ce qui concerne les dimensions supérieures, cette définition peut être facilement généralisée aux n -vecteurs et images multispectrales, par simple augmentation de la dimension de $\vec{u}$.

Finalement, grâce à l'utilisation du formalisme du gradient couleur, nous avons proposé une définition de signal monogène couleur qui préserve l'interprétation du signal monogène classique - amplitude, phase, orientation - en prenant en compte tous les contours, et sans

1. Une perspective intéressante pour nos travaux futurs est peut-être dans la considération conjointe de l'axe couleur lié à la transformée de Riesz, dont les passages par zero seront eux au niveau des lignes.

privilégier de couleur particulière. Aucun paramètre n'est à fixer pour conduire l'analyse. Une nouvelle information d'axe couleur local est apparue, et permet de séparer l'information couleur de l'information géométrique dans la caractérisation des structures locales. Conformément au signal analytique/monogène scalaire, on peut reconstruire le signal de départ à travers un modèle AM/FM paramétré par l'amplitude, la phase et l'axe couleur. La construction repose principalement sur l'utilisation de normes euclidiennes, ce qui est théoriquement simple, mais aboutit à une analyse indépendante de la couleur, là où les méthodes basées sur l'*intensité* échouent.

Avant d'étendre ce nouveau formalisme aux ondelettes, et pour terminer cette section, nous présentons une synthèse du travail très récemment présenté dans l'article [93], car il est dans le même esprit, bien que limité à des vecteurs de dimension 2.

Le signal quaternionique 1D

Une extension des notions d'amplitude et de phase pour des signaux de $\mathbb{R}$ dans $\mathbb{C}$ est proposée dans [93]. La méthode de construction est l'« annulation des fréquences négatives », qui est effectuée dans le spectre de la transformée de Fourier quaternionique 1D [93]. Même si cela n'est pas mis en valeur dans l'article, nous avons pu remarquer que cette définition résulte en la juxtaposition du signal d'origine avec sa transformée de Hilbert *marginale* :

$$s(t) = a(t) + \mathbf{j}_1 b(t) \quad \rightarrow \quad s_A(t) = a(t) + \mathbf{j}_1 b(t) + \mathbf{j}_2 \mathcal{H}a(t) + \mathbf{j}_3 \mathcal{H}b(t) \in \mathbb{H} \quad (4.30)$$

Posons $\mathcal{N} = \sqrt{(\mathcal{H}a)^2 + (\mathcal{H}b)^2}$. Une notion d'amplitude est obtenue par la représentation polaire de « Cayley-Dickson » d'un quaternion, et résulte en une enveloppe complexe définie par $(a + \mathbf{j}_1 b)/\cos(\mathcal{N})$. Ensuite, une notion de phase est définie par la représentation polaire classique (Euler) d'un quaternion, et on obtient $\varphi = \arg\{a + \mathbf{j}\sqrt{b^2 + \mathcal{N}^2}\}$. Notons que cette définition coïncide au signe près à la définition classique dans le cas réel $b(t) = 0$. Cette phase scalaire doit s'interpréter dans une certaine « direction » - dans le plan complexe - traduite par une notion d'*axe* défini par $(\mathbf{j}_1 b + \mathbf{j}_2 \mathcal{H}a + \mathbf{j}_3 \mathcal{H}b)/\sqrt{b^2 + \mathcal{N}^2}$.

Les informations ainsi extraites semblent correspondre expérimentalement à ce que l'on attend : une enveloppe lisse qui suit l'évolution de l'amplitude de l'oscillation, tout en suivant son orientation dans le plan complexe ; une fréquence instantanée (par dérivation de la phase) validée par comparaison avec un signal synthétique. Cependant, nous n'avons pas trouvé d'explication intuitive pour ces définitions. Le modèle nécessite deux représentations polaires différentes, et la façon de combiner les composantes du signal avec leurs transformées de Hilbert n'est pas évidente. De plus, à cause de l'algèbre des quaternions, une extension à des signaux 3-vectoriels semble impossible.

Maintenant que nous avons défini un signal monogène couleur plus propice à une analyse géométrique des structures couleur locales, nous proposons de l'utiliser dans un contexte multi-résolution.

4.3 ONDELETTES MONOGÈNES COULEUR

4.3.1 Définition

La définition d'une transformée en ondelettes monogènes couleur ne pose pas de difficulté car la construction du signal monogène couleur repose sur une analyse marginale, dont les composantes sont recombinaées de façon non-marginale pour faire apparaître les quantités interprétables (amplitude, phase, orientation, axe). En d'autres termes, la décomposition en ondelettes monogènes couleur nécessite dans un premier temps une décomposition monogène classique *marginale*, pour ensuite recombinaer les coefficients obtenus.

Nous avons donc choisi d'utiliser à nouveau les ondelettes splines polyharmoniques présentées à la section 3.3.2. A l'exception de la transformée proposée dans [69] qui est plus récente, les splines polyharmoniques de [144] sont les seules décompositions de la littérature capables de produire des sous-bandes monogènes.

Voici donc la procédure de décomposition. Il faut d'abord calculer les 6 décompositions suivantes :

$$c_{i,\mathbf{k}}^c = \langle s^c, \psi_{i,\mathbf{k}} \rangle = (\psi_i * s^c)(2^{-(i+1)}\mathbf{k}) \quad (4.31)$$

$$d_{i,\mathbf{k}}^c = \langle s^c, \psi_{i,\mathbf{k}}^{\mathcal{R}} \rangle = \{\mathcal{R}(\psi_i * s^c)\}(2^{-(i+1)}\mathbf{k}) \quad (4.32)$$

avec $c \in \{R, G, B\}$

puis calculer la norme de Riesz couleur multi-échelle :

$$\mathcal{N}_{i,\mathbf{k}} = \sqrt{|d_{i,\mathbf{k}}^R|^2 + |d_{i,\mathbf{k}}^G|^2 + |d_{i,\mathbf{k}}^B|^2} \quad (4.33)$$

Le signal monogène proposé est pour chaque échelle i :

$$c_M(\mathbf{k}) = \begin{bmatrix} c_{i,\mathbf{k}}^R & c_{i,\mathbf{k}}^G & c_{i,\mathbf{k}}^B & \mathcal{N}_{i,\mathbf{k}} \end{bmatrix}^T \rightarrow \begin{bmatrix} A_{i,\mathbf{k}} & \varphi_{i,\mathbf{k}} & \theta_{i,\mathbf{k}} & \vec{u}_{i,\mathbf{k}} \end{bmatrix}^T \quad (4.34)$$

L'amplitude, la phase et l'axe couleur doivent finalement être calculés par les équations (4.23), et l'orientation par l'équation (4.12). Notons que le coût algorithmique pour les 6 décompositions reste linéaire par rapport à la taille des données.

Nous illustrons la transformée en ondelettes monogènes couleur aux figures 4.9 et 4.10. On constate à nouveau que la phase évolue trop vite par rapport à l'échantillonnage pour pouvoir être visualisée correctement. Cela est associé à un codage compact de l'information, qui contribue à la propriété de quasi-invariance par translation/rotation de l'amplitude.

Cette nouvelle transformée est la première représentation spécialement définie pour les images couleur, offrant une analyse riche et interprétable. D'abord, le bancs de filtres est non-séparable, ce qui signifie que l'on n'a plus de classification arbitraire des détails horizontaux, verticaux, et diagonaux, contrairement aux schémas classiques, mais une sous-bande unique pour chaque échelle d'analyse. En complément, chaque coefficient se voit attribuée une mesure précise d'orientation locale, obtenue implicitement à travers le formalisme du gradient couleur. Ensuite, pour chaque échelle, la sous-bande est à valeurs vectorielles, dont les composantes traduisent des caractéristiques locales d'énergie, de géométrie et de colorimétrie, directement interprétables en termes d'amplitude des contours, de type de contour (bord/ligne), de couleurs impliquées dans ces contours. Notons que la méthode de construction non-marginale de cet outil aboutit à une prise en compte efficace des contours « isoluminants », qui disparaîtraient dans un schéma basé sur l'intensité. L'analyse est approximativement invariante par translation et rotation, ce qui la place dans la catégorie des représentations *cohérentes*, très utiles dans des contextes de reconnaissance de formes. L'information multi-échelle de cette transformée peut potentiellement être utilisée dans des schémas d'analyse nécessitant par exemple la caractérisation de points particuliers. La redondance de la représentation correspond à la quantité de données nécessaire pour reconstruire l'image divisée par la quantité de données initiale. Les quatre bancs de filtres de type pyramides obtenus pour trois canaux couleur aboutissent à une redondance de $4\frac{4}{3}/3 \approx 1.8$, ce qui est très raisonnable (inférieure à celle du Dual-Tree 1D), et permet de qualifier la représentation de *compacte*. L'analyse sous-jacente de cette transformée est conforme à une nouvelle généralisation du signal anaytique, spécialement défini pour les images couleur, mais cette fois, toutes les bandes fréquentielles sont analysées, afin de prendre en compte tout type de signal couleur.

Afin de faire de cette transformée un outil « complet » vis-à-vis de la problématique de cette thèse - définir une représentation offrant une analyse riche et une reconstruction cohérente - il faut définir une méthode de reconstruction de l'image couleur à partir des coefficients d'ondelettes monogènes.

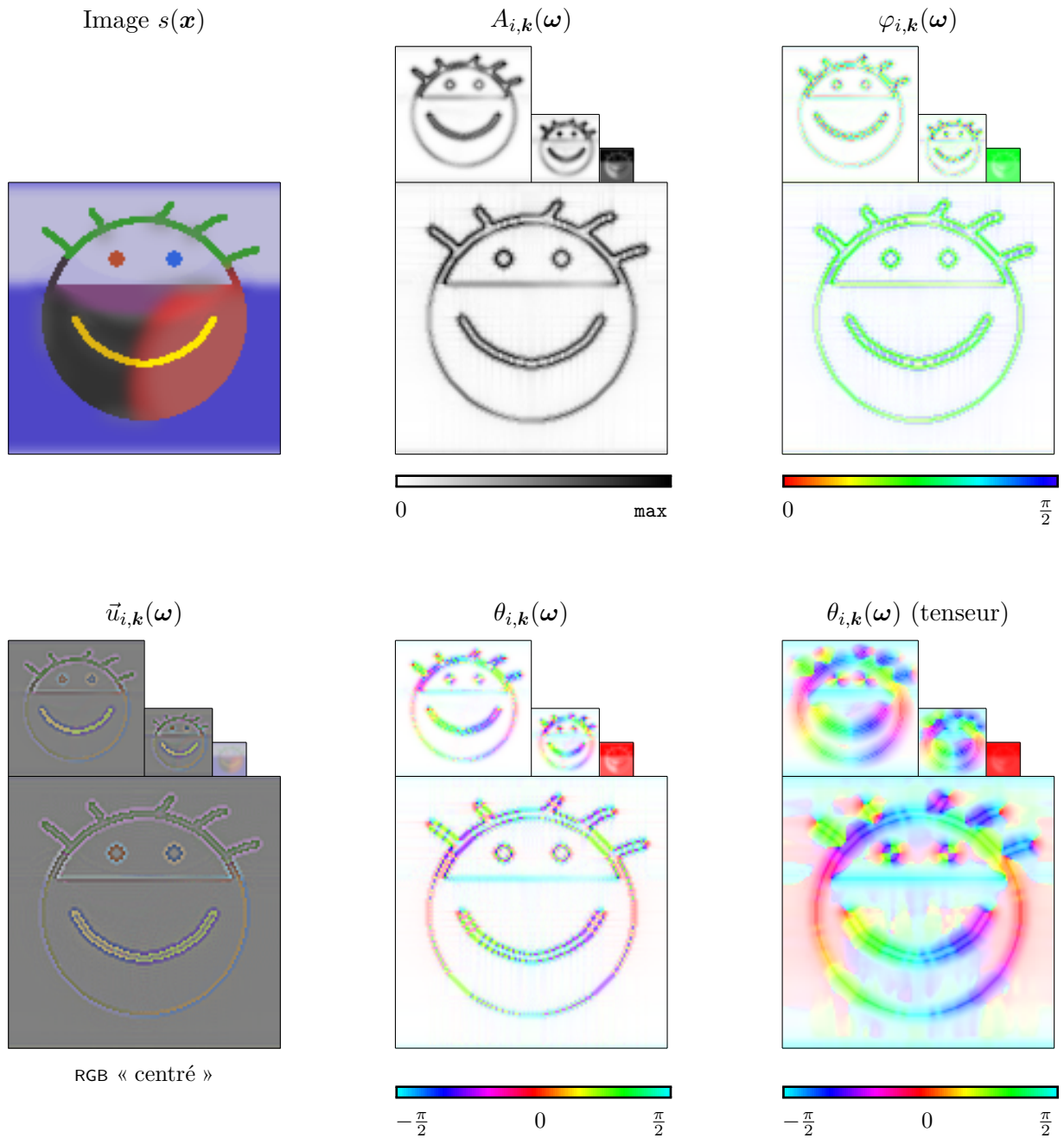


FIGURE 4.9 – Transformée en ondelettes monogène couleur proposée, sur une image synthétique ($\gamma = 3$, $i \in \{-1, -2, -3\}$). L'orientation améliorée par lissage du tenseur est également montrée.

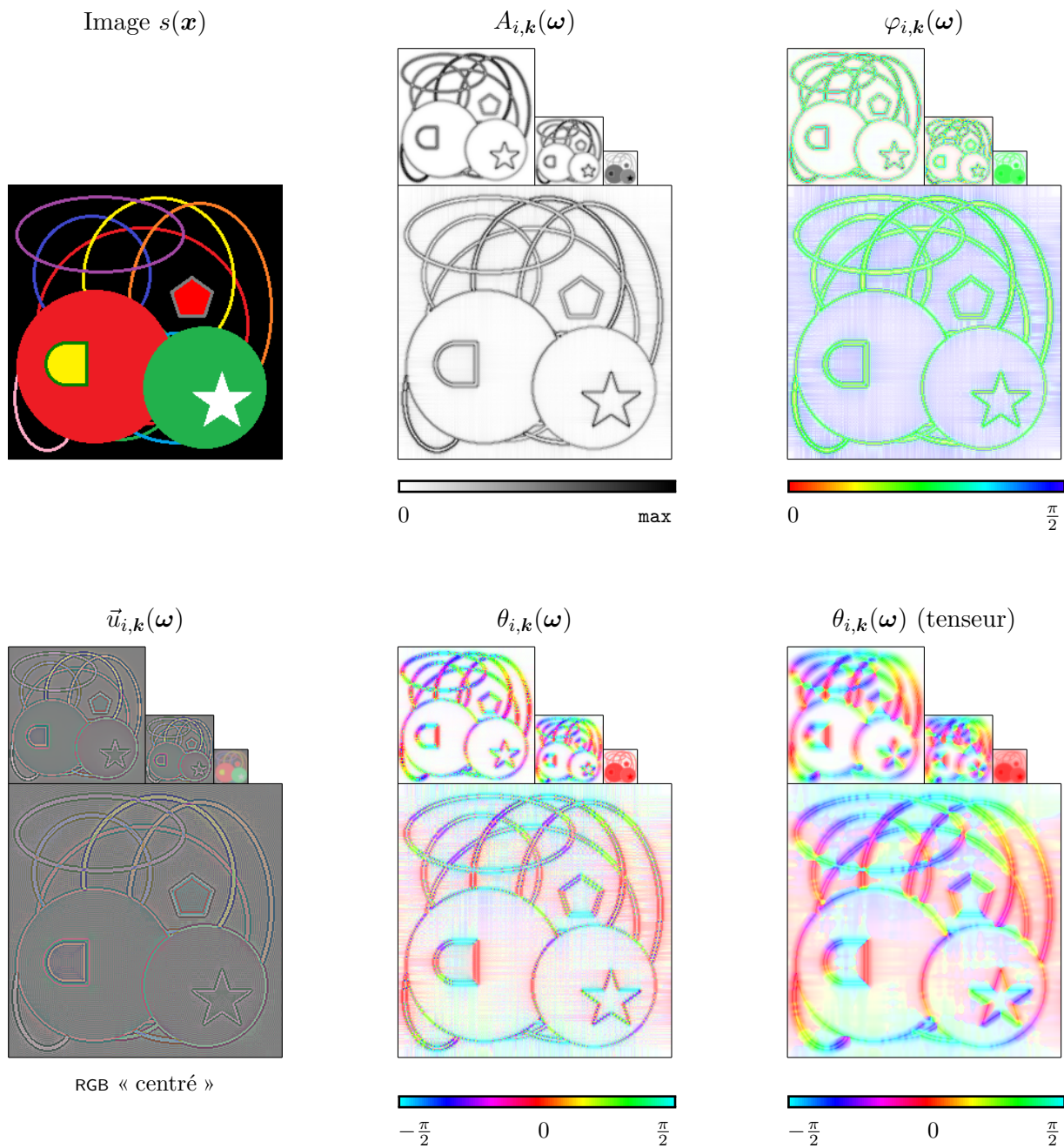


FIGURE 4.10 – Voir figure 4.9.

4.3.2 Reconstruction

Nous avons vu au chapitre précédent que la problématique de reconstruction du signal de départ dans le formalisme du signal monogène n'a pas été discutée dans la littérature. Depuis la proposition de M. Felsberg, jusqu'aux applications de la transformée de M. Unser, tous les travaux sont tournés vers l'*analyse* d'image. Cette dimension est pourtant capitale dans le champ de recherche sur les ondelettes, dont le succès en image repose en grande partie sur le débruitage rapide et la compression.

L'extension couleur donnée dans ce chapitre a pour point de départ l'analyse des contours par gradient, et pour composantes de base les outils scalaires liés au signal monogène d'une part et aux ondelettes splines polyharmoniques d'autre part. Du fait de cette dépendance aux outils susdits, cette extension hérite de cette ambiguïté sur la reconstruction, de l'approximation sur l'isotropie et l'invariance, et des choix qui ont été faits sur la forme des filtres d'ondelette. La notion de fusion des différentes reconstructions possibles, dans une perspective de façonner les atomes de base en fonction des paramètres monogènes, n'est pas donc pas définie *a priori*. Deux stratégies de fusion des informations sont néanmoins envisagées.

La première est naturellement dans la définition des bancs de filtres de splines polyharmoniques. Chaque banc de filtre possède un algorithme de reconstruction parfaite, ce qui permet dans notre cas de reconstruire l'image couleur de départ ainsi que sa transformée de Riesz marginale, à partir des composantes *cartésiennes* des coefficients monogènes $c_{i,k}^c$ et $d_{i,k}^c$. Cependant, les coefficients qui nous intéressent sont plutôt les composantes *polaires*, qui donnent l'analyse des structures de l'image. Ensuite, reconstruire la transformée de Riesz n'a pas d'intérêt *a priori*, on souhaiterait plutôt combiner les deux reconstructions vers une image reformée en fonction de l'ensemble des coefficients monogènes. Enfin, la définition du signal monogène couleur proposée ici ne permet pas de reconstruire les coefficients de la « partie Riesz » $d_{i,k}^c$, dont on n'a gardé que la norme $\mathcal{N}_{i,k}$ ².

Le second aspect est celui du modèle AM/FM, qui indique que le signal de départ peut être reconstruit par modulation d'une sinusoïde colorée : $s(\mathbf{x}) = A(\mathbf{x}) \cos(\varphi_2(\mathbf{x})) \vec{u}(\mathbf{x})$. Nous avons déjà soulevé dans le cas scalaire que l'orientation n'était pas utilisée, ce qui élimine l'idée d'orientation automatique des atomes de reconstruction.

La question de la reconstruction par représentation monogène fait donc partie des perspectives importantes pour nos travaux futurs. A ce stade de notre recherche, nous proposons une première méthode. Il s'agit d'utiliser le modèle AM/FM au niveau des *coefficients*, afin de reconstruire les *sous-bandes* de la partie primaire :

$$\begin{bmatrix} c_{i,k}^R \\ c_{i,k}^G \\ c_{i,k}^B \end{bmatrix} = A_{i,k} \cos(\varphi_{i,k}) \vec{u}_{i,k} \quad (4.35)$$

Dans un second temps, l'algorithme de synthèse des trois bancs de filtres primaires sera utilisée pour reconstruire l'image de façon unique. L'avantage de cette méthode et qu'elle concilie l'unicité de la reconstruction avec l'utilisation des paramètres A , φ et $\vec{u}$. Notons que, bien que la partie Riesz ne soit pas utilisée pour reconstruire l'image, elle est tout-de-même indispensable pour générer A et φ .

Finalement, grâce à une nouvelle extension du signal analytique, nous venons de définir une transformée en ondelettes pour les images couleur, dont les coefficients forment une analyse riche et permettent à la fois de reconstruire l'image de façon unique à partir de ces informations d'analyse - ce qui était le but de notre travail. Notons que très récemment, un article proposant une phase pour des signaux vectoriels 1D est apparu [98], ce qui montre le dynamisme de ce thème de recherche. L'étude de cet article fait partie des perspectives de cette thèse.

2. On peut retrouver la transformée de Hilbert dans la partie imaginaire du signal analytique.

Afin de progresser à travers une analyse critique de ce nouvel outil, nous proposons une étude expérimentale.

4.4 MISE EN PRATIQUE

Dans cette section, nous allons étudier la mise en pratique de la représentation en ondelettes monogènes couleur proposée. Nous avons choisi de mettre cette nouvelle représentation à l'épreuve dans deux applications typiques des transformées en ondelettes : la représentation parcimonieuse des signaux, et l'extraction invariante de points particuliers, qui éprouvera la qualité d'analyse des coefficients. L'aspect parcimonie constitue un travail préalable pour les problématiques de *codage hiérarchisé* qui nous intéressent dans le cadre du projet ANR « CAIMAN », auquel est associé cette thèse. Pour finir, nous ouvrirons le travail sur la problématique de la définition de schémas numériques associés aux représentations monogènes, qui constitue pour nous une perspective importante.

4.4.1 Représentation parcimonieuse des images couleur

La parcimonie d'une représentation correspond à la puissance descriptive de ses coefficients, associée à une forte concentration de l'information utile dans un petit nombre d'entre eux. Les transformées en ondelettes sont reconnues pour offrir une représentation compacte des images naturelles, c'est-à-dire que quelques coefficients de forte amplitude portent l'essentiel du contenu visuel. Suite à l'essor des ondelettes orthogonales (non-redondantes), qui se sont finalement montrées insuffisantes en termes de qualité d'analyse (variance) ainsi que de débruitage (oscillations), il a été remarqué que les représentations *redondantes* offraient des propriétés nécessaires pour certaines applications. L'on s'est donc intéressé à des représentations redondantes dans lesquelles l'information est la plus compacte possible. Par exemple, les ondelettes non-décimées (algorithme à trous) sont invariantes par translation, et donnent accès à une méthode de débruitage plus efficace [31, 71, 138]. L'utilisation plus générale de dictionnaires redondants pour le codage d'image donne dans certains cas de meilleures reconstructions en termes de qualité visuelle, pour une même quantité de données et dans un contexte bas-débit [50]. Cela inclut les ondelettes complexes [153], ainsi que les représentations en ondelettes *analytiques* que nous étudions, comme nous l'avons déjà montré avec les ondelettes quaternioniques (section 2.4). Ces dernières ont une redondance de $4\times$, et permettent de traiter les contours de façon plus lisse et plus ronde que les ondelettes orthogonales. Il vient que la phase des coefficients peut être quantifiée dûment sans introduire d'artéfacts significatifs.

Nous attendons des ondelettes monogènes couleur des résultats similaires, puisqu'elles ont été construites sur une extension couleur du signal monogène, lui-même évolution du signal quaternionique. En effet, le sens physique que nous donnons aux coefficients confère une grande puissance de description des formes géométriques couleur, avec peu de nombres. Cet outil est donc un bon candidat à la parcimonie.

Nous proposons ici de tester la faculté à résumer l'information par des reconstructions partielles. Plus précisément, nous allons conduire une procédure de sélection et quantification des coefficients, conjointement à une analyse des artéfacts correspondants qui auront été introduits dans l'image reconstruite. En plus d'analyser la compression de l'information, cela nous permettra d'identifier les éléments graphiques liés aux différentes composantes de notre représentation.

Stratégie de quantification

La quantification des coefficients d'ondelettes est classique avec les ondelettes réelles, et se base sur un modèle statistique pour optimiser cette procédure. Dans notre cas, les données sont inhabituelles, car il s'agit d'une norme euclidienne (A) d'une part, et d'*angles* d'autre part (φ_2 , α , β). Par exemple, pour une distribution gaussienne des composantes cartésiennes (qui est

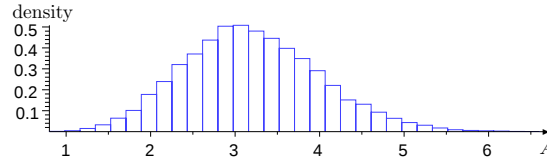


FIGURE 4.11 – Distribution « type Rayleigh » de l'amplitude A_i ($\gamma = 3$).

souvent l'hypothèse avec les ondelettes réelles), notre amplitude A suivra une distribution de Rayleigh, comme illustrée à la figure 4.11. Pour les angles, des outils statistiques existent (lois de von Mises), mais ne sont pas très connus dans le domaine du codage par ondelettes (voir aussi la discussion que nous avons déjà eue à la section 2.4.2, page 50, sur la quantification de la phase quaternionique). On comprend vite que l'ensemble des outils habituels pour quantifier des coefficients d'ondelettes ne peut pas s'appliquer dans notre cas.

Dans le cadre de ce travail expérimental, le but n'est pas de construire un schéma optimal, qui nécessite un travail dédié, mais bien de *tester* la transformée. Étant donnée la portée de cette application, nous pensons que la simplicité de la méthode est primordiale, d'une part pour ne pas noyer la contribution de la transformée en elle-même, et d'autre part car pour avoir une approche compression complète, l'outil proposé nécessite une réflexion approfondie, notamment pour modéliser statistiquement les coefficients. C'est pourquoi nous utiliserons une quantification uniforme de chaque composante A , φ_2 , α et β . Cette méthode inclut de façon sous-jacente une *sélection* des coefficients. En effet, les coefficients dont l'amplitude est dans le premier intervalle de quantification associé, verront leur amplitude mise à zéro, ce qui éliminera automatiquement les coordonnées polaires correspondantes puisque $A_{i,k} = c_{i,k}^R = c_{i,k}^G = c_{i,k}^B = 0$. Cela correspond donc à éliminer des coefficients « complets » - incluant l'amplitude, la phase et l'axe couleur.

Nous proposons de faire varier la précision de quantification pour les différentes composantes des coefficients, pour évaluer le nombre de bits nécessaires pour encoder de manière suffisante les coefficients monogènes. Ensuite, nous pouvons identifier les caractéristiques visuelles des défauts introduits selon la composante dégradée.

On utilisera 4 nombres de bits fixés pour toutes les échelles $(B_A, B_{\varphi_2}, B_\alpha, B_\beta)$, chacun correspondant à une composante. Le niveau de décomposition est 4 ($i \in \{-1, -2, -3, -4\}$), et la sous-bande basses-fréquences ne sera pas altérée. Ce dernier aspect est justifié en théorie car l'analyse monogène n'a pas de sens physique en basses-fréquences. Pour le point de vue pratique, la quantité de données concernée par cette sous-bande est très réduite (16×16 pour une image 256×256). De plus, puisque l'analyse monogène n'est pas à propos, on peut considérer les trois coefficients primaires $c_{i,k}^R$, $c_{i,k}^G$ et $c_{i,k}^B$, à la place des quatre coefficients monogènes $A_{i,k}$, $\varphi_{i,k}$, $\alpha_{i,k}$ et $\beta_{i,k}$, ce qui réduit encore la taille des données.

Analyse des résultats

La figure 4.12 illustre les reconstructions d'images couleur pour différentes configurations de $(B_A, B_{\varphi_2}, B_\alpha, B_\beta)$.

Dans la première colonne, il apparaît qu'une quantification de A (resp. φ , α , β) sur 6 (resp. 2, 3, 4) bits n'introduit pas de défaut visible. En ce qui concerne les angles φ , α et β , cette configuration correspond à des intervalles de quantification de $\pi/8$, ce qui peut paraître assez approximatif au regard de la qualité de reconstruction. Cela montre que l'information utile a bien été répartie à travers les quatre composantes des coefficients, et à travers la redondance spatiale des pyramides utilisées. En effet, si l'on comptabilise la redondance de $16:9$ de notre représentation, ainsi que le nombre total de $6 + 2 + 3 + 4 = 15$ bits utilisés ici pour coder les coefficients, on obtient un nombre moyen de bits par pixel d'environ 27, ce qui est de l'ordre de grandeur des 24 bits par pixel de départ, dans le format brut de l'image codée en RGB (8 bits par canal couleur). La cohérence de la représentation est donc dans cette première étape confirmée par la stabilité du contenu visuel vis-à-vis de la quantification.

Voyons l'effet d'une plus forte dégradation de l'amplitude A (deuxième colonne, figure 4.12).

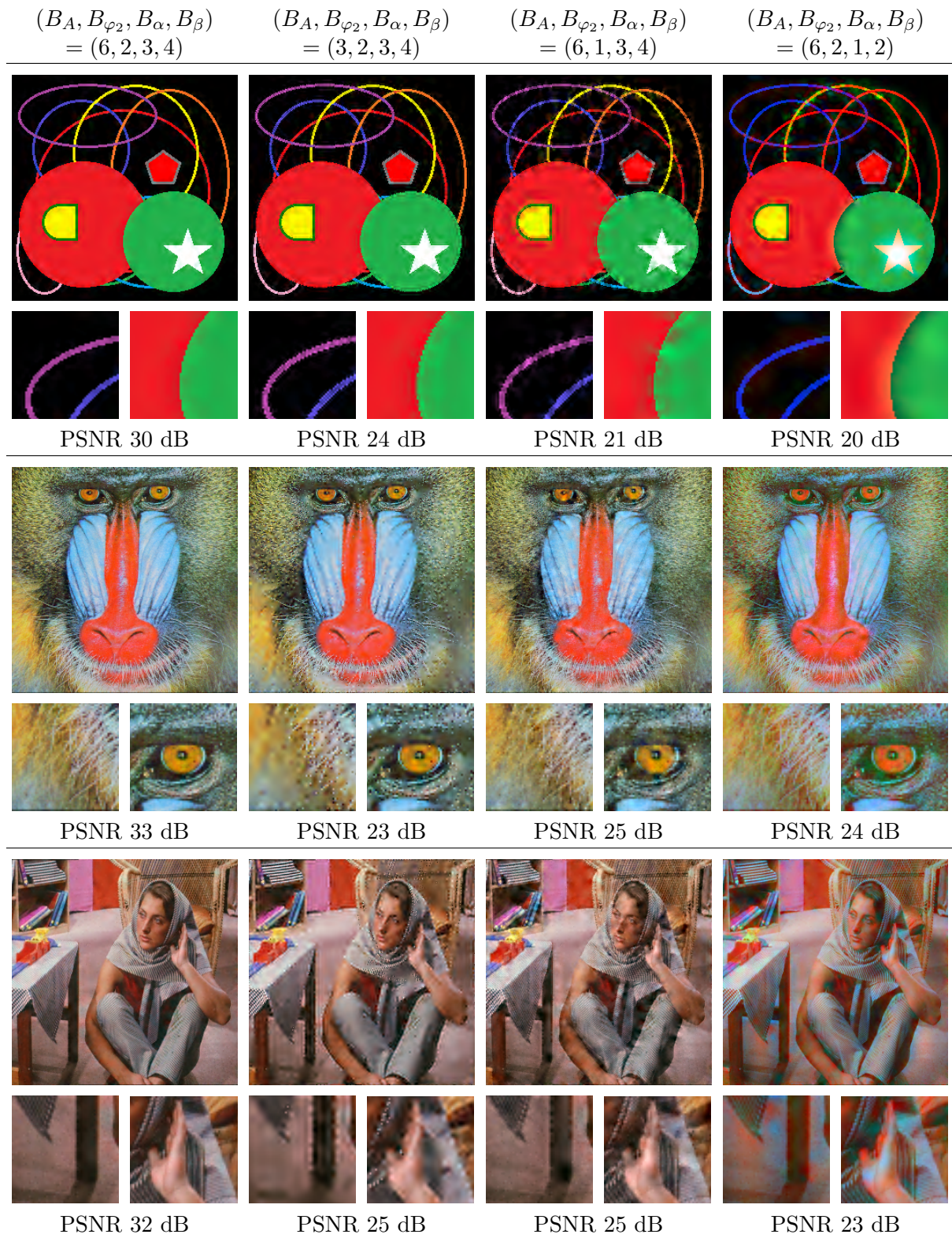


FIGURE 4.12 – Quantification uniforme des coefficients d'ondelettes monogènes couleur pour les images de test *circles*, *mandrill* et *barbara*.

Les bords de l'image **circles** subissent l'effet *pseudo-Gibbs*, qui fait apparaître des oscillations. De plus, certaines zones texturées sont lissées, comme les coins inférieurs de l'image **mandrill**, ou les vêtements de l'image **barbara**. Ces deux phénomènes sont classiques avec les transformées en ondelettes réelles, ce qui confirme que l'amplitude des ondelettes monogènes couleur est une donnée de même nature que le « module » dans le cas classique. Cependant, cette production d'oscillations n'est pas en accord avec le concept d'ondelettes analytiques, dont l'invariance est théoriquement liée à une réduction de ce phénomène [124]. Ce défaut est lié à la dimension numérique, et notamment à la méthode de *reconstruction*, qui diffère du Dual-Tree et de la QWT en ce qu'elle n'implique pas de combinaison d'atomes « pairs et impairs ». De plus, nous avons déjà indiqué que les filtres de reconstruction des ondelettes splines polyharmoniques sont très oscillants et très peu isotropes, à cause des contraintes de reconstruction parfaites dont les seuls degrés de libertés ont été utilisés pour affiner les filtres d'*analyse*. Ceci confirme qu'une réflexion sur la méthode numérique de décomposition et reconstruction devrait être menée de façon plus approfondie. Notons que malgré le caractère non-linéaire de cette opération, aucune fausse couleur n'a été introduite.

Pour observer l'effet d'une forte dégradation de la phase φ_2 , nous avons choisi de l'arrondir sur les deux valeurs 0 et $\pi/2$ (troisième colonne, figure 4.12). Cette quantification sur un seul bit de la phase introduit un phénomène basse-fréquences semblable à celui d'un papier mouillé, que l'on observe par exemple autour des bords des cercles de **circles** et du pied de table de **barbara**. Ce phénomène semble provenir des contours mais se répand largement sur toute l'image. La notion de « congruence de phase » permet de comprendre ce problème. La congruence de phase a été introduite comme détecteur de contours [108], sous l'hypothèse selon laquelle les contours fortement perçus par le système visuel humain correspondent aux maxima de congruence de phase. La particularité de ces points est donc que l'ensemble des sinusoides qui le composent, à différentes fréquences, y ont la *même phase*. Il s'est avéré que le signal analytique permettait de calculer cette mesure par l'amplitude locale, tout en donnant explicitement cette valeur de phase φ_2 au niveau des contours. Dans une décomposition en ondelettes analytique comme la notre, aux positions des maxima d'amplitude, c'est-à-dire des contours, la valeur de φ_2 doit donc être la même *d'une échelle à l'autre*. Cependant, s'il s'avère que cette congruence n'est pas parfaite, les phases ne seront pas identiques. Dans ce cas, la procédure de quantification radicale que nous avons employée risque d'arrondir ces valeurs sur des centroïdes différents, et ainsi accentuer des différences qui auraient dû rester négligeables. On obtient donc par reconstruction un mauvais alignement des oscillations qui composent les contours. Ceci confirme bien qu'il faut une stratégie plus évoluée pour définir un algorithme de compression comme par exemple une gestion multiéchelle des angles. De plus, une erreur de phase aux échelles grossières entraîne une erreur spatiale *large*, ce qui explique probablement que les artefacts observés soient principalement de nature basses-fréquences. Notons que nous avons observé le même genre de phénomène à la section 2.4.2, en dégradant la phase des coefficients d'ondelettes quaternioniques. Enfin, ici également, aucune fausse couleur n'a été introduite.

La dégradation des angles de l'*axe couleur* a été testée en quantifiant α et β sur 1 et 2 bits (quatrième colonne, figure 4.12). C'est sans grande surprise que de fausses couleurs sont introduites, témoignant du fait que cette donnée d'axe est la seule à porter une information colorimétrique. Observons par exemple les yeux du mandrill, la main de Barbara ou les lignes de l'image **circles**.

Pour conclure, ces différentes expérimentations montrent que la transformée en ondelettes monogène couleur que nous avons proposée est stable par rapport à la quantification, et permet d'encoder efficacement des éléments géométriques colorés. En éliminant pour chaque coefficient de faible amplitude A l'ensemble de ses coordonnées polaires, le contenu visuel de l'image est bien conservé, ce qui appuie la *parcimonie* supposée de cette représentation. La netteté des contours, ou bien la qualité des couleurs peuvent être contrôlées indépendamment, en ajustant la stratégie de quantification de l'amplitude et celle de l'axe couleur. Cet aspect peut être un avantage dans une perspective de compression d'image hiérarchisée (« scalable »). En revanche, nous

avons vu que la forme « oscillante » des artefacts liés à la dégradation de l'amplitude suggèrent d'approfondir la problématique de reconstruction par coefficients monogènes. En particulier, il serait intéressant de prendre en compte l'*orientation* θ afin de diriger de façon analytique les atomes de reconstruction. Malgré cela, constate que l'utilisation d'une méthode de représentation adaptée aux dimensions des images couleur permet d'obtenir une vraie cohérence entre les opérations et les modifications graphiques introduites.

Parmi les nombreuses pistes d'amélioration possibles, une modélisation statistique des coefficients ou une identification des lois par critère d'information permettrait d'optimiser la quantification. La corrélation spatiale de l'axe couleur, qui semble varier lentement à un changement de signe près, pourrait être étudiée, ainsi que la corrélation en général entre les différentes données redondantes de cette représentation. La qualité de la quantification des angles pourrait varier en fonction de l'amplitude, afin d'économiser des données sur la description de coefficients peu pertinents (d'amplitude faible). Enfin, des contraintes liées au système visuel humain comme une pondération dépendante de l'orientation pourrait facilement être introduite car l'orientation des structures liées à chaque coefficient est accessible via θ .

4.4.2 Analyse invariante d'images couleur

L'analyse d'image par ondelettes est en général destinée à manipuler plusieurs résolutions et directions, pour extraire des caractéristiques fréquentielles et/ou spatiales comme des points particuliers. Elle sert dans de nombreuses applications de reconnaissance de formes et de classification, comme une étape d'analyse « bas-niveau », à partir de laquelle des descripteurs éventuellement plus évolués pourront être construits, puis injectés dans un processus de décision. L'avantage de l'utilisation des ondelettes est dans la compacité des données et la rapidité de calcul. Cet aspect devient indispensable lorsque de grandes bases de données d'images sont à mettre en place. Comme nous l'avons dit les ondelettes orthogonales vont introduire un biais notamment dû à leur variance trop importante lorsque les images à analysées ne sont pas alignées de la même façon. C'est pourquoi des transformées aux meilleures capacité d'analyse sont préférées. La transformée proposée ici a été définie pour avoir une analyse invariante par translation/rotation, riche et compacte. Nous proposons ici de conduire une analyse de ses coefficients dans un objectif de description de l'information, pour mettre en valeur leur invariance et mieux illustrer les caractéristiques visuelles qu'ils portent. L'avantage principal de notre méthode est que la couleur est fondamentalement prise en compte, alors que la plupart des outils d'analyse d'image sont scalaires. De plus, la transformée monogène est liée à une interprétation physique des formes décrites, ce qui permettra potentiellement de prendre en compte correctement les aspects du système visuel humain.

Caractérisation des contours

Le modèle monogène est bien adapté aux *contours* des images couleur, que l'on détecte efficacement par les maxima de l'amplitude A . Les valeurs de phase, orientation et axe, sont particulièrement pertinentes en ces points particuliers, car elles caractérisent l'ensemble de la forme dans son voisinage, en un seul coefficient. C'est pourquoi nous proposons de faire une extraction de points particuliers associés aux contours de l'image dans le domaine des ondelettes monogènes couleur, afin d'analyser la relation concrète entre les formes géométriques réelles et les valeurs des coefficients. Cela nous permettra également de tester l'invariance par rotation.

Puisque nous nous focalisons sur des *contours* plutôt que des points isolés, la notion de « maximum » sera comprise dans une certaine direction, c'est-à-dire que nous incluerons les *points selle* comme étant des maxima. En pratique, on retiendra les coefficients dont l'amplitude est supérieure à au moins 6 de ses 8 voisins. Ainsi des contours fins et continus seront extraits, conformément au modèle monogène, et par opposition aux « points particuliers » classiques, plutôt destinés à détecter les coins et les jonctions.

Pour améliorer l'invariance, on utilisera la version non-décimée de la transformée, comme c'est souvent le cas pour la détection de points particuliers par exemple. En effet, bien que la transformée classique dans sa version non-décimée (« UDWT ») ait beaucoup été utilisée en débruitage, elle est également intéressante en analyse [138, 77, 107], car elle offre une meilleure *localisation* des structures.

Analyse des résultats

La figure 4.13 illustre la décomposition d'une image simple avec trois rotations différentes. L'image contient des structures typiques colorées, un bord et une ligne, constituant des cas élémentaires pour l'analyse monogène. La figure montre de gauche à droite : L'image, l'amplitude A pour trois échelles, de la plus grossière à la plus fine, la phase φ_2 , l'axe couleur et l'orientation, pour les mêmes trois échelles. Les phase/axe/orientation ne sont affichées que sur les maxima de l'amplitude. Afin de mieux comparer l'orientation dans les différentes rotations de l'image, un offset a été ajouté pour compenser ces rotations. Ainsi, une parfaite invariance par rotation doit résulter en des affichages *identiques* de θ pour chaque rotation.

Amplitude En observant l'amplitude de la décomposition de la figure 4.13, on remarque qu'elle a approximativement la même forme dans les différentes rotations. Les coefficients « tournent » sans introduire d'oscillation significative, alors que c'est le cas avec les ondelettes classiques. Ceci est dû d'abord au banc de filtres, qui sous forme d'une pyramide utilise des ondelettes aussi isotropes que possible, mais aussi au formalisme monogène, lui-même construit de façon isotrope. L'illustration confirme également l'invariance par changement de phase de A , qui se traduit par une réponse semblable autour des contours de type *bord* et *ligne*. Ceci n'est bien sûr plus vrai à l'échelle d'analyse la plus fine, dans laquelle la largeur de la ligne devient substantielle, ce qui se traduit par une double-réponse de l'amplitude, type gradient, correspondant à décrire cette « large ligne » comme un rectangle. Notons que cette amplitude constitue à notre connaissance le premier détecteur de contour couleur invariant par changement de phase.

Voyons maintenant les angles, pour lesquels nous n'affichons que les valeurs aux maxima de A .

Phase A partir de la correspondance théorique entre la valeur de la phase et le type des contours, on s'attend à obtenir des points rouges ($\varphi_2 = 0$) sur la ligne et des points bleus ($\varphi_2 = \pi/2$) sur les bords du rectangle. A l'échelle la plus grossière, φ_2 est cohérente avec cette perception que nous avons de l'image. Le lissage des coins qui est dû au filtrage isotrope produit des « bords courbés », mais la phase pour autant ne détecte plus la structure de bord (jaune-vert). A cette échelle, le voisinage d'un coin ne peut pas être décrit de façon satisfaisante par une représentation monogène. Aux échelles plus fines, on remarque que la ligne devient progressivement un rectangle, comme nous l'avons expliqué pour l'amplitude, et donc le « rouge » de la phase disparaît. L'invariance par rotation est parfaite à l'échelle -3 , où les bleus restent bleus et les rouges restent rouges, approximative à l'échelle -2 où des variations légères de valeurs de phase sont visibles (le bleu des bords devient cyan, voire parfois vert). A l'échelle -1 , l'invariance est parfaite, mais l'interprétation des structures paraît plus « pixelique ». Il semble qu'à cette échelle, la plus fine, toutes les structures soient vues comme des coins, ce qui pourrait s'expliquer par le fait que tous les contours sont « en escalier », puisque non-alignés sur la grille cartésienne. D'une façon plus générale, le contenu hautes-fréquences d'une image est très dépendant de son échantillonnage, et l'interprétation des représentations analytiques/monogènes est délicat car les « structures » locales réelles sont très « carrées », définies par quelques pixels, et ne ressemblent pas au modèle. De plus, le maximum de l'amplitude en ces points n'a peut-être de sens qu'à des positions intermédiaires aux coefficients, pour lesquelles nous n'avons pas accès à la phase. Pour conclure, la phase vérifie approximativement l'invariance par rotation.

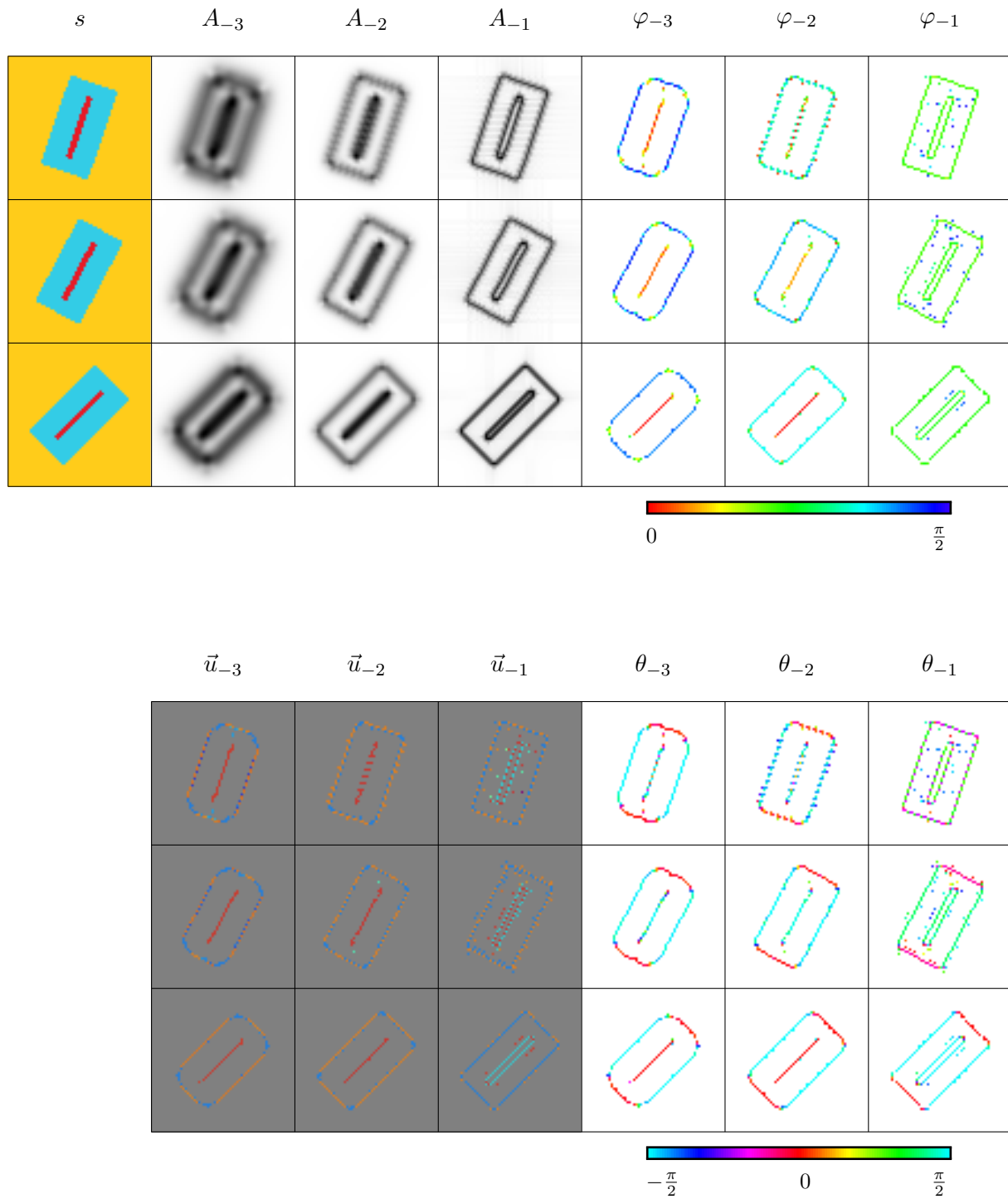


FIGURE 4.13 – Décomposition monogène couleur non décimée ($\gamma = 3$) de trois rotations d'une image de test simple (19°, 28°, 45°). Légende : voir figure 4.9. Pour φ , $\vec{u}$ et θ , on ne montre que les points où A atteint un maximum local.

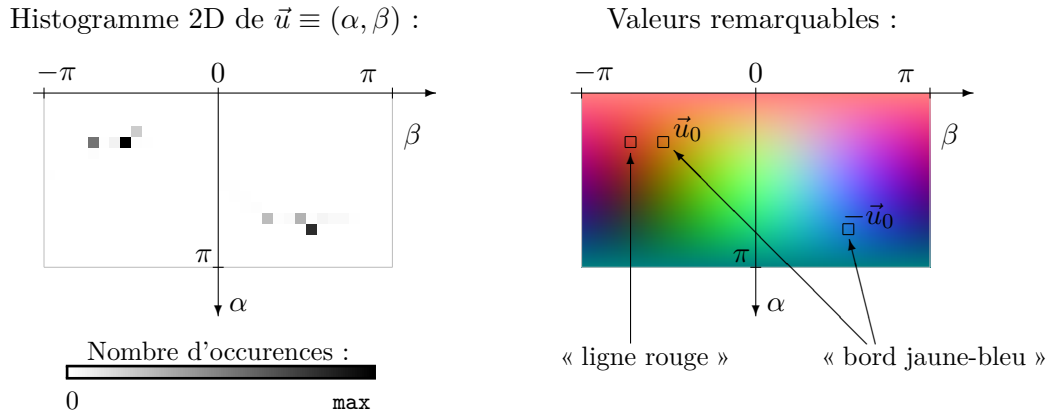


FIGURE 4.14 – Analyse des valeurs de $\vec{u}$ dans l'expérimentation de la figure 4.13. A gauche : Histogramme des couples (α, β) , parmi ceux qui n'ont pas été « grisés ». A droite : Identification du « code couleur » associé aux valeurs principalement observées.

Axe couleur Contrairement à l'amplitude et la phase, dont l'interprétation est identique à celles du signal monogène en niveaux de gris, l'axe couleur est une donnée nouvelle. Elle porte l'information colorimétrique locale de l'image, indépendamment de A et φ_2 . Comme nous l'avons vu à la section 4.2.3, l'axe couleur $\vec{u}$ pose un problème de représentation, et montre des changements de signe autour des bords. On constate en effet que les bords de notre objet sont décrits par un axe couleur dont l'illustration montre des valeurs « oranges » et d'autres « bleues ». Nous illustrons à la figure 4.14, à travers le calcul d'un histogramme des valeurs de $\vec{u}$ observées, que ces deux couleurs correspondent en fait au même axe à un signe près. On retrouve dans la partie droite de la figure les points bleus et oranges, qui correspondent donc à deux points opposés sur la sphère. Ainsi, même si l'illustration de $\vec{u}$ paraît confuse au premier abord, il vient que cette donnée est plus simple, puisque les valeurs sont bien localement *constantes* au signe près, conformément au modèle de signal que nous avons proposé à l'équation (4.25). Notons qu'une meilleure représentation pourrait être obtenue en fusionnant les points opposés, ce qui est possible en utilisant un *tore* au lieu d'une sphère, et qui permettrait de mieux gérer la double circularité de (α, β) .

Étudions enfin les valeurs particulières de $\vec{u}_0$. Dans le cas du bord du rectangle, le jaune vaut $I_1 = [255 \ 204 \ 26]$ et le bleu $I_2 = [51 \ 204 \ 230]$ dans l'espace RGB. Considérons l'axe déterminé par ces deux couleurs $\vec{u} = (I_1 - I_2)/\|I_1 - I_2\|$, et extrayons les coordonnées polaires α et β par l'équation (4.24). On obtient $\alpha = \pi/4$ et $\beta = -\pi/2$, ce qui correspond exactement aux valeurs mesurées (cf. figure 4.14). Le formalisme monogène proposé permet donc d'extraire une notion de différence de couleur, entre les deux zones concernées par ce bord. La couleur de la ligne rouge vaut $I_3 = [230 \ 26 \ 38]$, et l'axe couleur correspondant à cette ligne est déterminé par la différence entre ce rouge et le bleu I_2 . Pour $\vec{u} = (I_3 - I_2)/\|I_3 - I_2\|$, on obtient $\alpha \approx 0.31\pi$ et $\beta = -0.74\pi$. A nouveau, ces valeurs calculées analytiquement concordent avec le résultat de l'analyse monogène. Notons que cette caractérisation de différence de couleur entre les deux zones concernées par un bord ou une ligne ne permet pas de connaître la couleur *absolue* de ces éléments de l'image, qui est plutôt une information basses-fréquences. C'est pourquoi *plusieurs* profils couleur sont possibles pour une même valeur de $\vec{u}$. Pour conclure, l'analyse monogène couleur permet d'extraire une information colorimétrique cohérente, car constante au signe près, et conforme au modèle proposé d'oscillation couleur.

Orientation Nous attendons de l'analyse des orientations locales qu'elle affiche les mêmes valeurs pour les trois versions de l'image, puisque nous y avons ajouté un offset de compensation. Ainsi, les côtés courts du rectangles sont considérés *horizontaux* ($\theta = 0$), et les côtés longs, ainsi que la ligne centrale, sont *verticaux* ($\theta = \pm\pi/2$). A la première échelle, l'orientation détectée ne suit pas ce modèle dans la plupart des cas. Cela est dû au fait qu'à cette échelle, les détails de

l'image ne correspondent pas à notre analyse visuelle des images, comme déjà discuté à propos de la phase. Rappelons également que l'isotropie des ondelettes utilisées n'est qu'approximative, et donc risque d'entraîner une légère favorisation de certaines directions. Aux échelles suivantes, l'analyse est plus en accord avec notre perception, et les bords du rectangles sont bien détectés comme horizontaux et verticaux. La ligne, en revanche n'est pas bien détectée. Cela est dû au fait qu'au niveau d'une ligne, la transformée de Riesz est nulle. On peut revoir à ce sujet la direction du gradient et de $\mathcal{R}$ dans les figures précédentes (4.2, 4.4, 4.7, 4.9 et 4.10), où l'analyse est toujours pondérée au niveau des lignes par une norme nulle, correspondant à un passage par zéro du gradient ou de $\mathcal{R}$. Pourtant, l'information correcte est bien présente *autour* de la ligne, et peut être facilement « étalée » par un lissage du tenseur. Ce « post-traitement » est bien connu dans le cadre des approches par gradient. Nous illustrons à la figure 4.15 le même test d'invariance par rotation avec la version *décimée* de la transformée, en comparant avec l'utilisation du lissage du tenseur. On constate que cette technique de lissage permet de grandement améliorer la qualité de l'orientation. Les bords verticaux du rectangle et la ligne montrent une orientation proche de 0 (rouge), et les bords horizontaux montrent bien une orientation à $\pi/2$ (cyan), quelle que soit l'échelle analysée. La différence avec l'analyse « brute » est surtout constatée à la première échelle. À cette échelle, les contours sont « en escalier », ce qui implique que l'orientation mesurée est discontinue le long de ces contours. Cependant, la *moyenne* de ces orientations, implicitement effectuée par le lissage du tenseur, donne accès à une orientation plus globale, la « pente de l'escalier », qui traduit bien l'orientation du contour réel. Notons également que si l'on extrait l'amplitude et la phase à partir de la norme du tenseur lissé plutôt que de la norme de la transformée de Riesz ponctuelle (ce que nous avons fait dans l'illustration), l'amplitude apparaît plus régulière. Finalement, nous avons confirmé l'invariance par rotation approximative de la donnée d'orientation des ondelettes monogènes couleur. Cette approximation nécessitera néanmoins parfois un post-traitement comme un lissage du tenseur, en particulier pour l'orientation des structures de type *ligne* et pour les structures de la première échelle.

À travers cette application des ondelettes monogènes couleur, nous avons introduit une nouvelle façon d'analyser finement des points caractéristiques de l'image, en prenant ici l'exemple des contours. Ce travail ouvre la voie à une nouvelle génération de détecteurs de points d'intérêt pour les images couleur, basé sur une analyse géométrique fondamentale qui hérite du gradient couleur de S. Di Zenzo aussi bien que du signal monogène, apportant une dimension physique aux données. Ce dernier aspect différencie clairement notre approche des outils type SIFT [100], qui n'utilisent pas de notion *signal*, et des applications prenant en compte le système visuel humain sont à considérer. Il serait également intéressant de conduire une application en analyse de textures, à la lumière de ce qui a été fait avec les ondelettes quaternioniques (section 2.5), pour lesquelles l'invariance et la phase ont permis d'améliorer la reconnaissance.

4.5 BILAN

Le but global du travail de cette thèse - définir une représentation riche et inversible dans l'esprit du signal analytique - semble avoir été atteint avec les ondelettes monogènes couleur présentées dans ce chapitre. Nous venons de mettre à l'épreuve cette représentation dans deux champs d'application classiques, ce qui a permis de confirmer les propriétés avancées.

L'analyse fournie par les coefficients est approximativement invariante par translation et rotation, et directement interprétable en termes de caractéristiques géométriques et colorimétriques des structures locales. Le modèle sous-jacent est un contour couleur rectiligne dont les valeurs varient localement dans une direction *spatiale* unique (déterminée par θ) ainsi que dans une direction *colorimétrique* unique (déterminée par $\vec{u}$). La discrimination bord/ligne est assurée par la phase φ_2 et la détection de contour invariante par translation/rotation/changement de phase est fournie par A . Les données sont donc complémentaires et cohérentes avec notre perception.

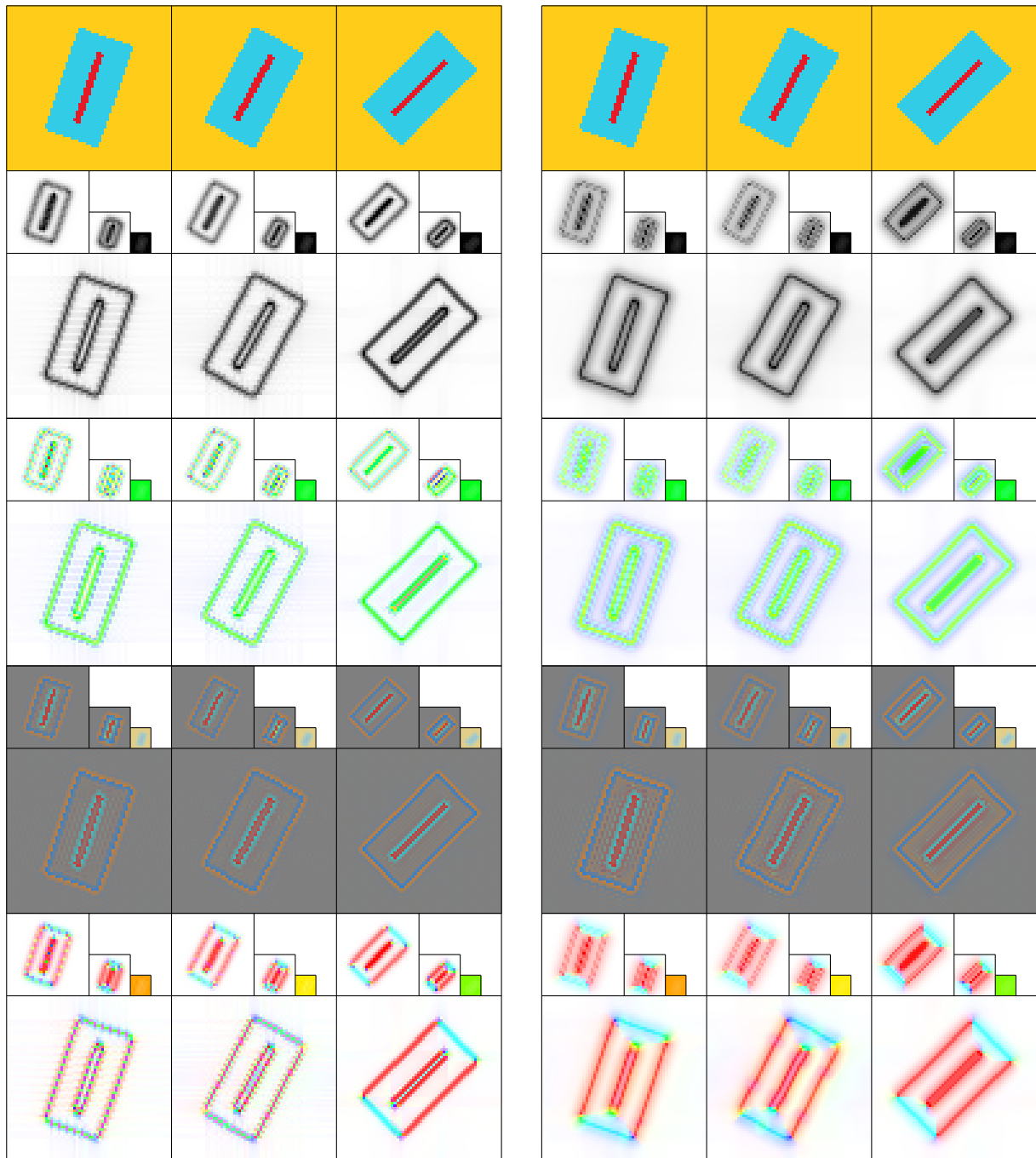


FIGURE 4.15 – Variance par rotation de la transformée en ondelettes monogènes couleur sans (gauche) et avec (droite) lissage du tenseur ($\gamma = 3$, lissage du tenseur gaussien avec $\sigma = 1.5$).

L'intérêt d'une prise en compte fondamentale de la couleur a été validé par l'absence de fausses couleurs dans les reconstructions partielles. L'organisation du contenu visuel à travers les coefficients permet de hiérarchiser les informations, et de contrôler indépendamment la qualité des contours et de la couleur. La parcimonie de la représentation a également été confirmée par la qualité des reconstructions obtenues par mise à zéro des coefficients de faible amplitude. Ces expérimentations ouvrent la voie vers de nouvelles méthodes de codage, de débruitage, de détection de points d'intérêts... pour lesquelles la prise en compte d'informations visuelles/physiques plus proche du système visuel humain (phase, orientation...) permettra probablement d'améliorer l'existant.

Cette dernière partie expérimentale a soulevé clairement la problématique du schéma numérique. Nous avons vu que les représentations monogènes manquent théoriquement d'une notion de reconstruction, qui n'est pas traitée dans la littérature - focalisée sur l'analyse. De plus, l'isotropie des bancs de filtres que nous utilisons est approximative, et les filtres de synthèse pâtissent de la qualité des filtres d'analyse qui « consomment » tous les degrés de liberté de la construction. Ces défauts liés au schéma numérique ont provoqué l'apparition d'*oscillations* dans les reconstructions partielles et un manque de précision dans les valeurs des coefficients.

Nous proposons pour finir ce mémoire de thèse d'ouvrir une piste de réflexion sur cette problématique discrète autour des représentations monogènes.

4.6 UNE NOUVELLE PISTE DE SCHÉMA NUMÉRIQUE

Nous avons vu que les outils de représentation monogène manquent de définitions claires en termes de schémas numériques. En particulier, la transformée de Riesz $\mathcal{R}$ est un opérateur *continu* et à réponse impulsionnelle *infinie*. En pratique, les filtrages sont réalisés dans le domaine de Fourier par FFT, et consistent à échantillonner la réponse fréquentielle de la transformée de Riesz sur $\omega \in [-\pi; \pi]^2$. En général, le signal à analyser est discret et fini, et l'opération revient donc à periodiser la réponse fréquentielle de $\mathcal{R}$, ce qui est équivalent à échantillonner sa réponse impulsionnelle. Cette méthode n'applique donc pas *exactement* une transformée de Riesz, mais il n'y a pas à notre connaissance de travaux sur sa discrétisation, ni sur les implications théoriques d'une telle mise à l'œuvre.

Pourtant, nous pensons que les représentations monogènes, qui sont des outils typiquement « signal », devraient comme la transformée de Fourier et la transformée en ondelettes profiter d'un schéma numérique maîtrisé. Afin d'ouvrir la voie sur cette thématique, nous proposons une réflexion basée sur deux constats :

- Il existe un lien fondamental entre la représentation monogène et la transformée de Radon [84] ;
- Des algorithmes discrets de transformée de Radon ont déjà été définis, notamment au laboratoire XLIM-SIC [19].

Plus précisément, le concept de l'analyse monogène est équivalent à une analyse de Hilbert 1D effectuée dans le domaine de Radon. De cette façon, la transformée de Radon est responsable de l'*isotropie*, problématique cruciale dans le monde discret, et propriété de base du concept monogène.

Dans [19], une transformée de Radon discrète avec reconstruction exacte est définie à l'aide de notions de géométrie analytique discrète. Nous allons étudier l'utilisation de cet outil bien établi pour réaliser une décomposition monogène. On s'attend par conséquent à ce que l'analyse soit numériquement stable et géométriquement cohérente. Ce travail ouvre la voie vers des alternatives au banc de filtres de M. Unser [144], dans le but de construire une méthode plus efficace et plus simple.

Après avoir défini la transformée de Radon et ses propriétés, nous proposerons une analyse monogène basée Radon.

4.6.1 Le domaine de Radon

Le domaine de Radon représente les fonctions 2D par un ensemble de fonctions 1D obtenues par projections de la première le long de différentes orientations. La transformée de Radon forme un lien fondamental entre les transformées de Fourier 1D et 2D, et permet de réaliser des filtrages *isotropes* à partir d'un filtre 1D.

Soit une fonction $s(\mathbf{x})$, sa projection dans le domaine de Radon le long de la direction θ est définie par :

$$s_\theta(t) = \int_{\mathbb{R}} s(\tau \sin \theta + t \cos \theta, -\tau \cos \theta + t \sin \theta) d\tau \quad (4.36)$$

Cette représentation est utilisée par exemple en reconstruction tomographique, où les données acquises sont formées par des projections spatiales dans plusieurs directions. La transformée de Radon inverse est donc utilisée.

Notons la propriété facilement vérifiable suivante :

$$s_\theta(-t) = s_{\theta+\pi}(t) \quad (4.37)$$

qui assure la cohérence circulaire de la représentation. Le « Fourier slice theorem » définit le lien avec la transformée de Fourier :

$$\hat{s}_\theta(f) = \hat{s}(f \cos \theta, f \sin \theta) \quad (4.38)$$

où $\hat{s}_\theta$ est la transformée de Fourier 1D de s_θ pour un θ fixé et $\hat{s}$ est la transformée de Fourier 2D de s . Dans le domaine de Fourier, la propriété (4.37) devient :

$$\hat{s}_\theta(-f) = s_{\theta+\pi}(f) \quad (4.39)$$

ce qui assure la consistance de $\hat{s}$. C'est pourquoi en pratique, les projections de Radon sont restreintes à $\theta \in [0; \pi[$.

Le « Fourier slice theorem » permet de déduire des propriétés intéressantes sur le filtrage isotrope. Soient deux signaux $s(\mathbf{x})$, $h(\mathbf{x})$ et leurs transformées de radon respectives $s_\theta(t)$, $h_\theta(t)$, la convolution dans chaque domaine est liée de la façon suivante :

$$(s_\theta * h_\theta)(t) = (s ** h)_\theta(t) \quad (4.40)$$

où $*$ est la convolution 1D et $**$ la convolution 2D. $(s ** h)_\theta(t)$ est donc la transformée de Radon de $(s ** h)(\mathbf{x})$.

Si le filtre h est isotrope $h(\mathbf{x}) = h_1(\|\mathbf{x}\|)$ alors sa transformée de Radon ne dépend pas de θ et est symétrique par rapport à t : $h_\theta(t) = h'_1(|t|)$. Cela signifie qu'un filtrage 1D dans le domaine de Radon avec un filtre symétrique est équivalent à un filtrage 2D isotrope dans le domaine spatial.

Un autre cas particulier est celui de la transformée de Riesz. Il a été montré dans [13] que $\mathcal{R}$ est équivalent à une transformée de Hilbert $\mathcal{H}$ (1D) indépendante sur chaque projection de Radon, combinée à une pondération de ces projections par une fonction sinusoïdale de θ :

$$\{\mathcal{R}s\}_\theta(t) = \{\mathcal{H}s_\theta\}(t)e^{j\theta} \quad (4.41)$$

La transformée de Hilbert est bien connue et déjà intégrée dans le domaine des ondelettes à travers l'algorithme Dual-Tree [26, 80] (section 1.2.3). La réalisation d'une analyse monogène est donc réduite à un problème plus classique dans le domaine de Radon. Nous présentons maintenant un algorithme discret de transformée de Radon.

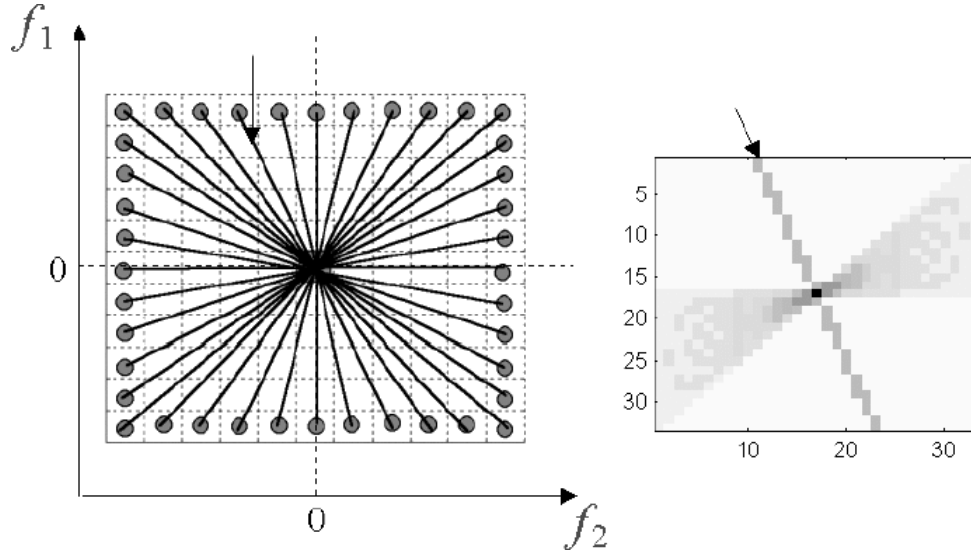
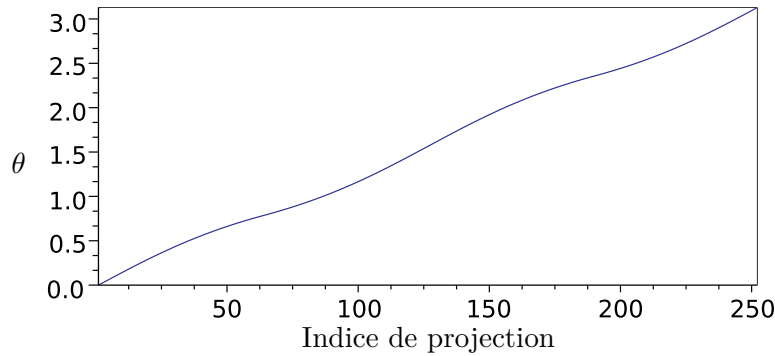


FIGURE 4.16 – Extraction de lignes sur la transformée de Fourier d'un signal 2D.

FIGURE 4.17 – Valeurs de θ pour chaque projection de Radon d'un signal quelconque de taille 127×127 .

4.6.2 Algorithme discret

Le travail de [19] consiste à définir une décomposition *discrète* en utilisant des notions de géométrie discrète. En se basant sur le « Fourier slice theorem », on peut considérer des lignes discrètes dans le domaine de Fourier avant d'appliquer une transformée de Fourier inverse 1D sur chaque ligne extraite. Les lignes sont organisées comme illustré à la figure 4.16. Un paramètre d'épaisseur arithmétique peut être utilisé pour contrôler à la fois la redondance et la connexité des lignes. Dans tous les cas, une reconstruction exacte est garantie et l'algorithme est simple et rapide.

Dans ce travail, nous avons besoin de connaître les valeurs de θ associées aux lignes extraites (afin de calculer l'exponentielle complexe de l'équation (4.41)). Comme illustré à la figure 4.17, θ n'évolue pas linéairement avec les indices de projection, à cause de la nature cartésienne de l'organisation des données. Nous avons calculé sa valeur géométriquement, conformément au schéma de la figure 4.16, à partir d'arc-tangentes des rapports de coordonnées. Nous verrons que dans le cas d'une analyse monogène, cette différence avec le monde continu n'est pas un problème et n'empêche pas de construire un algorithme précis.

Le calcul d'une analyse monogène basée sur une transformée de Radon discrète peut maintenant être défini comme suit :

- Calculer la transformée de Radon de s ;
- Appliquer un filtre passe-bande 1D pour sélectionner une échelle (ceci est équivalent à un filtrage isotrope) ;
- Calculer la transformée de Hilbert de chaque projection s_θ ;
- Multiplier chaque projection par $e^{j\theta}$ (à partir du calcul de θ décrit plus haut) ;

- Calculer la transformée de Radon inverse sur la partie réelle et la partie imaginaire séparément (on obtient s_{r_1} et s_{r_2});
- Convertir s , s_{r_1} et s_{r_2} en coordonnées sphériques conformément au signal monogène.

Les données obtenues expérimentalement sont très « habituelles » : l'amplitude traduit la présence d'éléments visuels dans s alors que la phase apporte une description supplémentaire des ces éléments - la classification bord/ligne. L'orientation locale vient compléter l'analyse. Nous avons pu mesurer expérimentalement que les données sont presque identiques au calcul classique d'une analyse monogène par FFT. L'erreur est de l'ordre de 10^{-4} pour l'amplitude et de 10^{-2} pour la phase et l'orientation, avec de plus grandes différences pour les coefficients d'amplitude faible qui sont naturellement instables. Notons que le schéma est rapide, grâce à la simplicité de l'algorithme de Radon que nous utilisons.

4.6.3 Discussion

Ce travail préliminaire montre qu'une analyse monogène discrète peut être effectuée simplement et rapidement dans le domaine de Radon en utilisant les algorithmes existants de transformée de Radon. La transformée utilisée offre une reconstruction exacte grâce à la notion de ligne discrète, ce qui est une propriété importante dans notre champ de recherche. L'amélioration par rapport à la méthode par FFT n'est pas claire à cette étape, mais pourrait devenir substantielle par une évolution vers un schéma *multi-échelle*, ou vers des dimensions supérieures (voir extension 3D de l'algorithme de Radon [68]).

En fait, la définition de transformées en *ondelettes* monogènes et analytiques souffre encore aujourd'hui d'un manque de définitions discrètes. Une solution réside peut-être dans l'application d'une transformée en ondelettes complexes 1D dans le domaine de Radon. Cependant, il faudrait prendre en compte le ré-échantillonnage des sous-bandes, et l'algorithme de transformée de Radon inverse ne permet pas ce genre de changements de dimension. Plus concrètement, il faudrait reconstruire une image plus petite avec le même nombre de projections, dans une version raccourcie. Pour obtenir une représentation *monogène* des sous-bandes, cette étape de reconstruction est indispensable.

La considération des coefficients obtenus par décomposition en ondelettes des projections de Radon *sans Radon inverse* correspond à une représentation en *ridgelets*. Dans le cas complexe, on obtient des *ridgelets complexes* [28], qui consistent à appliquer l'algorithme Dual-Tree 1D dans le domaine de Radon. Dans cet article [28], une application en débruitage par seuillage est proposée. Contrairement aux ondelettes monogènes, les ridgelets complexes ne sont ni isotropes, ni localisées, et n'offrent pas *a priori* d'interprétation visuelle de la phase.

Des *curvelets monogènes* ont également été proposées dans [141] comme une transformée « monogène directionnelle ». Les atomes de curvelets sont combinés avec leur transformée de Riesz pour améliorer l'analyse. L'avantage cité est que la notion de phase est ajoutée à la grande résolution angulaire des curvelets. Cela correspond donc à une transformée de Gabor munie de nombreuses directions et d'un système d'échantillonnage et de reconstruction. Au contraire, une transformée en ondelettes monogène classique s'attachera à intégrer *toutes* les orientations possibles dans une même analyse, à travers un concept plus évolué de signal analytique 2D.

Pour conclure, nous pensons que la prochaine étape est de définir une méthode discrète de ré-échantillonnage dans le domaine de Radon. Dans une perspective d'utiliser des ondelettes 1D sur les projections de Radon, un sous-échantillonnage par exemple consistera à retirer un échantillon sur deux sur *chaque* projection. On aura donc supprimé la moitié des données. La dimension de l'image reconstruite doit alors vraisemblablement avoir réduit sa taille « de moitié », ce qui suggère un échantillonnage en quinconce. Le problème est ouvert. Un autre axe connexe est celui du filtrage isotrope dans le domaine de Radon, qui à notre connaissance n'a pas été traité dans la littérature.

CONCLUSION

Nous nous sommes intéressés dans cette thèse aux représentations des signaux par ondelettes analytiques peu redondantes permettant une reconstruction parfaite. Ce genre d'outil est destiné à offrir un compromis entre la compacité des données et la cohérence de l'analyse qu'elles constituent, grâce à une amplitude invariante et une phase dont une interprétation géométrique peut être déduite. L'algorithme Dual-Tree qui est apparu en 1999 a ouvert la voie vers des transformées analytiques efficaces, et le but de notre travail était de définir une extension de ce genre d'outils pour les images couleur. L'approche *signal* que nous avons nécessité deux étapes d'extension des outils.

La première est spatiale, et consiste à étendre les outils 1D pour les images en niveaux de gris. Nous avons vu que les extensions 2D classiques (type Gabor) dont le Dual-Tree 2D offre une implantation efficace, souffrent d'une définition insatisfaisante du signal analytique. La seconde concerne l'aspect vectoriel des images couleurs, alors que les outils classiques sont scalaires. Sur ce point, aucun outil de représentation tel que ceux que nous étudions n'a été proposé à notre connaissance. Il s'avère que l'algèbre des complexes est trop limitée pour généraliser de façon élégante les outils de traitement du signal, c'est pourquoi nous avons étudié les approches quaternioniques.

Historiquement, les quaternions en traitement d'image ont d'abord concerné la couleur, à travers le travail initié par S. Sangwine et poursuivi par exemple dans notre équipe. Un banc de filtres couleur non-marginal a été finalement proposé, mais il est apparu que le formalisme quaternionique n'était pas indispensable dans la définition, qui pouvait également s'écrire dans des espaces euclidiens classiques. De plus, il s'agissait d'une approche indépendante de la notion de signal analytique qui est au centre de cette thèse.

Le formalisme quaternionique de T. Bülöw se propose de rester limité aux images en niveaux de gris, afin d'utiliser la puissance de l'algèbre des quaternions dans l'analyse géométrique des signaux 2D. La transformée de Fourier quaternionique décompose les images dans une base d'atomes plus riche que les classiques ondes planes, richesse qui se retrouve dans le domaine des ondelettes avec la QWT, dont les coefficients possèdent une amplitude invariante par translation accompagnée d'une phase de trois angle. Ces trois angles permettent de déterminer finement la forme locale du signal, en incluant la possibilité d'une forme isotrope, impossible avec Gabor.

Après avoir approfondi l'étude des approches quaternioniques du point de vue des atomes de reconstruction, nous avons proposé deux applications de cette nouvelle transformée. La première consistait à mettre en évidence la flexibilité des atomes de reconstruction, par oppositions aux ondelettes horizontales, verticales et diagonales du schéma séparable classique (DWT). Les expérimentations de quantification des coefficients ont montré que la QWT restituait mieux les contours que la DWT, pour une quantité de données équivalentes. La seconde consistait à mettre à l'épreuve les nouvelles informations offertes par une telle décomposition dans un contexte de classification de textures. À nouveau, l'information de la phase a permis de surpasser la DWT, grâce à une mesure de « directionnalité » accessible simplement dans la QWT, et impossible avec les ondelettes classiques.

Les ondelettes quaternioniques bénéficient d'un schéma de décomposition et reconstruction simple et bien défini. En revanche, ce schéma est encore trop lié aux lignes et aux colonnes de l'image, ce qui se traduit par une variance par rotation de l'analyse quaternionique. De plus, la prise en compte de la couleur par les quaternions, qui fait notamment partie des travaux

antérieurs dans notre équipe, n'a pas abouti à une représentation en ondelettes couleur satisfaisante. C'est pourquoi il semble qu'à son tour, l'algèbre des quaternions soit trop limitée.

Le formalisme monogène de M. Felsberg dépasse cette limite en utilisant d'une part des algèbres encore plus générales, et d'autre part une nouvelle notion de signal analytique isotrope, et donc invariant par rotation. L'interprétation de la phase monogène est plus simple que celle de la phase quaternionique, mais en contrepartie la mise en évidence d'atomes de reconstruction n'a pas été faite. La dimension numérique a néanmoins été engagée avec le banc de filtres monogène de M. Unser, qui permet une reconstruction parfaite, sans toutefois proposer de méthode de fusion des informations monogènes en une reconstruction univoque. Le travail de synthèse que nous avons mené sur le concept monogène et sur cet algorithme a révélé une faiblesse en termes de schéma numérique, en particulier en ce qui concerne la reconstruction de l'image, qui reste à l'heure actuelle un sujet ouvert.

Afin de donner une première solution de représentation analytique pour les images couleur, nous avons défini une extension couleur des ondelettes monogènes en nous basant sur l'analyse de Clifford. Il s'agit d'une transformée en ondelettes non-séparable et non-marginale, qui profite donc des propriétés d'invariance par translation et rotation du formalisme monogène. Les sous-bandes obtenues sont composées de trois composantes liées à l'énergie couleur, ainsi qu'à deux composantes correspondant à une analyse géométrique.

Suite à cette définition, nous avons proposé d'appliquer les ondelettes monogènes couleur dans un contexte de débruitage. En comparaison avec les ondelettes séparables qui introduisent de fortes oscillations et les ondelettes non-décimées qui font référence en débruitage, la transformée monogène couleur a montré des performances intermédiaires, se plaçant comme un compromis entre la redondance des données et la qualité du débruitage. Cependant, la difficulté que nous avons eue à définir une méthode de reconstruction, ainsi que le manque de régularité des artefacts introduits montrent clairement la limite en termes de schéma numérique monogène. De plus, la notion de phase dans cette démarche n'est pas suffisamment aboutie pour avoir une caractérisation physique des coefficients. C'est pourquoi nous avons décidé de reprendre le concept de signal analytique et monogène pour les signaux couleur, afin de définir une caractérisation du signal maîtrisée.

Nous avons proposé dans la dernière partie de nous inspirer des outils de géométrie différentielle étendus aux images multivaluées, qui représente probablement un des outils couleur les plus efficaces à l'heure actuelle dans la communauté de traitement d'image, en termes d'analyse de la géométrie locale. Sur ce principe, nous avons défini une transformée de Riesz couleur, que nous avons ensuite identifiée à une opération de déphasage. La considération des images couleur à travers des sous-bandes fréquentielles nous a amené à construire un modèle oscillatoire de signal couleur, localement paramétré par une direction variable dans l'espace couleur : « l'axe couleur ». Nous avons ensuite construit une nouvelle notion de phase couleur, définie par réduction à un cas scalaire grâce à la séparation de l'information colorimétrique portée par l'axe couleur, et de l'information oscillatoire. Grâce à la transformée de Riesz couleur, nous avons pu définir une procédure simple d'extraction des caractéristiques physiques locales des images couleur : amplitude, phase, orientation, axe couleur. Ces caractéristiques s'interprètent de la façon suivante :

L'amplitude traduit une notion d'énergie à une certaine échelle, elle se comporte donc comme un détecteur de contours. Il réagit de façon invariante aux lignes et aux bords, et détecte toutes les différences de couleur, même les contours isoluminants.

La **phase** traduit une notion physique oscillatoire que l'on peut interpréter en termes de caractérisation du type de contour, dont un modèle continu entre les structures de type « bord » et les structures de type « ligne » est considéré.

L'orientation indique la direction principale de la structure. Elle est identique à une direction de gradient couleur.

L'axe couleur, qui est un vecteur unitaire, contient toute l'information colorimétrique liée

à un voisinage spatial, à une échelle donnée. Cette information est différentielle, c'est-à-dire qu'elle traduit la position relative dans l'espace colorimétrique des deux couleurs impliquées dans la structure locale de bord ou de ligne.

L'extension dans le domaine des ondelettes a pu se faire naturellement, car toute la partie calculatoire du formalisme monogène couleur que nous avons développé se place dans la recombinaison des coefficients obtenus par un traitement marginal. On a alors pu utiliser directement les bancs de filtres scalaires, que nous avons appliqué marginalement, pour finalement calculer de façon non-marginale les caractéristiques monogènes couleur. L'ensemble de ces caractéristiques est interprétable en termes visuels, et contient l'information suffisante pour reconstruire l'image correspondante. Nous avons donc finalement construit une représentation en ondelettes monogènes couleur peu redondante et inversible.

Afin de positionner ce nouvel outil d'un point de vue expérimental, nous avons proposé une mise en pratique de cette transformée à travers deux applications classiques des ondelettes.

L'aspect compact de la transformée a d'abord été testée par des reconstructions partielles, effectuées après différentes configurations de quantification des coefficients. Cela a permis de voir que les différentes composantes des coefficients monogènes couleur portent bien des informations de géométrie et de couleur physiquement interprétables. En effet, les dégradations appliquées sur certaines composantes affectaient certaines caractéristiques de l'image, comme la netteté des contours, ou la précision des couleurs. Nous avons mis en lumière que l'ajustement de la quantification d'une composante permettait de contrôler indépendamment ces aspects graphiques. Cette indépendance garantit en particulier que si la quantification de l'axe couleur est suffisante, la dégradation des autres composantes n'introduit pas de fausse couleur dans l'image. Ce résultat préliminaire ouvre la voie vers de nouvelles méthodes de codage d'image, où des quantités directement liées à des informations visuelles importantes permettront de compresser efficacement les images couleur, en maîtrisant les artefacts introduits par un fort taux de compression.

La deuxième expérimentation consistait à vérifier la qualité d'analyse des structures locales de l'image, et en particulier l'invariance par rotation. Nous avons vu que la description fournie par l'ensemble des coefficients constitue effectivement une analyse géométrique fine, associant les propriétés du gradient couleur à celles du signal monogène scalaire. Ainsi, une détection de contours multi-échelle est effectuée, tout en classifiant ces contours selon leur parité (bord/ligne), leur orientation, et les couleurs relatives qu'ils impliquent. L'invariance par rotation est vérifiée, à l'exception de la première échelle qui est particulière, car elle implique des structures très « pixeliques », qui ne correspondent pas au modèle de signal. Notre représentation monogène couleur préfigure donc une nouvelle famille de détecteurs de points d'intérêt, intégrant cette notion nouvelle de phase couleur.

Dans les deux champs applicatifs étudiés, les problèmes liés au schéma numérique se sont faits ressentir. La méthode de reconstruction que nous avons proposée semble insuffisante pour reconstruire les images à partir d'atomes monogène réellement pondérés et déphasés. En pratique, seule la partie primaire, associée à des ondelettes paires et isotropes, est utilisée pour la synthèse. Il n'y a pas comme avec la QWT ou le Dual-Tree de combinaison de parties paires et impaires, aboutissant à un modelage adaptatif des atomes. Cette limite, que nous avons mise en évidence notamment par la considération d'applications de *traitement*, n'est pas traitée dans la littérature. La définition d'une transformée en ondelettes monogène *complète* nécessite de déterminer les atomes de reconstruction. Ce travail fait partie de nos perspectives.

Enfin, la transformée en ondelettes monogène couleur que nous avons finalement proposée est la première transformée analytique spécialement dédiée aux images couleur, et profite des différentes étapes d'amélioration des représentations analytiques que nous avons étudiées. À savoir, cette transformée est non-séparable, non-marginale, approximativement invariante par translation et rotation, peu redondante ($\approx 1.8\times$) et inversible. Finalement, les bonnes propriétés de cette nouvelle transformée et la concordance de l'analyse avec notre perception valident la démarche de définir un outil vectoriel, en considérant un signal couleur dès le début de la

construction. En ce qui concerne le choix d'une piste de recherche autour du signal analytique, il est à l'origine de la *complémentarité* des informations obtenues par l'amplitude, la phase et l'axe couleur.

D'un point de vue codage d'image hiérarchisé, la richesse de la décomposition ouvre toutes les possibilités pour ordonner les coefficients en fonction de l'amplitude par exemple, mais aussi de leur orientation ou de leur phase, qui peuvent permettre de faire intervenir certaines contraintes du système visuel. Ces derniers aspects se placent tout-à-fait dans le cadre du projet ANR CAIMAN lié à cette thèse.

Les perspectives de ce travail sont les suivantes.

La reconstruction par atomes monogènes doit être définie de façon univoque, afin d'identifier une méthode de fusion des deux bancs de filtres. Cette méthode doit selon nous correspondre à une combinaison linéaire des composantes cartésiennes des filtres monogènes, permettant de conférer à l'atome reconstruit sa phase et son orientation. Cette dimension qui était automatique dans la transformée quaternionique est absente des outils monogènes à l'heure actuelle, et pose problème dans les application de traitement. Nous pensons que les ondelettes monogènes doivent être liées à une nouvelle famille d'atomes de reconstruction, dont la forme serait définie par les caractéristiques monogènes du coefficient associé. Notre travail futur consistera plus généralement à définir un schéma numérique complet pour la décomposition et la reconstruction monogènes.

Le modèle d'oscillation couleur proposé pour l'extension des représentations monogènes pourrait être généralisé, afin de résoudre notamment les problèmes de passage par zero de certaines composantes, qui aboutissent à une caractérisation parfois instable de certaines données. Cela est dû à notre avis au fait que nous n'exploitons pas entièrement les données d'analyse offertes par la transformée de Riesz couleur. Plus particulièrement, un deuxième axe couleur pourrait probablement être extrait, ce qui permettrait d'enrichir l'analyse colorimétrique de la représentation monogène couleur. En effet, à ce stade, l'extension couleur proposée est limitée à un modèle de signal selon lequel les oscillations locales évoluent le long d'un axe linéaire. Il semble satisfaisant pour les structures de type bord et ligne, mais pourrait probablement englober d'autres types de structures couleur par une caractérisation plus évoluée. De plus, si un travail doit être fait autour du choix de l'espace colorimétrique, aspect très important vis-à-vis des problématiques liées au système visuel humain, nous pensons qu'il devrait se faire conjointement à la définition du modèle couleur.

Des perspectives plus applicatives sont également envisagées. Des descripteurs de points d'intérêts plus évolués pourraient être définis à partir des coefficients d'ondelettes monogènes couleur, et être comparés à des outils comme SIFT.

Une modélisation statistique des coefficients monogènes pourrait permettre de définir une procédure optimisée de quantification, afin d'évoluer vers un schéma de codage d'image hiérarchisé, prenant en compte des contraintes du système visuel humain.

La réflexion sur la modélisation oscillatoire des images couleur, plus généralement des signaux multidimensionnels et multivalués, ouvre la voie également sur une nouvelle définition de transformée de Fourier couleur, guidée par l'interprétation physique des coefficients qui permettra notamment de faire le lien avec la perception.

PUBLICATIONS ASSOCIÉES

Le travail présenté dans ce mémoire de thèse a été valorisé tout au long de ces trois dernières années à travers 9 publications dans des actes de conférences avec comité de lecture, dont 7 internationales, 2 publications dans des revues internationales (dont une à paraître), et 2 chapitres de livres à paraître.

Articles de revues

- [133] R. Souillard and P. Carré. Quaternionic wavelets for texture classification. *Elsevier Pattern Recognition Letters*, 32(13) :1669–1678, 2011.
- [137] R. Souillard, P. Carré, and C. Fernandez-Maloigne. Vector extension of monogenic wavelets for geometric representation of color images. *IEEE Transactions on Image Processing (à paraître)*, 2012.

Articles de conférences

- [127] R. Souillard and P. Carré. Ondelettes quaternioniques pour la classification de textures. *Actes Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'10)*, pages 222–229, Caen France, 2010.
- [129] R. Souillard and P. Carré. Quaternionic wavelets for texture classification. *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*, pages 4134–4137, Dallas, TX, USA, 2010.
- [128] R. Souillard and P. Carré. Quaternionic wavelets for image coding. *Proc. Europ. Sig. Proc. Conf. (EUSIPCO'10)*, pages 125–129, Aalborg, Denmark, 8 2010.
- [130] R. Souillard and P. Carré. Color extension of monogenic wavelets with geometric algebra : application to color image denoising. *Proc. Int'l Conf. Clifford Algebras and Appli. in Math. Physics (ICCA9)*, pages 15–20, Weimar, Germany, July 2011.
- [132] R. Souillard and P. Carré. Ondelettes monogènes couleur pour l'analyse d'image. Actes GRETSI 2011, number ID52, Bordeaux, France, September 2011.
- [131] R. Souillard and P. Carré. Color monogenic wavelets for image analysis. *Proc. IEEE Int'l Conf. on Image Processing*, pages 277–280, Brussels, Belgium, September 2011.
- [136] R. Souillard and P. Carré. Tensor based generalization of monogenic wavelets for coherent multiscale local phase analysis of color images. *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*, pages 1501–1504, Kyoto, Japon, March 2012.
- [134] R. Souillard and P. Carré. Color monogenic wavelet representation based on a tensor-like use of the riesz transform : application to image coding. *Proc. Applied Geometric Algebras in Computer Science and Engineering*, number 16, La Rochelle, France, 2012.
- [135] R. Souillard and P. Carré. A discrete approach to monogenic analysis through radon transform. *Proc. Applied Geometric Algebras in Computer Science and Engineering*, number 40, La Rochelle, France, 2012.

Chapitres de livres

- R. Souillard, P. Carré and C. Fernandez-Maloigne, The definition of a discrete color monogenic wavelet transform, in *Advances In Heuristic Signal Processing And Applications*, Springer (à paraître).

- R. Soulard and P. Carré, Color extension of monogenic wavelets with geometric algebra : application to color image denoising , in E. Hitzer, S. Sangwine (eds.), "Quaternion and Clifford Fourier transforms and wavelets", Trends in Mathematics, Birkhauser, Basel, 2013 (à paraître)

RÉSULTATS DE CLASSIFICATION DE TEXTURE

Nous exposons ici la table complète des taux de reconnaissance (en pourcentages) obtenus par classe, en termes de « taux moyen » ainsi que d'écarts-types.

Classe	DWT		QWT Mag.		QWT Phase		QWT		CWT	
	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.
Brod. 1	73.00	13.91	86.40	11.29	99.47	1.82	93.87	7.85	93.20	8.21
Brod. 2	60.53	11.32	73.73	8.88	40.40	10.75	80.60	8.32	81.87	7.41
Brod. 3	72.87	10.56	82.00	8.61	56.40	13.38	98.07	3.04	64.93	11.63
Brod. 4	67.20	14.33	65.60	13.56	97.07	5.22	97.67	6.17	100.00	0.00
Brod. 5	67.73	12.42	42.27	12.25	33.33	10.72	63.73	10.35	70.93	10.83
Brod. 6	99.20	2.73	100.00	0.00	99.00	3.47	100.00	0.00	99.87	1.33
Brod. 7	41.73	10.87	23.47	9.36	48.40	10.06	47.20	11.24	60.13	12.35
Brod. 8	95.73	3.73	89.87	6.73	99.73	1.31	99.27	2.10	99.47	1.82
Brod. 9	56.07	13.64	82.47	9.02	79.67	11.58	90.67	8.63	82.47	9.46
Brod. 10	61.60	11.92	77.67	9.35	37.53	12.45	85.13	8.68	60.13	11.25
Brod. 11	81.80	8.20	91.87	5.73	91.40	6.52	97.93	3.51	100.00	0.00
Brod. 12	80.27	9.13	88.87	8.48	46.53	9.28	95.20	5.61	96.33	4.58
Brod. 13	37.40	9.98	19.40	10.03	14.93	6.84	34.93	12.29	51.67	11.23
Brod. 14	31.93	12.40	21.47	9.36	48.67	10.05	71.87	9.55	12.60	8.42
Brod. 15	77.00	10.52	68.07	9.77	48.47	12.17	70.33	9.01	77.67	9.73
Brod. 16	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 17	100.00	0.00	99.00	2.39	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 18	95.20	6.97	98.07	3.58	92.80	5.58	99.47	1.82	84.40	9.00
Brod. 19	81.93	9.95	49.93	12.66	72.73	11.01	86.87	7.49	65.40	11.75
Brod. 20	97.73	4.57	100.00	0.00	99.93	0.67	100.00	0.00	99.67	1.46
Brod. 21	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 22	86.13	8.08	52.93	11.56	83.13	10.95	90.00	7.31	64.93	11.47
Brod. 23	44.07	11.13	37.07	12.69	32.93	12.02	56.27	11.30	28.40	9.78
Brod. 24	96.53	5.14	98.13	3.80	88.93	7.23	99.93	0.67	84.53	10.24
Brod. 25	70.87	12.70	86.33	6.79	91.80	5.35	93.27	5.40	91.93	4.47
Brod. 26	81.47	11.55	95.53	6.00	90.13	6.46	96.93	4.96	90.93	8.76
Brod. 27	41.00	13.04	28.00	9.89	15.07	7.68	41.80	11.25	11.93	8.17
Brod. 28	71.20	11.79	46.20	14.00	46.27	13.43	72.80	12.16	45.13	11.04
Brod. 29	64.67	10.78	67.27	12.39	33.33	13.57	79.33	9.16	49.27	11.83
Brod. 30	14.00	8.76	17.80	9.29	27.47	9.20	23.80	9.10	0.73	2.10
Brod. 31	15.00	9.30	23.73	10.09	29.13	11.20	32.53	10.09	14.20	7.32
Brod. 32	99.27	2.10	98.33	3.05	79.80	10.49	99.80	1.14	95.33	3.73

Classe	DWT		QWT Mag.		QWT Phase		QWT		CWT	
	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.
Brod. 33	73.40	8.45	69.73	9.82	40.53	13.79	78.53	9.02	51.00	9.01
Brod. 34	99.13	2.94	98.80	2.90	100.00	0.00	100.00	0.00	97.80	6.07
Brod. 35	88.87	6.90	67.80	10.64	54.73	12.65	91.60	8.35	85.20	10.05
Brod. 36	67.20	11.00	52.80	12.81	49.00	12.90	69.67	12.22	70.60	13.27
Brod. 37	85.87	13.54	90.27	6.39	90.93	8.29	96.00	3.42	94.73	4.47
Brod. 38	60.07	12.34	89.13	5.58	79.67	8.44	94.33	4.87	94.20	5.97
Brod. 39	55.80	11.48	60.87	11.97	30.60	10.93	66.60	11.39	71.33	10.74
Brod. 40	32.53	11.92	31.27	11.48	42.80	10.60	61.00	12.30	30.33	9.91
Brod. 41	44.33	10.94	59.40	10.89	83.33	9.55	84.20	8.46	31.00	11.35
Brod. 42	34.60	10.62	37.73	7.41	30.47	10.17	48.73	8.30	47.27	7.99
Brod. 43	17.13	9.44	22.27	12.51	12.93	8.88	19.73	10.72	9.47	7.23
Brod. 44	28.67	10.91	41.40	10.90	35.07	11.24	41.73	12.45	34.00	12.48
Brod. 45	5.93	5.27	22.47	9.83	9.73	5.87	20.07	9.16	23.13	9.31
Brod. 46	88.07	10.94	85.13	10.50	95.27	5.71	94.07	10.11	74.73	12.18
Brod. 47	99.93	0.67	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 48	95.00	4.17	56.60	13.11	88.33	6.24	86.33	7.11	41.40	11.69
Brod. 49	100.00	0.00	94.87	6.06	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 50	73.67	9.16	62.27	9.34	81.33	7.28	77.13	8.11	48.07	11.38
Brod. 51	84.87	9.61	81.27	8.67	87.53	9.69	90.47	8.70	95.27	3.94
Brod. 52	39.33	12.15	81.60	12.68	100.00	0.00	99.33	3.21	98.40	6.15
Brod. 53	97.53	3.37	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 54	45.47	11.39	45.73	12.21	42.73	10.30	71.13	10.12	71.87	7.90
Brod. 55	98.87	2.52	100.00	0.00	96.87	4.49	100.00	0.00	100.00	0.00
Brod. 56	92.47	7.85	97.00	3.33	100.00	0.00	99.47	1.82	100.00	0.00
Brod. 57	97.93	5.42	91.33	7.06	47.33	12.52	98.13	4.03	100.00	0.00
Brod. 58	13.13	7.61	11.07	6.98	13.67	7.89	20.13	8.63	9.87	6.73
Brod. 59	48.33	13.04	30.87	10.15	18.60	8.49	33.47	9.62	14.67	7.34
Brod. 60	38.80	10.40	57.67	11.27	12.33	6.86	62.53	12.13	57.47	12.74
Brod. 61	38.53	11.43	27.93	11.04	23.53	10.14	42.73	11.57	38.20	12.49
Brod. 62	48.07	12.47	70.00	8.34	27.60	10.38	79.40	6.00	53.13	12.11
Brod. 63	26.87	10.66	23.20	8.55	25.80	10.71	36.00	11.80	16.60	9.64
Brod. 64	69.80	10.36	76.73	9.87	97.73	3.57	88.73	8.08	85.60	10.83
Brod. 65	94.67	3.79	98.73	2.95	99.93	0.67	99.80	1.14	90.93	10.27
Brod. 66	51.00	12.04	54.07	11.56	84.07	11.17	92.47	7.50	40.33	12.51
Brod. 67	46.13	12.16	72.60	10.55	80.93	8.48	84.93	8.13	74.67	10.42
Brod. 68	94.73	6.52	100.00	0.00	100.00	0.00	100.00	0.00	84.47	12.14
Brod. 69	44.00	10.89	47.47	10.98	35.40	8.77	55.87	11.60	53.53	9.96
Brod. 70	69.07	13.38	63.13	13.25	85.53	8.15	73.27	11.03	68.07	10.77
Brod. 71	71.40	10.98	84.07	9.70	39.13	10.28	90.13	7.67	80.60	8.53
Brod. 72	76.13	8.75	63.07	13.91	60.93	8.89	76.60	10.31	56.67	12.41
Brod. 73	20.80	9.44	21.67	9.11	10.60	7.70	35.87	11.08	26.07	9.43
Brod. 74	40.73	11.79	63.80	11.54	83.13	9.50	94.93	7.41	21.67	10.44
Brod. 75	67.87	11.19	80.00	11.33	94.87	3.77	97.40	3.40	23.20	9.55
Brod. 76	95.93	6.89	99.53	2.55	100.00	0.00	100.00	0.00	84.47	13.10
Brod. 77	98.67	2.68	96.40	3.34	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 78	96.47	5.31	83.53	6.25	93.33	5.44	88.27	6.44	87.40	6.96
Brod. 79	87.53	8.77	93.00	7.30	100.00	0.00	97.07	5.94	71.53	15.19
Brod. 80	70.73	12.50	62.13	15.91	98.33	3.33	91.13	6.07	86.80	9.19
Brod. 81	90.73	7.92	92.20	6.36	98.27	3.09	99.93	0.67	97.67	6.24
Brod. 82	100.00	0.00	99.87	0.94	99.93	0.67	100.00	0.00	84.27	7.12
Brod. 83	99.93	0.67	93.00	7.95	100.00	0.00	100.00	0.00	98.53	2.78
Brod. 84	99.20	2.89	90.47	7.17	100.00	0.00	99.53	2.37	97.73	3.57
Brod. 85	93.00	5.14	97.07	5.04	100.00	0.00	100.00	0.00	100.00	0.00
Brod. 86	42.73	13.17	27.67	10.57	56.27	10.78	63.07	11.15	86.33	9.25
Brod. 87	88.73	8.67	79.53	10.48	77.47	8.93	94.93	4.84	68.53	11.45
Brod. 88	25.40	8.35	8.87	6.36	31.33	10.82	21.67	9.01	7.33	5.65
Brod. 89	18.53	9.41	7.00	6.45	27.93	10.06	11.60	7.32	4.40	4.94
Brod. 90	33.80	10.39	18.67	9.33	56.07	11.65	46.80	12.50	28.53	9.38
Brod. 91	21.93	9.68	9.53	5.94	25.87	11.65	34.07	9.89	10.93	8.29
Brod. 92	60.47	12.44	55.67	13.18	70.80	11.08	83.13	11.70	69.73	13.18
Brod. 93	89.00	7.18	87.53	8.67	96.13	4.85	94.80	5.89	94.27	6.14
Brod. 94	51.13	14.06	82.27	8.11	76.60	11.82	94.13	5.47	90.80	7.01

Classe	DWT		QWT Mag.		QWT Phase		QWT		CWT	
	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.
Brod. 95	90.87	8.61	90.07	11.55	97.93	3.24	97.67	5.14	96.20	3.82
Brod. 96	49.47	11.57	65.73	9.98	76.87	10.91	66.87	10.31	79.47	8.24
Brod. 97	36.33	9.11	25.13	9.47	11.47	8.26	33.33	10.47	51.33	12.80
Brod. 98	21.13	9.38	22.20	9.04	38.00	11.43	32.00	13.06	10.80	7.08
Brod. 99	18.53	8.97	18.73	9.46	12.13	6.99	32.93	11.87	10.27	7.37
Brod. 100	33.07	11.83	21.53	7.63	26.67	11.25	44.60	13.19	42.53	10.15
Brod. 101	51.33	12.97	50.20	12.33	66.53	13.33	69.07	16.26	58.40	13.00
Brod. 102	46.27	15.30	52.80	12.09	42.47	14.39	68.53	13.83	36.53	12.15
Brod. 103	66.07	12.47	60.40	14.83	52.80	11.97	65.47	13.35	57.60	14.54
Brod. 104	63.13	16.78	47.00	14.03	62.27	13.71	52.73	16.08	60.13	14.74
Brod. 105	82.53	8.41	69.40	11.69	71.07	9.77	74.00	9.87	82.47	8.24
Brod. 106	69.53	16.12	54.73	13.44	67.13	14.22	56.20	12.97	79.07	10.42
Brod. 107	28.67	9.64	72.67	10.74	35.33	11.03	73.27	11.82	61.67	11.66
Brod. 108	16.00	8.04	28.07	10.65	25.00	9.30	32.00	11.13	14.73	7.66
Brod. 109	67.33	11.59	48.67	12.97	57.13	15.92	56.73	14.42	27.73	9.60
Brod. 110	51.87	15.30	61.73	14.45	48.73	15.00	64.93	15.76	60.93	9.89
Brod. 111	64.33	11.50	49.07	12.52	73.67	13.65	96.40	4.07	62.80	11.22
TC12 1	84.03	4.10	85.22	4.30	87.03	5.10	95.24	3.69	89.20	4.52
TC12 2	83.64	4.83	80.31	4.29	84.39	3.52	93.13	2.75	85.22	4.09
TC12 3	98.54	1.24	91.13	3.24	73.58	4.67	98.67	1.17	93.58	2.29
TC12 4	92.10	3.32	81.01	4.83	76.43	5.18	92.91	4.35	93.06	3.93
TC12 5	93.31	5.85	95.13	5.60	94.23	5.84	95.05	6.43	94.56	6.88
TC12 6	94.54	4.65	96.06	5.68	94.55	5.28	95.15	5.87	94.27	5.64
TC12 7	90.39	5.46	96.31	4.23	91.37	4.80	95.57	5.37	96.19	4.78
TC12 8	100.00	0.00	95.50	2.45	92.74	3.96	98.09	2.15	88.22	3.92
TC12 9	97.62	3.25	87.99	3.95	87.82	3.81	97.65	1.86	96.86	3.65
TC12 10	63.06	3.88	65.06	5.14	75.69	4.54	80.52	4.15	75.99	4.38
TC12 11	79.15	4.75	84.98	4.64	83.35	4.42	92.15	4.98	91.91	5.89
TC12 12	99.02	1.43	90.13	3.11	93.03	2.70	99.00	0.80	97.59	2.39
TC12 13	86.97	3.11	62.61	4.45	77.67	4.53	89.08	3.43	90.60	3.51
TC12 14	77.90	4.17	92.66	5.41	65.12	4.78	93.91	4.44	94.26	4.29
TC12 15	46.31	4.38	65.77	4.35	53.89	4.57	66.83	3.62	72.33	4.24
TC12 16	61.33	4.64	62.08	4.19	64.26	4.80	83.92	4.00	79.89	4.93
TC12 17	62.94	4.81	56.98	5.09	48.24	3.95	77.28	4.55	75.99	5.06
TC12 18	80.47	4.28	83.58	5.16	83.35	4.45	93.68	2.85	89.09	4.39
TC12 19	73.21	3.70	70.41	4.15	67.75	4.74	87.17	3.28	90.96	4.20
TC12 20	86.43	3.40	72.28	4.39	69.34	4.37	91.57	2.71	79.49	3.64
TC12 21	94.49	2.11	85.91	2.71	98.32	0.95	99.48	0.63	97.42	1.13
TC12 22	96.48	2.53	90.79	3.10	96.11	3.76	98.38	1.96	96.07	3.82
TC12 23	81.90	5.15	84.83	3.80	69.30	5.55	85.93	5.07	67.65	6.21
TC12 24	77.39	4.19	77.97	3.88	62.87	5.23	78.71	4.03	64.58	4.49
TC14 1	91.04	9.97	91.78	9.71	99.92	0.39	99.90	0.59	82.82	7.11
TC14 2	67.96	8.15	76.50	7.75	91.78	4.92	95.90	6.31	88.04	5.01
TC14 3	67.90	9.27	68.22	11.55	87.24	6.73	91.04	5.68	91.06	5.73
TC14 4	98.90	2.06	98.30	3.39	100.00	0.00	100.00	0.00	100.00	0.00
TC14 5	76.70	8.55	57.40	8.14	99.96	0.28	99.98	0.20	86.54	4.41
TC14 6	89.72	6.89	98.08	4.48	99.98	0.20	99.86	0.59	100.00	0.00
TC14 7	91.40	5.47	97.10	3.49	99.20	1.02	99.66	0.95	99.48	1.29
TC14 8	95.08	6.57	95.00	10.04	99.14	3.27	99.20	3.88	99.70	1.25
TC14 9	78.74	7.77	51.76	9.09	98.28	2.80	98.78	1.95	88.96	5.65
TC14 10	96.98	5.12	82.26	9.28	98.54	1.20	98.80	0.98	99.98	0.20
TC14 11	94.34	7.43	90.02	7.99	97.58	2.97	98.66	2.38	100.00	0.00
TC14 12	72.94	10.35	74.18	9.04	93.64	9.47	92.68	8.66	90.60	7.06
TC14 13	96.34	3.86	90.00	4.98	99.98	0.20	99.62	0.89	83.00	8.87
TC14 14	68.56	7.72	84.62	6.96	98.36	2.76	96.92	3.45	97.44	5.11
TC14 15	99.62	1.29	89.70	8.60	99.56	0.83	99.44	0.99	100.00	0.00
TC14 16	91.50	6.51	73.08	10.08	99.06	1.46	99.54	1.68	90.86	7.05
TC14 17	59.12	10.81	54.70	10.30	81.92	7.83	84.50	7.62	68.70	10.40
TC14 18	86.02	8.58	53.02	10.47	90.74	4.26	91.48	4.19	85.10	8.96
TC14 19	30.86	9.33	31.00	7.13	44.42	9.12	66.38	8.56	43.84	8.78
TC14 20	41.32	7.19	29.12	7.08	47.96	6.70	53.06	8.24	51.36	7.54
TC14 21	18.52	6.03	15.42	4.52	43.58	6.78	43.42	6.82	31.14	6.53

Classe	DWT		QWT Mag.		QWT Phase		QWT		CWT	
	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.	Moy.	Ec. Ty.
TC14 22	52.26	10.45	40.68	9.54	43.74	7.57	66.54	9.32	44.78	8.35
TC14 23	64.08	7.85	69.48	7.35	95.42	7.53	95.54	6.70	76.62	8.68
TC14 24	87.20	5.53	72.44	6.72	91.48	5.36	92.76	5.08	77.94	8.35
TC14 25	43.84	7.51	30.60	7.81	87.70	7.51	85.34	8.00	89.44	8.68
TC14 26	33.52	6.97	27.46	7.18	18.74	6.85	26.96	6.09	18.48	6.16
TC14 27	26.00	7.06	24.10	6.91	67.68	7.22	64.96	7.83	40.52	6.78
TC14 28	99.84	1.23	96.12	5.26	100.00	0.00	100.00	0.00	99.82	0.58
TC14 29	88.94	10.94	82.58	12.06	92.88	4.43	93.38	4.26	100.00	0.00
TC14 30	61.88	10.26	50.72	8.54	45.16	7.19	52.68	9.12	72.02	7.43
TC14 31	37.12	8.73	29.38	7.21	44.36	8.13	48.74	8.42	64.18	8.79
TC14 32	39.14	9.36	38.82	10.13	21.98	7.33	35.30	8.03	42.30	7.91
TC14 33	59.20	7.51	47.78	8.66	95.58	4.43	87.52	6.29	56.12	8.12
TC14 34	50.36	8.86	47.70	9.44	46.26	9.44	50.62	9.80	77.82	9.35
TC14 35	39.88	7.87	21.54	4.64	25.30	5.64	34.42	8.60	40.56	7.78
TC14 36	34.28	6.90	26.44	7.28	17.94	6.24	26.62	7.86	40.20	8.22
TC14 37	31.78	7.91	35.60	7.92	59.90	9.26	62.74	9.71	49.78	9.01
TC14 38	89.00	7.50	72.34	8.15	72.84	9.99	84.40	8.31	89.70	5.53
TC14 39	28.72	6.93	24.04	6.61	31.10	6.99	37.14	8.74	51.82	7.98
TC14 40	21.30	7.31	17.08	6.74	28.68	7.40	33.36	8.22	41.98	10.42
TC14 41	70.12	7.36	49.14	7.29	43.52	7.85	55.64	8.22	59.04	8.68
TC14 42	30.98	8.21	31.82	8.29	36.36	8.52	39.60	9.51	46.26	9.97
TC14 43	20.76	6.72	23.50	7.13	35.48	7.59	37.22	8.43	38.44	8.04
TC14 44	33.66	8.47	33.86	7.94	26.30	7.39	33.22	9.44	50.52	8.26
TC14 45	49.62	11.52	37.70	8.61	50.28	8.99	54.30	9.67	63.38	10.09
TC14 46	64.78	9.09	45.12	8.85	53.96	9.17	63.86	7.84	84.70	6.61
TC14 47	29.88	8.34	27.28	8.68	20.92	6.20	35.48	9.32	33.46	7.76
TC14 48	57.86	12.74	58.34	11.75	24.72	6.93	42.80	9.07	42.96	9.40
TC14 49	54.92	11.28	53.72	9.59	27.44	7.49	43.60	9.51	49.72	9.35
TC14 50	51.98	9.73	45.32	9.94	28.44	6.78	43.86	9.82	40.72	7.67
TC14 51	44.38	9.09	40.02	9.20	33.26	8.98	44.04	10.01	41.30	8.43
TC14 52	39.98	10.17	43.30	8.00	23.06	7.70	32.02	8.40	29.44	6.87
TC14 53	74.98	10.92	68.34	10.05	47.08	9.40	67.88	8.77	61.24	10.24
TC14 54	96.16	6.27	88.98	6.98	96.26	2.57	99.10	1.59	99.50	1.25
TC14 55	70.70	8.45	68.94	8.51	43.34	8.63	53.00	9.80	81.84	6.43
TC14 56	68.12	10.48	58.52	7.85	36.18	8.64	49.52	10.15	51.52	9.58
TC14 57	89.82	5.00	81.22	6.33	98.36	2.37	98.30	2.38	98.26	2.30
TC14 58	26.24	7.34	17.62	5.98	26.52	5.56	23.84	7.81	27.38	6.50
TC14 59	19.04	5.64	10.96	5.09	29.56	7.54	24.06	7.00	16.66	5.48
TC14 60	22.84	8.20	26.84	8.96	22.36	6.84	30.08	8.49	23.98	6.24
TC14 61	19.74	5.86	16.46	5.30	18.46	5.50	20.58	5.64	25.30	6.37
TC14 62	13.90	4.89	11.92	4.92	15.46	4.71	14.00	5.36	21.86	5.64
TC14 63	20.88	6.58	25.30	7.52	21.04	5.34	26.74	7.50	28.32	6.71
TC14 64	25.10	8.25	25.24	8.09	29.78	7.72	35.52	9.27	26.70	7.68
TC14 65	23.72	7.53	20.34	7.11	23.70	6.26	23.14	7.13	19.74	7.42
TC14 66	32.62	7.39	44.58	8.69	31.00	7.92	51.74	7.36	48.48	7.61
TC14 67	23.64	6.50	19.44	5.90	24.44	5.89	20.10	6.18	23.14	6.05
TC14 68	26.22	7.98	29.32	7.83	15.60	5.87	27.42	8.91	19.58	6.41

TABLE A.1 – Résultats détaillés de le procédure de classification de texture proposée.

LISTE DES FIGURES

1.1	Signal analytique 1D	17
1.2	Filtre en quadrature de Gabor h et sa réponse fréquentielle $\hat{h}$	19
1.3	Banc de filtres à échantillonnage critique 1D. Dans cet exemple la profondeur de décomposition choisie est de trois niveaux.	22
1.4	Banc de filtres « dual-tree » 1D, sur un seul niveau de décomposition.	24
1.5	Importance de la phase dans Fourier (le contraste des deux images de droite a été renforcé).	27
1.6	Représentation en ondelettes séparables et découpage de spectre associé.	30
1.7	Implantation séparable du banc de filtres dual-tree. Pour construire les sous-bandes complexes, les sorties des quatre bancs de filtres 2D doivent être combinées en deux sorties complexes $(A + D) + \mathbf{j}(B - C)$ et $(A - D) + \mathbf{j}(C - B)$	32
2.1	Transformée de Fourier quaternionique de T. Bülow. La fréquence $\omega = [0 \ 0]$ est au centre des illustrations, le contraste de A a été renforcé.	38
2.2	(b) et (c) : Parties réelle et imaginaire des <i>ondes planes orientées</i> , base de la transformée de Fourier classique. (d,e,f,g) : Parties réelle et imaginaires quaternioniques des <i>boîtes à œufs étirées</i> , atomes de la QFT. Chaque imagerie est une fonction du repère spatial (a), oscillant autour du « gris moyen ».	39
2.3	Atomes de QFT (équation (2.20)) pour différentes valeurs de ψ ($\omega = 2\pi[0.025 \ 0.075]$, $\varphi = 0$, $\theta = 0$).	40
2.4	Signal quaternionique analytique $s_Q = Ae^{j_1\varphi + j_2\theta + j_3\psi}$ associé à s . Le contraste des spectres a été renforcé.	41
2.5	Implantation séparable du banc de filtres dual-tree. Les 4 bancs de filtres 2D résultants offrent deux interprétations possibles : 1) Une décomposition par trois ondelettes quaternioniques orientées horizontalement, verticalement et diagonalement [24]. Les sous-bandes à valeurs quaternioniques sont donc $A + \mathbf{j}_1B + \mathbf{j}_2C + \mathbf{j}_3D$. 2) Une décomposition par 6 ondelettes complexes type Gabor orientées vers $k\pi/6$ [124]. Les sous bandes sont donc $(A + D) + \mathbf{j}(B - C)$ et $(A - D) + \mathbf{j}(C - B)$	45
2.6	QWT d'une image réelle. Le contraste de l'amplitude a été renforcé.	46
2.7	Haut : Signal analytique associé à un signal 2D s à bande étroite. Bas : Sous-bandes horizontales, verticales et diagonales de l'échelle de la QWT de s contenant la majorité de l'énergie. Les sous-bandes QWT sont zoomées.	47
2.8	Effet de la quantification uniforme de l'amplitude de la DWT/ QWT. L'abscisse des courbes débit-distortion est en bits par pixels, correspondant à $\log_2(K)$	52
2.9	Effet de la quantification uniforme de la phase QWT. Chaque reconstruction correspond à la quantification d'une seule des trois angles.	53
2.10	Histogrammes moyens des angles de phase QWT dans chaque sous-bande. Les courbes sont disposées comme dans l'affichage des sous-bandes QWT figure 2.6. Les coefficients ont été sélectionnés selon un seuillage sur l'amplitude avec $T = 0.02$	54
2.11	Résultats comparatifs de quantification DWT <i>vs.</i> QWT, à quantités de données égales. « bpp » (bits per pixel) est le nombre moyen de bits utilisés pour un pixel de l'image.	57

2.12	Performances de la classification basée sur la mesure de l'amplitude : résultats de validation croisée pour quelques classes des trois bases.	63
2.13	Comparaison des performances de reconnaissance pour les mesures basées sur l'amplitude DWT (blanc), l'amplitude QWT (gris clair), la phase QWT (gris foncé) et la combinaison amplitude/phase QWT (noir).	64
2.14	Performances du descripteur final QWT basé sur la combinaison amplitude/phase, en comparaison avec le descripteur DWT.	66
2.15	Performances de classification DWT/CWT/QWT	67
3.1	Analyse par tenseur de structure de l'image <i>s</i> . Méthode par gradient (1ère ligne), méthode par filtres en quadrature séparables en coordonnées polaires (2ème ligne), et gains fréquentiels des filtres utilisés (3ème ligne).	74
3.2	Analyse par filtres en quadrature orientables (première ligne), réalisée par 5 filtres, dont nous illustrons la réponse impulsionnelle (deuxième ligne, gauche). L'affichage en couleurs de la phase φ_q suit la correspondance du schéma en bas à droite.	78
3.3	Filtres en quadrature sphériques « DoP » de M. Felsberg (Voir aussi figure 3.2).	83
3.4	Analyse monogène multi-échelle par ondelettes DoP.	84
3.5	Spline polyharmonique, ondelette isotrope et sa transformée de Riesz [144]	92
3.6	Transformée en ondelettes monogène (MWT) [144] sur une image synthétique ($\gamma = 3, i \in \{-1, -2, -3\}$).	93
3.7	Schéma de décomposition en ondelettes monogènes couleur. Chaque canal couleur est décomposé par la transformée primaire symbolisée par un « bloc ψ », et la somme « $R + G + B$ » est analysée par la « partie Riesz » (blocs ψ_{r1} et ψ_{r2}).	96
3.8	MWT couleur d'images synthétiques. La partie Riesz est représentée dans le système couleur HSV avec son module dans l'intensité et l'argument (modulo π) dans la teinte. La valeur de γ est fixée à 3 expérimentalement.	97
3.9	Histogrammes (barres) des sous-bandes primaires et du module des sous-bandes de la partie Riesz ; modèles de densités de probabilités (lignes).	100
3.10	Image peppers dégradée par un bruit blanc gaussien additif (SNR=83 dB) et distribution des coefficients de la première échelle du canal rouge c^R	101
3.11	Histogramme du module de la partie Riesz à la première échelle pour l'image peppers bruitée.	102
3.12	Histogrammes du module de la partie Riesz sur trois échelles.	102
3.13	Débruitage de l'image house . Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.	104
3.14	Débruitage de l'image peppers . Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.	105
3.15	Débruitage de l'image parrot . Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.	106
3.16	Débruitage de l'image mandrill . Première ligne : Image bruitée, débruitage par ondelettes orthogonales. Seconde ligne : débruitage par ondelettes non-décimées, approche proposée.	107
4.1	Filtrage linéaire (moyennage) et non-linéaire (médian) d'une image couleur. (Fenêtre de taille 7×7).	111

4.2	Gradient d'une image couleur $s = [s^R \ s^G \ s^B]$. Les quatre premières colonnes montrent l'utilisation d'un gradient classique sur les canaux couleur de s ainsi que sur son intensité. La cinquième colonne illustre le gradient vectoriel. Dans les 5 cas, on affiche la norme (2ème ligne) et la direction (3ème ligne) du gradient. .	113
4.3	Illustration d'un signal à valeurs dans $\mathbb{R}^3$: exemple 1D. On identifie les vecteurs comme des éléments du cube RGB, en plaçant $[0, 0, 0]$ au centre, c'est-à-dire sur le « gris moyen ». Les « couleurs » observées n'en sont donc pas à proprement parler.	116
4.4	Transformée de Riesz couleur proposée.	116
4.5	Modèle oscillant de signal vectoriel 1D lié à une rupture entre deux couleurs. . .	117
4.6	Restriction de la phase à $[0; \pi/2[$, et interprétation correspondante.	120
4.7	Signal monogène couleur proposé. La saturation de φ_2 est contrôlée par A , celle de $\theta_{\mathcal{R}}$ par $\mathcal{N}_{\mathcal{R}}$. Les valeurs de $\vec{u}$ sont illustrées par la carte colorée décrite à la figure 4.8. Celles dont l'amplitude correspondante A est faible sont remplacées par un pixel gris.	122
4.8	Illustration des valeurs de l'axe $\vec{u}$ par intégration de la sphère dans le cube RGB.	123
4.9	Transformée en ondelettes monogène couleur proposée, sur une image synthétique ($\gamma = 3$, $i \in \{-1, -2, -3\}$). L'orientation améliorée par lissage du tenseur est également montrée.	126
4.10	Voir figure 4.9.	127
4.11	Distribution « type Rayleigh » de l'amplitude A_i ($\gamma = 3$).	130
4.12	Quantification uniforme des coefficients d'ondelettes monogènes couleur pour les images de test circles , mandrill et barbara	131
4.13	Décomposition monogène couleur non décimée ($\gamma = 3$) de trois rotations d'une image de test simple (19° , 28° , 45°). Légende : voir figure 4.9. Pour φ , $\vec{u}$ et θ , on ne montre que les points où A atteint un maximum local.	135
4.14	Analyse des valeurs de $\vec{u}$ dans l'expérimentation de la figure 4.13. A gauche : Histogramme des couples (α, β) , parmi ceux qui n'ont pas été « grisés ». A droite : Identification du « code couleur » associé aux valeurs principalement observées. .	136
4.15	Variance par rotation de la transformée en ondelettes monogènes couleur sans (gauche) et avec (droite) lissage du tenseur ($\gamma = 3$, lissage du tenseur gaussien avec $\sigma = 1.5$).	138
4.16	Extraction de lignes sur la transformée de Fourier d'un signal 2D.	141
4.17	Valeurs de θ pour chaque projection de Radon d'un signal quelconque de taille 127×127	141

BIBLIOGRAPHIE

- [1] P. Abry and P. Flandrin. Multiresolution transient detection. In *Proc. IEEE Int'l Symp. Time-Freq. Time-Scale Anal.*, pages 225–228, Oct. 1994.
- [2] M. D. Adams and R. Ward. Wavelet transforms in the JPEG-2000 standard. In *Proc. IEEE Pacific Rim Conf. on Comm., Comput. and Sig. Proc.*, volume 1, pages 160–163, Victoria, BC, Canada, Aug. 2001.
- [3] S. Arivazhagan and L. Ganesan. Texture classification using wavelet transform. *Pat. Rec. Lett.*, 24(9-10) :1513–1521, 2003.
- [4] A. Ayache. Some methods for constructing nonseparable, orthonormal, compactly supported wavelet bases. *Applied and Computational Harmonic Analysis*, 10(1) :99–111, 2001.
- [5] E. Bayro-Corrochano. The theory and use of the quaternion wavelet transform. *J. of Math. Im. and Vis.*, 24(1) :19–35, 2006.
- [6] E. Bedrosian. A product theorem for Hilbert transforms. *Proc. IEEE*, 51 :868–869, 1963.
- [7] A. Belaid, D. Boukerroui, Y. Maingourd, and J. F. Lerallut. Phase-based level set segmentation of ultrasound images. *Trans. Info. Tech. Biomed.*, 15(1) :138–147, Jan. 2011.
- [8] A. Benazza-Benyahia, J.-C. Pesquet, and M. Hamdi. Vector-lifting schemes for lossless coding and progressive archival of multispectral images. *IEEE Trans. Geosci. Remote Sens.*, 40(9) :2011–2024, 2002.
- [9] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proc. IEEE Int'l Conf. Comput. Vis.*, pages 433–438, London, Jun. 1986.
- [10] G. Blanchet and L. Moisan. An explicit sharpness index related to global phase coherence. In *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing (ICASSP 2012)*, 2012.
- [11] A. C. Bovik. *Handbook of Image and Video Processing*. Academic Press, May 2000.
- [12] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(1) :55–73, 1990.
- [13] F. Brackx, N. De Schepper, and F. Sommen. The two-dimensional clifford-fourier transform. *J. Math. Imaging Vision*, 26(1-2) :5–18, 2006.
- [14] P. Brodatz. *Textures : A Photographic Album for Artists and Designers*. Dover publications, New York, 1966.
- [15] T. Bülow. Hypercomplex spectral signal representation for the processing and analysis of images. *Thesis*, Aug. 1999.
- [16] T. J. Burns, S. K. Rogers, D. W. Ruck, and M. E. Oxley. Discrete, spatiotemporal, wavelet multiresolution analysis method for computing optical flow. *Opt. Eng.*, 33(7) :2236–2247, 1994.
- [17] A. Calderón and A. Zygmund. On the existence of certain singular integrals. *Acta Math.*, 1952.
- [18] J. F. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8 :679–698, 1986.

- [19] P. Carré and E. Andres. Discrete analytical ridgelet transform. *Sig. Proc.*, 84 :2165–2173, 2004.
- [20] P. Carré and P. Denis. Quaternionic wavelet transform for colour images. In *Proc. SPIE, Wavelet Applications in Industrial Processing IV*, volume 6383, Boston, Massachusetts, USA, Oct. 2006.
- [21] T. Çelik and T. Tjahjadi. Multiscale texture classification using dual-tree complex wavelet transform. *Pat. Rec. Lett.*, 30(3) :331–339, 2009.
- [22] T. Çelik and T. Tjahjadi. Bayesian texture classification and retrieval based on multiscale feature vector. *Pat. Rec. Lett.*, 32(2) :159–167, 2011.
- [23] W. L. Chan, H. Choi, and R. G. Baraniuk. Quaternion wavelets for image analysis and processing. In *Proc. IEEE Int’l Conf. on Image Processing*, pages 3057–3060, 2004.
- [24] W. L. Chan, H. H. Choi, and R. G. Baraniuk. Coherent multiscale image processing using dual-tree quaternion wavelets. *IEEE Trans. Image Process.*, 17(7) :1069–1082, Jul. 2008.
- [25] S. Charneau. Algèbre géométrique et lancer de rayons. Master’s thesis, Université de Poitiers, France, 2004.
- [26] K. Chaudhury and M. Unser. Construction of Hilbert transform pairs of wavelet bases and Gabor-like transforms. *IEEE Trans. Signal Process.*, 57(9) :3411–3425, 2009.
- [27] C. Chaux, L. Duval, and J.-C. Pesquet. Image analysis using a dual-tree M -band wavelet transform. *IEEE Trans. Image Process.*, 15(8) :2397–2412, Aug. 2006.
- [28] G. Chen and B. Kégl. Image denoising with complex ridgelets. *Pat. Rec.*, 40(2) :578–585, 2007.
- [29] N. Chenouard and M. Unser. 3D steerable wavelets and monogenic analysis for bioimaging. In *Proc. IEEE Int’l Symp. Biomed. Im. (ISBI) : From Nano to Macro*, pages 2132–2135, Chicago IL, USA, Apr. 2011.
- [30] A. Cohen and I. Daubechies. Non-separable bidimensional wavelet bases. *Revista Matemática Iberoamericana*, 9(1) :51–137, 1993.
- [31] R. R. Coifman and D. L. Donoho. Translation-invariant de-noising. Technical report, Yale University and Stanford University, 1995.
- [32] A. Croisier, D. Esteban, and C. Galand. Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. In *Proc. Int. Symp. Info., Circuits and Systems*, Patras, Greece, 1976.
- [33] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.
- [34] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vis. Res.*, 20(10) :847–856, 1980.
- [35] J. A. Davis, D. E. McNamara, D. M. Cottrell, and J. Campos. Image processing with the radial Hilbert transform : theory and experiments. *Optics Letters*, 25(2) :99–101, 2000.
- [36] P. De Rivaz and N. G. Kingsbury. Complex wavelet features for fast texture image retrieval. In *Proc. IEEE Int’l Conf. on Image Processing*, pages 109–113, 1999.
- [37] G. Demarcq. *Le signal monogène couleur : Théorie et applications*. PhD thesis, Université de La Rochelle, France, 2010.
- [38] G. Demarcq, L. Mascarilla, M. Berthier, and P. Courtellemont. The color monogenic signal : Application to color edge detection and color optical flow. *J. Math. Im. Vis. (JMIV)*, 40(3) :269–284, 2011.
- [39] G. Demarcq, L. Mascarilla, and P. Courtellemont. The color monogenic signal : a new framework for color image processing. In *Proc. IEEE Int’l Conf. on Image Processing*, 2009.

- [40] P. Denis. *Quaternions et Algèbres Géométriques, de nouveaux outils pour les images numériques couleur*. These, Université de Poitiers, Dec. 2007.
- [41] P. Denis, P. Carré, and C. Fernandez-Maloigne. Spatial and spectral Quaternionic approaches for Colour Images. *Computer Vision and Image Understanding*, 107(1-2) :74–87, 2007.
- [42] R. Deriche. Using Canny’s criteria to derive a recursively implemented optimal edge detector. *Int’l J. of Comp. Vis.*, 1(2) :167–187, 1987.
- [43] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33(1) :116–125, 1986.
- [44] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory*, 41(3) :613–627, 1995.
- [45] T. A. Ell and S. J. Sangwine. Hypercomplex fourier transforms of color images. *IEEE Trans. Image Process.*, 16(1) :22–35, 2007.
- [46] J.-C. Faugère, F. Moreau de Saint-Martin, and F. Rouillier. Design of regular nonseparable bidimensional wavelets using Gröbner basis techniques. *IEEE Trans. Signal Process.*, 46(4) :845–856, Apr. 1998.
- [47] M. Felsberg. Low-level image processing with the structure multivector. *Thesis*, 2002.
- [48] M. Felsberg and G. Sommer. The monogenic signal. *IEEE Trans. Signal Process.*, 49(12) :3136–3144, 2001.
- [49] F. Fernandes. *Directional, Shift-Insensitive, Complex Wavelet Transforms with Controllable Redundancy*. PhD thesis, Rice University, Aug. 2001.
- [50] R. M. Figueras i Ventura, P. Vanderghelynst, and P. Frossard. Low-rate and flexible image coding with redundant representations. *IEEE Trans. Image Process.*, 15(3) :726–739, 2006.
- [51] D. Fleet, A. Jepson, and M. Jenkin. Phase-based disparity measurement. *CVGIP : Image Understanding*, 53(2) :198–210, 1991.
- [52] O. Fleischmann, L. Wietzke, and G. Sommer. Image analysis by conformal embedding. *J. of Math. Im. and Vis.*, 40(3) :305–325, 2011.
- [53] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS intercommission workshop*, Interlaken, Jun. 1987.
- [54] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(9) :891–906, 1991.
- [55] D. Gabor. Theory of communication. *J. Inst. Elec. Eng.*, 93(3) :429–457, 1946.
- [56] X. Gao, T. Nguyen, and G. Strang. Theory and lattice structure of complex paraunitary filterbanks with filters of (hermitian-)symmetry/antisymmetry properties. *IEEE Trans. Signal Process.*, 49(5) :1028–1043, 2001.
- [57] I. Gelfand and G. Shilov. *Generalized functions.*, volume 1. Academic press, New York, USA, 1964.
- [58] R. Gopinath. The phaselet transform-an integral redundancy nearly shift-invariant wavelet transform. *IEEE Trans. Signal Process.*, 51(7) :1792–1805, 2003.
- [59] P.-H. Gosselin. Quaternions et images couleur. Master’s thesis, Université de Poitiers, 2002.
- [60] G. H. Granlund. In search of a general picture processing operator. *Computer Graphics and Image Processing*, 8(2) :155–173, 1978.
- [61] G. H. Granlund and H. Knutsson. *Signal processing for computer vision*. Kluwer, 1995.
- [62] A. Grossman and J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15(4) :723–736, 1984.

- [63] S. L. Hahn. Multidimensional complex signals with single-orthant spectra. *Proc. IEEE*, 80(8) :1287–1300, 1992.
- [64] S. L. Hahn. Hilbert transforms in signal processing. *Artech House, Boston, London*, 1996.
- [65] W. R. Hamilton. On Quaternions, or on a New System of Imaginaries in Algebra. *Philosophical Magazine*, 25(3) :489–495, 1844.
- [66] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. Syst., Man, Cybern.*, 3(6) :610–621, 1973.
- [67] J. P. Havlicek, D. S. Harding, and A. C. Bovik. Extracting essential modulated image structure. In *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, pages 1014–1018, 1996.
- [68] D. Helbert, P. Carré, and É. Andrès. 3-D discrete analytical ridgelet transform. *IEEE Trans. Image Process.*, 15(12) :3701–3714, 2006.
- [69] S. Held, M. Storath, P. Massopust, and B. Forster. Steerable wavelet frames based on the riesz transform. *IEEE Trans. Image Process.*, 19(3) :653–667, 2010.
- [70] D. Hubel and T. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J. Physio.*, 160 :106–154, 1962.
- [71] W. L. Hwang and S. Mallat. Characterization of self-similar multifractals with wavelet maxima. *Appl. Comp. Harm. Anal.*, 1(4) :316–328, 1994.
- [72] L. Jacques, L. Duval, C. Chaux, and G. Peyré. A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity. *Sig. Proc.*, 91(12) :2699–2730, 2011.
- [73] B. Jähne. *Digital Image Processing 6th Edition*. Springer, Berlin, 2005.
- [74] N. Joshi, S. Bond, and M. Brady. The segmentation of colorectal MRI images. *Medical Image Analysis*, 14(4) :494–509, 2010.
- [75] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet. Vector lifting schemes for stereo image coding. *IEEE Trans. Image Process.*, 18(11) :2463–2475, Nov. 2009.
- [76] G. Karlsson, M. Vetterli, and J. Kovačević. Non-separable two-dimensional perfect reconstruction filter banks. In *Proc. SPIE Conf. Visual Comm. and Im. Proc.*, volume 1001, pages 187–199, 1988.
- [77] M. Khansari, H. R. Rabiee, M. Asadi, and M. Ghanbari. Object tracking in crowded video scenes based on the undecimated wavelet features and texture analysis. *EURASIP J. Adv. Sig. Proc.*, 2008.
- [78] F. W. King. *Hilbert Transforms, volume I*. Cambridge University Press, 2009.
- [79] N. G. Kingsbury. The dual-tree complex wavelet transform : a new efficient tool for image restoration and enhancement. In *Proc. Europ. Sig. Proc. Conf. (EUSIPCO’98)*, pages 319–322, Rhodes, Sep. 1998.
- [80] N. G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Appl. Comp. Harm. Anal.*, 10(3) :234–253, 2001.
- [81] H. Knutsson and G. H. Granlund. Fourier domain design of line and edge detectors. In *Proc. Int’l. Conf. Pat. Rec. (ICPR’80)*, Miami, Florida, Dec. 1980.
- [82] H. Knutsson. *Filtering and reconstruction in image processing*. PhD thesis, Linköping University, 1982.
- [83] H. Knutsson, C.-F. Westin, and M. Andersson. Representing local structure using tensors ii. In *Proc. 17th Scand. conf. Im. analysis*, pages 545–556, 2011.
- [84] U. Köthe and M. Felsberg. Riesz-transforms versus derivatives : On the relationship between the boundary tensor and the energy tensor. In J. W. R. Kimmel, N. Sochen, editor, *Proc. Scale-Space, LNCS 3459*, pages 179–191. Springer, 2005.

- [85] J. Kovačević and M. Vetterli. Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathcal{R}^n$. *IEEE Trans. Inf. Theory, sp. iss. Wavelet Transforms and Multiresolution Signal Analysis*, 38(2) :533–555, 1992.
- [86] P. Kovessi. Image features from phase congruency. *Videre : J. Comp. Vis. Research*, 1(3), 1999.
- [87] U. Köthe. Edge and junction detection with an improved structure tensor. In B. Michaelis and G. Krell, editors, *Pattern Recognition*, volume 2781 of *Lecture Notes in Computer Science*, pages 25–32. Springer Berlin Heidelberg, 2003.
- [88] V. Labunets, A. Maidan, E. Labunets-Rundblad, and J. Astola. Colour triplet-valued wavelets and splines. In *Proc. Int'l Symp. Im. Sig. Proc. Analysis (ISPA'01)*, pages 535–540, Pula, Croatia, Jun. 2001.
- [89] K. G. Larkin. Natural demodulation of two-dimensional fringe patterns : II. stationary phase analysis of the spiral phase quadrature transform. *J. Opt. Soc. Am.*, 18(8) :1871–1881, 2001.
- [90] K. G. Larkin, D. Bone, and M. A. Oldfield. Natural demodulation of two-dimensional fringe patterns : I. general background to the spiral phase quadrature transform. *J. Opt. Soc. Am.*, 18(8) :1862–1870, 2001.
- [91] N.-E. Lasmar and Y. Berthoumieu. Multivariate statistical modeling for texture analysis using wavelet transforms. In *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*. IEEE, 2010.
- [92] W. Lawton. Application of complex valued wavelet transforms to subband coding. *IEEE Trans. Signal Process.*, 41 :3566–3568, 1993.
- [93] N. Le Bihan and S. J. Sangwine. Quaternionic spectral analysis of non-stationary improper complex signals. In *Proc. Int'l Conf. Clifford Algebras and Appl. in Math. Physics (ICCA9)*, Jul. 2011. Weimar, Germany.
- [94] J. Lebrun and M. Vetterli. High-order balanced multiwavelets : theory, factorization, and design. *IEEE Trans. Signal Process.*, 49(9) :1918–1930, 2001.
- [95] J. Li, H. Zhao, K. Jiang, X. Zhou, and X. Tong. Multiscale stereo analysis based on local-color-phase congruency in the color monogenic signal framework. *Opt. Lett.*, 35(13) :2272–2274, 2010.
- [96] J. Li, H. Zhao, C. Shi, and X. Zhou. A multi-model stereo similarity function based on monogenic signal analysis in poisson scale space. *Math. Prob. Eng.*, 2011.
- [97] J. Li, H. Zhao, X. Zhou, and C. Shi. Robust stereo image matching using a two-dimensional monogenic wavelet transform. *Opt. Lett.*, 34(22) :3514–3516, 2009.
- [98] J. M. Lilly. Modulated oscillations in three dimensions. *IEEE Trans. Signal Process.*, 59(12) :5930–5943, 2011.
- [99] J. M. Lina and M. Mayrand. Complex Daubechies wavelets. *Appl. Comp. Harm. Anal.*, 2(3) :219–229, 1995.
- [100] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int'l J. of Comp. Vis.*, 60(2) :91–110, 2004.
- [101] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11 :674–693, 1989.
- [102] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition : The Sparse Way*. Academic Press, 3 edition, Dec. 2008.
- [103] T. Maltaverne, P. Delachartre, and A. Basarab. Motion estimation using the monogenic signal applied to ultrasound elastography. In *Proc. IEEE Eng. Med. Biol. Soc. Conf. (EMBC'10)*, pages 33–36, Buenos Aires, Argentina, 2010.

- [104] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Royal Soc. London. Series B, Biological Sciences*, 207(1167) :187–217, 1980.
- [105] B. Matschkal and J. Huber. Spherical logarithmic quantization. *IEEE Trans. Aud., Sp. and Lang. Proc.*, 18(1) :126–140, Jan. 2010.
- [106] Y. Meyer. *Ondelettes et opérateurs, volume 1*. Hermann, 1990.
- [107] G. Mirajkar and B. Barbadekar. Automatic segmentation of brain tumors from MR images using undecimated wavelet transform and Gabor wavelets. In *Proc. IEEE Elec. Circ. Syst.*, pages 702–705, 2010.
- [108] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pat. Rec. Lett.*, 6(5) :303–313, 1987.
- [109] M. N. Nabighian. Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms : Fundamental relations 780–786. *Geophysics*, 49(6), 1984.
- [110] R. Neelamani, A. Baumstein, and W. S. Ross. Adaptive subtraction using complex-valued curvelet transforms. *Geophysics*, 75(4) :V51–V60, 2010.
- [111] F. Nicolier, O. Laligant, and F. Truchetet. Discrete wavelet transform implementation in fourier domain for multidimensional signal. *J. Electronic Imaging*, 11(3) :338–346, 2002.
- [112] S. C. Olhede and G. Metikas. The hyperanalytic wavelet transform. Technical Report TR-06-02, Imperial College London, Department of Mathematics, Statistics Section, 2008.
- [113] S. C. Olhede and G. Metikas. The monogenic wavelet transform. *IEEE Trans. Signal Process.*, 57 :3426–3441, 2009.
- [114] A. Oppenheim and J. Lim. The importance of phase in signals. *Proc. IEEE*, 69(5) :529–541, 1981.
- [115] S.-C. Pei and C.-M. Cheng. Novel block truncation coding of image sequences for limited-color display. In *Proc. Int’l Conf. Im. Analysis and Processing (ICIAP’97)*, volume II, 1997.
- [116] C. Perwass. *Geometric Algebra with Applications in Engineering*, volume 4 of *Geometry and Computing*. Springer Berlin Heidelberg, 2009.
- [117] F. Petit, A.-S. Capelle-Laizé, and P. Carré. Underwater image enhancement by attenuation inversion with quaternions. In *Proc. IEEE Int’l Conf. Acous., Speech, and Signal Processing*, Tapei, Tai’wan (China), 2009.
- [118] F. Petit, A.-S. Capelle-Laizé, and P. Carré. Hue-based quaternionic criterion for focused-color extraction. In *Proc. IEEE Int’l Conf. on Image Processing*, Hong-Kong, Hong-Kong, Sep. 2010.
- [119] S. Pohlig. Fourier transform phase coding of images. *IEEE Trans. Acoust., Speech, Signal Process.*, 28(3) :339–341, 1980.
- [120] E. Ravelli and L. Daudet. Embedded polar quantization. *IEEE Signal Process. Lett.*, 14(10) :657–660, Oct. 2007.
- [121] M. Riesz. Sur les fonctions conjuguées. *Math. Zeitschr*, 27 :218—244, 1927.
- [122] S. J. Sangwine. Fourier transforms of colour images using quaternion, or hypercomplex, numbers. *Electronics Letters*, 1996.
- [123] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Trans. Image Process.*, 5(11) :1582–1586, 1996.
- [124] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury. The dual-tree complex wavelet transform - a coherent framework for multiscale signal and image processing. *IEEE Signal Processing Mag.*, 22(6) :123–151, 2005.

- [125] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. Live image quality assessment database.
- [126] N. Sochen, R. Kimmel, and R. Malladi. A general framework low level vision. *IEEE Trans. Image Process.*, 7 :310–318, 1998.
- [127] R. Souillard and P. Carré. Ondelettes quaternioniques pour la classification de textures. In *Actes Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'10)*, pages 222–229, Caen France, 2010.
- [128] R. Souillard and P. Carré. Quaternionic wavelets for image coding. In *Proc. Europ. Sig. Proc. Conf. (EUSIPCO'10)*, pages 125–129, Aalborg, Denmark, Aug. 2010.
- [129] R. Souillard and P. Carré. Quaternionic wavelets for texture classification. In *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*, pages 4134–4137, Dallas, TX, USA, 2010.
- [130] R. Souillard and P. Carré. Color extension of monogenic wavelets with geometric algebra : application to color image denoising. In *Proc. Int'l Conf. Clifford Algebras and Appli. in Math. Physics (ICCA9)*, pages 15–20, Weimar, Germany, Jul. 2011.
- [131] R. Souillard and P. Carré. Color monogenic wavelets for image analysis. In *Proc. IEEE Int'l Conf. on Image Processing*, pages 277–280, Brussels, Belgium, Sep. 2011.
- [132] R. Souillard and P. Carré. Ondelettes monogènes couleur pour l'analyse d'image. In *Actes GRETSI 2011*, number ID52, Bordeaux, France, Sep. 2011.
- [133] R. Souillard and P. Carré. Quaternionic wavelets for texture classification. *Pat. Rec. Lett.*, 32(13) :1669–1678, 2011.
- [134] R. Souillard and P. Carré. Color monogenic wavelet representation based on a tensor-like use of the riesz transform : application to image coding. In *Proc. Applied Geometric Algebras in Computer Science and Engineering*, number 16, La Rochelle, France, Jul. 2012.
- [135] R. Souillard and P. Carré. A discrete approach to monogenic analysis through radon transform. In *Proc. Applied Geometric Algebras in Computer Science and Engineering*, number 40, La Rochelle, France, Jul. 2012.
- [136] R. Souillard and P. Carré. Tensor based generalization of monogenic wavelets for coherent multiscale local phase analysis of color images. In *Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing*, pages 1501–1504, Kyoto, Japon, Mar. 2012.
- [137] R. Souillard, P. Carré, and C. Fernandez-Maloigne. Vector extension of monogenic wavelets for geometric representation of color images. *IEEE Trans. Image Process. (in press)*, 2012.
- [138] J.-L. Starck, J. M. Fadili, and F. Murtagh. The undecimated wavelet decomposition and its reconstruction. *IEEE Trans. Image Process.*, 16(2) :297–309, 2007.
- [139] J.-L. Starck, F. Murtagh, and M. Fadili. *Sparse Signal and Image Processing : Wavelets, Curvelets and Morphological Diversity*. Cambridge University Press, Cambridge, UK, 2010.
- [140] E. W. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton Univ. Press, NJ, 1971.
- [141] M. Storath. Directional multiscale amplitude and phase decomposition by the monogenic curvelet transform. *SIAM J. Imaging Sciences*, 4(1) :57–78, 2011.
- [142] D. Tschumperlé and R. Deriche. Vector-valued image regularization with pdes : A common framework for different applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4) :506–517, 2005.
- [143] M. Tuceyran and A. K. Jain. *Texture analysis*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1993.
- [144] M. Unser, D. Sage, and D. Van De Ville. Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform. *IEEE Trans. Image Process.*, 18(11) :2402–2418, 2009.

- [145] M. Unser and D. Van De Ville. The pairing of a wavelet basis with a mildly redundant analysis via subband regression. *IEEE Trans. Image Process.*, 17(11) :1–13, 2008.
- [146] R. Vafin and W. Kleijn. Entropy-constrained polar quantization and its application to audio coding. *IEEE Trans. Speech Audio Process.*, 13(2) :220–232, Mar. 2005.
- [147] D. Van De Ville, T. Blu, and M. Unser. Isotropic polyharmonic b-splines : Scaling functions and wavelets. *IEEE Trans. Image Process.*, 14(11) :1798–1813, 2005.
- [148] D. Van De Ville and M. Unser. Complex wavelet bases, steerability, and the marr-like pyramid. *IEEE Trans. Image Process.*, 17(11) :2063–2080, 2008.
- [149] R. L. C. Van Spaendonck, T. Blu, R. G. Baraniuk, and M. Vetterli. Orthogonal Hilbert transform filter banks and wavelets. In *Proc. IEEE Int’l Conf. Acous., Speech, and Signal Processing*, pages 505–508, 2003.
- [150] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice Hall PTR, Englewood Cliffs, New Jersey, 1995.
- [151] J. Ville. Théorie et applications de la notion de signal analytique. *Cables et transmission*, 2A(1) :61–74, 1948.
- [152] A. Vo, S. Orintara, and T. Nguyen. Using phase and magnitude information of the complex directional filter bank for texture image retrieval. In *Proc. IEEE Int’l Conf. on Image Processing*, pages 61–64, 2007.
- [153] J.-Y. Yang, J. Xu, F. Wu, Q. Dai, and Y. Wang. Image coding using 2-D anisotropic dual-tree discrete wavelet transform. In *Proc. IEEE Int’l Conf. on Image Processing*, pages 165–168, 2007.
- [154] J. Yang, Y. Wang, W. Xu, and Q. Dai. Image coding using dual-tree discrete wavelet transform. *IEEE Trans. Image Process.*, 17(9) :1555–1569, 2008.
- [155] R. Yu and H. Ozkaramanli. Hilbert transform pairs of orthogonal wavelet bases : Necessary and sufficient conditions. *IEEE Trans. Signal Process.*, 53(12) :4723–4725, 2005.
- [156] L. Zhang, L. Zhang, Z. Guo, and D. Zhang. Monogenic-lbp : A new approach for rotation invariant texture classification. In *Proc. IEEE Int’l Conf. on Image Processing*, pages 2677–2680, 2010.
- [157] X. P. Zhang, M. D. Desai, and Y. N. Peng. Orthogonal complex filter banks and wavelets : some properties and design. *IEEE Trans. Signal Process.*, 47(4) :1039–1048, 1999.
- [158] J. Zhou, Y. Xu, and X. Yang. Quaternion wavelet phase based stereo matching for uncalibrated images. *Pat. Rec. Lett.*, 28(12) :1509–1522, 2007.

INDEX

- algèbres géométriques, 83
- Analyse multirésolution (AMR), 21
- analytique, signal, 18
- axe couleur, 117, 121

- banc de filtres monogène de M. Unser, 92
- banc de filtres orthogonal, 21
- Bedrosian, théorème de, 17
- bivecteur, 85
- boîte à œufs, 39

- Cauchy-Riemann, équations de, 79, 94
- Cayley-Dickson, 124
- Cayley-Dickson, représentation de, 36
- Clifford, algèbres de, 83
- cohérence, 72
- congruence de phase, 28

- dual-tree, 24
- dual-tree 2D, 30

- espace colorimétrique, 112

- fausses couleurs, 111
- filtre en quadrature 1D, 19
- filtres de H. Knutsson, 73
- filtres en quadrature sphériques, 82
- filtres orientables, 76
- flot optique, 46
- Fourier 1D, transformée de, 16
- Fourier 2D, transformée de, 26
- Fourier fenêtrée, transformée de, 19
- Fourier quaternionique couleur, transformée de, 37
- Fourier quaternionique, transformée de, 38
- fréquence locale, 46

- Gabor 1D, ondelettes de, 20
- Gabor 1D, transformée de, 19
- Gabor 2D, ondelettes de, 28
- Gabor quaternionique, atomes de, 42
- gradient, 70
- gradient complexe, 79, 91

- gradient couleur, 113

- harmonique, fonction, 79, 94
- Hilbert couleur, transformée de, 118
- Hilbert partielle, transformée de, 27
- Hilbert, transformée de, 17
- holomorphe, fonction, 79, 94

- i1D, 70

- k-ppv, 59

- Laplacien fractionnel, 81, 91
- lifting vectoriel, 112

- marginal, traitement, 111
- monogène couleur, signal, 94, 121
- monogène, fonction, 79, 94
- monogène, signal, 81, 87
- monogènes continues, ondelettes, 82
- monogènes couleur, ondelettes, 96, 124
- Morlet, ondelettes de, 20

- Ondelettes complexes, 23
- Ondelettes orthogonales, 21
- ondelettes séparables, 29

- parcimonie, 49, 129
- phase couleur, 96, 119
- phase quaternionique, 39
- phase restreinte, 78, 119
- Poisson, noyau de, 82
- primaire, ondelette, 92
- pseudoscalaire, 85

- Q-Shift, filtres, 24
- quaternioniques couleur, ondelettes, 37
- quaternion, 35
- quaternionique 1D, signal, 124
- quaternionique analytique, signal, 41
- quaternioniques, ondelettes, 43

- Riesz couleur, transformée de, 115
- Riesz, ondelette, 92

Riesz, transformée de, 80

segmentation de textures, 48

signal simple, 70

sous-bande couleur, 116

splines polyharmoniques, 89

STFT, 19

tenseur de structure, 70

théorème du décalage 1D, 16

théorème du décalage 2D, 26

théorème du décalage QFT, 38

validation croisée, 60

RÉSUMÉ - ABSTRACT

Titre Ondelettes analytiques et monogènes pour la représentation des images couleur

Auteur Raphaël SOULARD, laboratoire XLIM-SIC (UMR 7252), Université de Poitiers

Discipline Automatique, traitement du signal - Traitement du signal et des images

Résumé De nombreux algorithmes de traitement d'image numérique (compression, restauration, analyse) sont basés sur une représentation en ondelettes. Les outils mathématiques disponibles étant souvent pensés pour des signaux 1D à valeurs scalaires (comme le son), ils sont mal adaptés aux signaux 2D vectoriels comme les images couleur. Les méthodes les plus répandues utilisent ces outils indépendamment sur chaque ligne et chaque colonne (méthodes « séparables »), de chaque plan couleur (méthode « marginale ») de l'image. Ces techniques trop simples ne donnent pas accès aux informations visuelles élémentaires, aboutissant à des traitements qui risquent d'introduire des artefacts rectangulaires et de fausses couleurs. Notre axe de recherche se situe autour des représentations analytiques qui utilisent un modèle oscillatoire des signaux. Ces outils de traitement du signal sont connus pour être bien adaptés à la perception humaine (auditive et visuelle), et leur extension à des dimensions supérieures est un sujet encore très actif, qui révèle des propriétés intéressantes pour l'analyse de la géométrie locale. Dans cette thèse, nous faisons une revue des ondelettes analytiques existantes pour l'image en niveaux de gris (dites complexes, quaternioniques et monogènes), et nous proposons des expérimentations qui valident leur intérêt pratique. Nous définissons ensuite une extension vectorielle qui permet de manipuler facilement le contenu géométrique d'une image couleur, ce que nous validons à travers des expérimentations en codage et analyse d'image.

Mots-clés ondelettes analytique monogène quaternionique dual-tree phase couleur

Title Analytic and monogenic wavelets for color image representation

Author Raphaël SOULARD, XLIM-SIC laboratory (UMR 7252), University of Poitiers, France

Discipline Signal and image processing

Abstract Many digital image processing algorithms (compression, restoration, analysis) are based on a wavelet representation. Available mathematical tools are often designed for 1D and scalar-valued signals (e.g. sound) so ill-adapted for 2D vector signals such as images. Most methods use those tools independently on every row and column ("separable methods") of each color channel ("marginal methods") of the image. These too simple techniques cannot give access to elementary visual information, sometimes resulting in rectangular artifacts or false colors. Our topic is about analytic representations using an oscillatory model for signals. These signal processing tools are known to fit well the human perception (auditory and visual), and their extension to higher dimensions is still an active topic revealing interesting properties for local geometry analysis. In this thesis we review existing analytic wavelets for grayscale images (complex, quaternionic and monogenic) and we propose experiments that validate their practical interest. We then define a vector extension that handles well the geometric content of a color image, what we further validate through experiments of image coding and analysis.

Keywords wavelets analytic monogenic quaternionic dual-tree phase color

