



Construction et analyse de transformées discrètes en ondelettes : spécificité de l'image et de la couleur

Philippe Carré

► To cite this version:

Philippe Carré. Construction et analyse de transformées discrètes en ondelettes : spécificité de l'image et de la couleur. Traitement du signal et de l'image [eess.SP]. Université de Poitiers, 2011. tel-00652608v2

HAL Id: tel-00652608

<https://theses.hal.science/tel-00652608v2>

Submitted on 2 Jan 2012

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

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

MÉMOIRE

pour l'obtention de l'
HABILITATION À DIRIGER DES RECHERCHES
Université de Poitiers
(Faculté des Sciences Fondamentales et Appliquées)

École Doctorale : **Sciences Pour l'Ingénieur**
Secteur de Recherche : **Traitement du signal et des images**

par
PHILIPPE CARRÉ

Construction et analyse de transformées discrètes en ondelettes : spécificité de l'image et de la couleur

Soutenue 17 Novembre 2011 devant la Commission d'Examen composée de :

Ludovic MACAIRE, Professeur, Université de Lille, LAGIS	Président du Jury
Jean-Marc CHASSERY, Directeur de Recherche CNRS, GIPSALAB	Rapporteur
William PUECH, Professeur, Université Montpellier, LIRMM	Rapporteur
Frédéric TRUCHETET, Professeur, Université de Bourgogne, LE2I	Rapporteur
Yannick BERTHOUMIEU, Professeur ENSEIRB-MATMECA, IMS	Examineur
Christine FERNANDEZ-MALOIGNE, Professeur, Université de Poitiers, XLIM-SIC	Garant

Préambule

Je remercie vivement Monsieur Jean-Marc CHASSERY, Directeur de Recherche CNRS au GIPSALAB de Grenoble, Monsieur William PUECH, Professeur au LIRMM de Montpellier et Monsieur Frédéric TRUCHETET, Professeur au LE2I du Creusot, de m'avoir fait l'honneur d'accepter d'être rapporteurs de ce mémoire. J'ai eu l'immense plaisir durant ces dix années de pouvoir échanger avec eux au cours de jurys de thèses, de discussions informelles dans le cadre d'écoles thématiques ou de congrès, sur des sujets concernant la sécurité des données multimédia, l'intégration de concepts géométriques dans les ondelettes, le lien avec la dimension physique, ou tout simplement sur le métier de chercheur.

Je témoigne toute ma gratitude à Monsieur Yannick BERTHOUMIEU, Professeur à l'IMS de Bordeaux et à Monsieur Ludovic MACAIRE, Professeur au LAGIS de Lille, de m'avoir fait l'honneur de siéger à la commission d'examen de mon HDR. J'ai eu l'opportunité et le plaisir de pouvoir échanger avec eux dans différents projets scientifiques concernant la dimension couleur, son intégration dans les outils du signal et la compréhension de la représentation suivant un angle probabiliste.

Je remercie tout particulièrement Christine Fernandez-Maloigne qui me guide et m'accompagne depuis de nombreuses années dans ce métier, tout d'abord en tant que directrice de thèse, puis en tant que directrice du laboratoire et collègue dans l'équipe. Son aide, ses conseils, sa confiance et son soutien ont été sources de grandes satisfactions, tant professionnelles que personnelles. Pour tout cela je lui dois un immense merci.

Les travaux présentés dans ce document sont, pour un grand nombre d'entre-eux, le fruit d'encadrement de thèses ou de séjours post-doctoraux. Je tiens donc à remercier et à associer à ce travail tous les jeunes collègues que j'ai eu le plaisir d'encadrer : Patrice Denis, Philippe Gosselin, François Largeteau, Guillaume Lebrun, Aldo Maalouf, François Mourougaya, Alice Parisi, Frédéric Petit, Sylvain Rousseau, Raphaël Soulard et Abdul Wadood.

J'ai eu la chance de construire et d'échanger sur différents projets à la fois scientifiques et pédagogiques avec de nombreux collègues : Eric Andrès, Bertrand Augereau, Michel Berthier, Anne-sophie Cappelletti, Christian Chatellier, Philippe Gaborit, David Helbert, Philippe Meseure, Christian Olivier, Yannis Pousset, Clency Perrine, Noel Richard, Rodolphe Vauzelle et d'autres encore. Ils ont tous été moteurs dans la construction commune d'une maquette d'enseignement ou d'un article.

Je remercie tous les membres du laboratoire XLIM-SIC qui sont mes collègues de travail depuis 11 ans, et particulièrement mes collègues de l'équipe Icones qui ont accepté de m'accompagner et de me supporter comme animateur de l'équipe depuis maintenant 4 ans. Je formule des remerciements spéciaux à Françoise Perrain et Virginie Caillet-Bichet qui m'aident sur de nombreux aspects administratifs, et à Sylvie Duclaud pour tout le travail qu'elle fournit pour faire "tourner" le Master.

Je remercie également mes amis, avec qui j'ai partagé de nombreux moments conviviaux et précieux,

sportifs ou culinaires : Jean et Hélène, Michel, Jean-Louis, Aline et Benoît, Anne-Sophie et Manu, Isabelle et Laurent, Céline et Julien.

Je voulais dédier ce mémoire à mes proches disparus trop vite, ma soeur et mon père. Ce dernier m'a donné le goût de ce métier passionnant et a été source de motivation. Je souhaite également saluer ma mère qui a toujours été, qui est et restera présente à mes côtés. Je la remercie pour son soutien sans faille.

Enfin, pour terminer, mes pensées vont vers ma femme, Sylvie, et mes deux enfants, Lilou et Jules, pour les joies nombreuses qu'ils m'apportent au quotidien, leur patience face à ma soif sans fin d'altitude et d'immensité, pour leur soutien dans les moments plus difficiles, et tout simplement pour leur amour.

Table des matières

1	Introduction	8
2	Prise en compte de la géométrie dans les transformées en ondelettes	13
2.1	Rappel : la transformée de Fourier, la transformée en ondelettes et les évolutions en "Let" . . .	14
2.1.1	La transformée de Fourier	14
2.1.2	La transformée en ondelettes	15
2.1.2.1	Analyse multirésolution 2D	15
2.1.2.2	Ondelettes et bancs de filtres	16
2.1.2.3	Singularités et ondelettes	18
2.1.3	X-lets et autres	20
2.1.3.1	Transformée géométrique à fonctions d'analyses fixes	20
2.1.3.2	Transformées adaptatives	22
2.2	Transformée géométrique, une illustration numérique : la transformée Ridgelet discrète . . .	30
2.2.1	Principe de la transformée Ridgelet	30
2.2.1.1	Définition de la transformée Ridgelet	30
2.2.1.2	La transformée de Radon	31
2.2.2	L'aspect discret de la transformée Ridgelet	33
2.2.2.1	Principe	33
2.2.2.2	La transformée de Radon discrète	35
2.2.2.3	Inversion de la transformée Radon discrète	36
2.2.2.4	Transformée de Radon dans le cadre de la Fast Slant Stack	37
2.2.2.5	La reconstruction dans le cadre de la Fast Slant Stack	42
2.2.3	La stratégie de Poitiers : la transformée Ridgelet analytique ou DART	43
2.2.3.1	Définitions des droites discrètes 2-D	43
2.2.3.2	Stratégie de calcul de la transformée de Radon discrète	45
2.2.3.3	Inversion de la transformée de Radon analytique discrète	46
2.2.3.4	Illustration de la transformée de Radon analytique discrète	47
2.2.4	Transformée Ridgelet et applications	47
2.2.4.1	Calcul de la transformée DART	47
2.2.4.2	Applications de la transformée Ridgelet 2-D dans la restauration d'images	48
2.2.5	Evolution de la DART	53
2.2.5.1	Transformée locale	53
2.2.5.2	Vers une transformée Curvelet	53

2.2.5.3	DART 3-D	56
2.2.6	Discussion	57
2.3	Une transformée en ondelettes géodésiques	61
2.3.1	Filtrage bilatéral et extension à travers la distance géodésique	62
2.3.1.1	Présentation du filtrage bilatéral	62
2.3.1.2	Introduction de la notion de variété	62
2.3.1.3	Généralisation du filtrage bilatéral	64
2.3.2	Notion de distance géodésique en traitement d'images	65
2.3.3	Extraction de la topologie	67
2.3.3.1	Principe	67
2.3.3.2	Graphe des k plus proches voisins	68
2.3.3.3	Construction du graphe non orienté	69
2.3.3.4	Recherche du plus court chemin et extraction de la distance géodésique	71
2.3.4	Filtrage bilatéral géodésique	73
2.3.5	Transformée en ondelettes 2-D géodésiques	75
2.3.5.1	Lifting scheme	75
2.3.5.2	Décimation par treillis quinconces	79
2.3.5.3	Introduction de la distance géodésique	79
2.3.6	Analyse de la transformée géodésique	80
2.3.7	Application à la restauration d'image	83
2.3.7.1	Introduction de redondance	83
2.3.7.2	Application de la transformée géodésique au débruitage	84
2.3.8	Intégration de la couleur	86
3	Prise en compte de la couleur dans les outils issus du traitement du signal	90
3.1	Définitions, éléments de base et premières manipulations couleur.	91
3.1.1	L'algèbre des quaternions	91
3.1.2	Manipuler des couleurs à l'aide des Quaternions	93
3.1.2.1	Transformations de $\mathbb{R}^3$	93
3.1.2.2	Quaternions et images couleur	94
3.1.2.3	La forme de Cayley-Dickson	97
3.1.2.4	Modification couleur à partir des quaternions	98
3.1.3	Les algèbres de Clifford	99
3.1.3.1	Définition et construction	99
3.1.3.2	Un contexte calculatoire	101
3.1.4	Algèbres géométriques pour les images couleur	105
3.1.4.1	Teinte, saturation et intensité à partir d'un vecteur couleur	105
3.1.4.2	Opérations sur les caractéristiques couleur	106
3.2	Filtrage quaternionique et cliffordien : application à la détection de contours couleur	107
3.2.1	Convolution	108
3.2.2	Les filtres quaternioniques et leurs extensions pour la détection de contours couleur	109
3.2.3	Extension de l'opérateur de Sangwine par des opérations géométriques	110
3.3	Les transformées de Fourier couleur	117

3.3.1	Analyse spatio-chromatique d'images couleur	117
3.3.1.1	La transformée de Fourier spatio-chromatique	117
3.3.1.2	Notion de chemin couleur	118
3.3.2	Transformées de Fourier quaternioniques	122
3.3.2.1	Définition	123
3.3.2.2	Définition numérique de l'espace de Fourier quaternionique	124
3.3.2.3	Interprétation du spectre quaternionique	125
3.3.2.4	Applications de Fourier quaternionique	130
3.3.3	Approche fréquentielle par algèbres de Clifford pour les images couleur	131
3.3.3.1	$\mathcal{G}_2$ et images couleur	132
3.3.3.2	$\mathcal{G}_3$ et images couleur	133
3.3.3.3	Une transformée de Fourier couleur utilisant $\mathcal{G}_4$	135
3.4	Conclusion	139
4	Ondelettes analytiques	141
4.1	Signal analytique et ses extensions	141
4.1.1	Rappel sur le signal analytique 1-D	142
4.1.1.1	Lien avec l'analyse de fonction.	144
4.1.1.2	Interprétation géométrique du signal analytique	145
4.1.2	Extension de la notion de signal analytique aux images en niveaux de gris	145
4.1.3	Extension de la notion de signal analytique par transformée de Fourier quaternionique	147
4.1.3.1	Transformée de Fourier Quaternionique pour les images niveaux de gris .	148
4.1.3.2	Le signal analytique quaternionique 2D.	149
4.2	ondelettes Complexes	151
4.2.1	Principe	151
4.2.2	Déploiement numérique des ondelettes complexes	152
4.2.3	Extension 2D : Dual-Tree 2D	154
4.3	Ondelettes quaternioniques	156
4.3.1	Principe	157
4.3.2	Expression numérique de la QWT	159
4.3.3	Information associée à la QWT	160
4.3.4	Estimation de flot optique par QWT	161
4.4	L'analyse de textures par QWT	161
4.4.1	Existant sur la discrimination quaternionique	162
4.4.2	Utilisation des mesures de norme et de phase de la QWT	163
4.4.3	Utilisation de la QWT pour la classification de textures : mesures	164
4.5	Codage d'image par QWT	169
4.5.1	Principe général du codeur	169
4.5.2	Quantification	169
4.5.3	Organisation du flux en fonction du canal	172
4.6	Discussion sur la transformée QWT	173
4.7	Les perspectives : Ondelettes monogéniques couleur	173
4.7.1	Transformée de Riesz	173

4.7.2	Le signal monogène	175
4.7.3	Ondelettes Monogéniques	178
4.7.4	Introduction de la dimension couleur	181
4.7.5	Monogénique, Radon et Ridgelet	184
4.8	Conclusion	184
5	Sécurité	186
5.1	Introduction	186
5.2	Tatouage vectoriel	187
5.2.1	Introduction au tatouage d'images	187
5.2.2	Définition générale de l'algorithme vectoriel basé ondelettes	190
5.2.2.1	Insertion de la marque	190
5.2.2.2	Détection	194
5.2.3	Détection non-décimée	195
5.2.4	Définition du seuil d'acceptation	196
5.2.4.1	Préambule	196
5.2.4.2	Calcul de la probabilité de fausse alarme	197
5.2.4.3	Mise en place du seuil en fonction de P_f	198
5.2.4.4	Évolution avec la détection non-décimée	199
5.2.5	Mesures de robustesse de l'algorithme vectoriel basé ondelettes	200
5.2.5.1	Mesure de fausse détection	201
5.2.5.2	Robustesse aux différentes attaques	202
5.3	Résistance aux attaques géométriques	211
5.3.1	Attaques géométriques provoquant une translation	211
5.3.2	Solution de la littérature contre les attaques par rotation et changement d'échelle	213
5.3.3	Evolution de la méthode de tatouage vectorielle	215
5.3.4	Les détecteurs de points d'intérêt	216
5.3.4.1	Différents détecteurs mono-échelle	216
5.3.4.2	Détections des points d'intérêt multi-échelles	217
5.3.5	Description du voisinage des points d'intérêt	221
5.3.5.1	Le Jet Local	221
5.3.5.2	Le descripteur SIFT	222
5.3.6	Intégration de la mesure SIFT dans le cadre du tatouage	224
5.4	Augmentation de la robustesse par utilisation des codes correcteurs	227
5.4.1	Construction de la marque à partir de la signature : introduction des codes	228
5.4.2	Mesure de la robustesse du tatouage avec les codes correcteurs	232
5.4.2.1	Compression JPEG	232
5.4.2.2	Bruit blanc Gaussien additif	234
5.4.2.3	Attaque par filtrage passe-bas	234
5.4.2.4	Attaque par modification de la teinte	235
5.4.2.5	Attaque par modification de la Saturation	236
5.4.2.6	Attaque par modification de l'intensité	237
5.5	Conclusion	241

5.6	Perspective : la stéga-analyse par métriques de qualité	242
6	Conclusion et perspectives	246
7	Annexes	251
7.1	annexe : bancs de filtres quaternioniques	251
7.1.1	Préambule	251
7.1.2	L'existant sur les ondelettes quaternioniques	252
7.1.3	Opération de filtrage : produit de convolution	253
7.1.4	Définition du sous-échantillonnage	254
7.1.5	Définition du sur-échantillonnage	255
7.1.6	Banc de filtres quaternionique à reconstruction parfaite	256
7.1.7	Un exemple : le banc de filtres quaternioniques de Shannon	257
7.2	Annexe : Image de Test pour les techniques de tatouage	259

Chapitre 1

Introduction

A la suite de ma thèse, j'ai été nommé Maître de Conférences à l'Université de Poitiers au début du quadriennal 2000-2003, dans le département SIC (Signal Image et Communication) de l'UMR IRCOM. Ce laboratoire était structuré autour de quatre thèmes (équipes) complémentaires et transverses intitulés : "Génie logiciel pour l'infographie", "Signaux et Canaux de transmission", "Lumières, Matières, Aspect" et "Formes et structures". J'ai intégré cette dernière équipe, alors constituée de dix enseignants-chercheurs et coordonnée par E. Andrès. A ma prise de fonction, l'objectif scientifique était de développer et de renforcer les outils théoriques et numériques d'analyse d'image dans le cadre spécifique des images multicanaux. Dans ce contexte j'ai développé de nouveaux axes de recherche notamment autour de la sécurité et de l'utilisation des Quaternions pour le traitement des images couleur et j'ai établi des premières collaborations tant universitaires qu'industrielles que nous retrouverons dans la suite de ce mémoire. Au début du quadriennal suivant (2004-2007) le laboratoire, alors autonome sous le statut de FRE, et dirigé par P. Lienhardt, s'est structuré autour de trois équipes : "Modélisation géométrique et animation", "Ondes-Images-Communications" et "Images Couleur dynamiques et statiques". Cette dernière équipe, constituée maintenant de huit enseignants-chercheurs, a toujours pour objectif le traitement et l'analyse d'images multicanaux mais avec une couverture complète de la chaîne à savoir : acquisition, outils théoriques, informations complémentaires (mouvement, contexte ...) et perception. Durant cette période j'ai pris en charge l'animation du groupe concernant la définition des outils théoriques et l'intégration dans ces modèles de la dimension psychovisuelle et couleur. Dans ce cadre j'ai élargi mon spectre de collaboration et de recherche en proposant des modèles de représentation multiéchelle plus complexes intégrant d'autres outils comme la géométrie discrète ou les équations aux dérivées partielles. Enfin, depuis janvier 2008, le laboratoire SIC est devenu un département de l'UMR XLIM, institut de recherche créé à partir de plusieurs laboratoires dont l'IRCOM. Le département SIC est constitué de 3 équipes-projets : "Informatique Graphique (IG)", "SyStèmes de COMMUNICATIONS sans fil (SYSCOM)" et "Images Couleur, mOuvemeNt, reliEf et Surfaces (ICONES)". J'ai pris la direction de l'équipe ICONES. Cette dernière est constituée désormais de 12 enseignants-chercheurs permanents et 2 chercheurs associés étrangers.

Les activités de recherche du projet ICONES sont organisées autour du traitement, de la caractérisation et de l'analyse de signaux et images multi-sources et multi-composantes avec une spécificité concernant la couleur. J'ai proposé et mis en place une structuration de l'équipe suivant deux axes :

- Définition d'outils mathématiques adaptés à la manipulation de données vectorielles, et associés à l'introduction de paramètres physiques issus, d'une part, de modèles d'aspect des surfaces analysées, d'autre part, de modèles du système visuel humain.

- Mise en place d'évaluations objectives et subjectives de la qualité, à la fois des traitements développés, mais aussi des supports de reproduction des images ou vidéos couleur, selon des conditions normalisées au sein de salles dédiées.

Plus personnellement, la problématique centrale de mes travaux est l'étude de méthodes de décomposition discrète Temps-échelle. Ils s'inscrivent pleinement dans l'équipe projet ICONES, en connexion avec les autres outils fondamentaux développés dans le premier axe de l'équipe (approche variationnelle et stochastique) et intègrent les modélisations physiques et psychovisuelles étudiées dans l'autre axe de l'équipe.

Dans ce cadre, je me suis concentré, dès mon intégration dans le laboratoire sur les décompositions atomiques discrètes restructurables définies dans le domaine général des ondelettes : discrètes car la conception de décomposition directement dans un espace discret permet d'obtenir un espace de représentation numérique qui possède de bonnes propriétés (orthogonalité si possible, stabilité, rapidité de calcul, etc) ; atomiques car la décomposition d'un signal numérique sous la forme d'une somme d'éléments de base autorise une reconstruction partielle d'une des composantes du signal et permet une description simple du contenu de l'information ; restructurables car la possibilité "d'inverser" la transformation ouvre le champ des applications possibles.

Toutes ces méthodes possèdent maintenant un cadre général et notamment la transformée en ondelettes standard (FWT) qui a connu un véritable succès ces vingt dernières années car elle offre une représentation fine des images naturelles pour un coût de calcul dérisoire. Cependant, il existe encore certaines limites. Tout d'abord les performances de la méthode 2-D sont limitées par la directionnalité spatiale naïve des fonctions de bases, conséquence d'une construction séparable suivant l'horizontale et la verticale, ce qui ne permet pas une analyse directionnelle. Ensuite la transformée FWT 2-D standard échoue dans la construction d'une représentation efficace d'éléments géométriques 2-D (comme les contours) de l'image.

Afin d'obtenir une représentation efficace des informations présentes dans l'image, il faut que les fonctions de base soient directionnelles et/ou prennent en compte l'aspect géométrique des structures. De plus la généralisation de ces transformées à des signaux à valeurs vectorielles, extension autre que par une simple stratégie marginale, reste une question ouverte. Les approches variationnelles ont depuis longtemps montré les défauts d'une stratégie marginale dans le cadre de la couleur avec la création de fausses couleurs, la suppression de détails "colorés" ou tout simplement une représentation incomplète de l'information. Cette question va au-delà de la décomposition en ondelettes puisque très peu d'outils classiques du signal sont définis avec précision pour les images couleur.

C'est à toutes ces questions que j'ai tenté d'apporter des éléments de réponse durant mes dix années de recherche en tant que Maître de Conférences à travers les thèses que j'ai co-encadrées et les projets que j'ai animés ou auxquels j'ai participé. L'objectif recherché est une représentation liée à une structure de données sous forme de primitives (si possible non redondantes), facilement calculable par un algorithme, telle que l'ensemble de ces primitives permettent la reconstruction de la donnée d'origine. Il faut par ailleurs que l'on puisse ordonner ces primitives (notion de reconstruction partielle) et que l'on puisse donner un sens à ces primitives (étape de compréhension de la donnée). Dans ce cadre, mon travail va se découper suivant trois axes de recherche et un axe applicatif fort. Je les décris succinctement ci-dessous.

Approche géométrique. Dans un premier temps, nous avons étudié un outil qui permet la représentation d'images d'une manière optimale (au sens des ruptures linéaires) : la transformée Ridgelet. Cette transformée construit une représentation qui code efficacement les discontinuités linéaires 2-D comme par exemple

les contours. A ce titre, elle s'inscrit bien dans la classe des analyses multirésolution géométriques. Cette nouvelle transformée s'est développée ces dernières années à travers la restauration des images. Grâce à une collaboration avec les chercheurs en géométrie discrète du département XLIM-SIC nous avons proposé une transformée Ridgelet/Curvelets paramétrable, rapide et inversible, obtenant des résultats en termes de restauration de données 2-D très satisfaisant, notamment en vidéo. De plus, la simplicité et la rigueur dans la définition apportée par la géométrie discrète nous a permis d'étendre ce travail à toute dimension en collaboration avec un jeune collègue D. Helbert. Au delà de la transformée Ridgelet, ce processus de représentation nous permet de manipuler l'image à travers des projections 1-D et comme nous le verrons, nous trouvons différentes intersections avec d'autres approches, notamment la transformée de Riesz.

Suivant ce même but, à savoir capter la géométrie pour adapter la transformée, nous avons proposé dans le cadre d'un partenariat avec France Telecom R&D (maintenant OrangeLab) (S. Pateux) une représentation en ondelettes géodésiques qui se propose d'introduire une modélisation de la topologie dans un schéma numérique de calcul du type *Lifting Scheme*. Comme nous le verrons, ce schéma original très souple nous permet d'intégrer de nombreuses informations complémentaires. Il s'appuie sur la construction d'un graphe qui doit capter l'organisation structurelle de l'image à traiter. De ce graphe, une distance entre chaque pixel de l'image est estimée et introduite dans l'algorithme de décomposition.

Ces deux points ont notamment été traités à travers une thèse que j'ai co-encadrée, ayant pour objectif une représentation optimisée des images (G. Lebrun 2009). La géométrie détectée a aussi permis de contrôler les processus de diffusion liés aux approches par Equations aux Dérivés Partielles. Ceci a fait l'objet de la thèse de A. Maalouf (2008), et a été le fruit d'une collaboration avec un collègue de XLIM-SIC, B. Augereau.

Couleur et MultiCanaux. Comme nous l'avons mentionné, l'intégration de la "dimension" couleur est peu prise en compte actuellement dans les algorithmes modernes d'analyse Espace-Echelle (si ce n'est par une approche marginale). Mais plus généralement cette prise en compte de la dimension vectorielle des données est absente de tous les outils élémentaires issus du signal : la définition d'une transformée de Fourier couleur reste par exemple un sujet ouvert. L'objet de ce second axe a été d'examiner des approches plus globales pour l'extension des briques élémentaires issus du signal dans le cadre des images couleur. Suivant cette idée, j'ai débuté dès 2001 une étude concernant l'utilisation des quaternions. Cette recherche a concerné la redéfinition des traitements classiques couleur en termes quaternionique, comme le filtrage, la modification des caractéristiques couleur Ceci a fait l'objet de la thèse de P. Denis (2007), notamment à travers la redéfinition des manipulations couleur avec un formalisme lié à l'algèbre géométrique. Notons que l'utilisation des quaternions/ondelettes complexes en imagerie couleur constitue actuellement un sujet de recherche très ouvert qui répond aussi aux questions de directionalités évoquées dans le troisième chapitre de ce mémoire.

Ces travaux sur la redéfinition des outils pour les données vectorielles s'inscrit dans le travail mené en collaboration avec le laboratoire de Mathématiques et d'Image (MIA) de La Rochelle dans le cadre de la fédération PRIDES (Programme Régional de recherche en Images, Données et Systèmes). Cela se traduit aussi par une collaboration avec la Direction Générale de l'Armement (DGA) avec pour objectif d'étudier les techniques de *Compressed Sensing* dans le domaine de l'analyse d'images hyperspectrales pour la détection d'anomalies et d'objets et pour la fusion d'images (Thèse de S. Rousseau, débutée en 2009).

Approche orientée et notion de phase. Un dernier aspect absent de l'approche classique de la transformée en ondelettes est l'adéquation entre le système visuel humain (SVH) et l'espace transformée. Le simple

découpage horizontal, vertical et selon les deux diagonales, lié à l'approche classique, n'est pas suffisant à la fois en termes de description du signal mais aussi par rapport à ce que l'on connaît du système visuel humain. L'objectif serait de définir une représentation associée à des directions d'analyse plus adaptatives, invariante soit à la translation soit à la rotation. Dans le cadre des analyses directionnelles, la transformée en ondelettes complexes suivant l'arbre dual constitue une réponse possible. A partir de ces travaux, nous étudions une approche alternative reposant sur les travaux de Bulow et al. et Choi et al. sur le signal analytique 2-D à partir des nombres quaternioniques. Ce travail permet de mettre en place une transformée en ondelettes quaternioniques pour la compression et la classification plus en accord avec l'aspect SVH et introduisant une nouvelle information : une notion de phase. La généralisation de ce travail nous mène à la définition d'une ondelette monogénique qui peut être considérée comme l'une des décompositions les plus abouties en termes d'information signal. Nous décrivons les nombreuses questions qui restent ouvertes, et la première d'entre-elles concerne l'interprétation des informations et l'intégration de la couleur. L'objectif est de construire une information traduisant la notion de phase, si importante dans l'analyse de signaux 1-D. Ce travail est aussi un élément important de la proposition du nouveau codeur dans le cadre de l'ANR CAIMAN en collaboration avec Thalès et le sujet de thèse de R. Soulard (thèse en cours, débutée en 2009).

Les algorithmes numériques temps-échelle offrent de nombreuses solutions dans différentes problématiques. De ce fait, des collaborations ou des transferts de technologie ont été initiés et je privilégie la mise en place ou le renforcement de collaborations transversales avec des chercheurs d'autres domaines. Je me suis aussi efforcé de mener une activité contractuelle importante à travers le montage et la gestion de contrats important de recherche. Ces activités me permettent d'éprouver ma recherche dans le cadre de problématique à forte contrainte mais aussi de la valoriser à travers des brevets. Enfin cette activité contractuelle m'a permis d'obtenir le financement de nombreuses thèses (3 sur la période évaluée). L'un des exemples les plus significatifs concerne le tatouage et plus généralement la sécurité. Ceci constitue mon dernier axe de recherche.

Sécurité Cet axe a débuté dès mon recrutement en 2000 d'une manière opportune à travers une thèse CIFRE, réalisée par A. Parisi (soutenue en 2004) que j'ai encadré sous la responsabilité de C. Fernandez-Maloigne et dont le partenaire industriel était France Telecom R&D. Depuis plusieurs années maintenant, j'ai développé un travail sur la protection électronique des images et vidéos, qui a donné lieu à deux brevets sur le tatouage basé ondelettes, et intégrant des données psychovisuelles de perception des couleurs. Cette étude a nécessité une réflexion sur la meilleure stratégie possible pour modifier des coefficients d'ondelettes mais aussi le moyen de rendre le schéma robuste aux transformations géométriques de type rotation ou changement d'échelle. Nous avons pour cela étudié les descripteurs SIFT et notamment leur robustesse face aux transformations numériques. Ces travaux se sont poursuivis dans le cadre d'une collaboration avec le Département Mathématiques Informatique (DMI) de XLIM. L'objectif visé est l'intégration de différentes familles de code correcteur d'erreur afin d'améliorer la robustesse de notre stratégie de tatouage. Ceci a fait l'objet de la thèse de A. Wadood (soutenue en 2011) que j'ai co-encadrée avec P. Gaborit du département DMI de XLIM.

Au final, tous ces travaux ont le même objectif : la définition d'une représentation numérique performante en ondelettes pour l'analyse d'images couleur. Ils ont été en partie réalisés au cours des thèses de David Helbert, Alice Parisi, François Mourougaya, Patrice Denis, Aldo Maalouf, Guillaume Lebrun, Wadood Abdul, Sylvain Rousseau et Raphel Soulard, que j'ai encadrées [105, 171, 162, 62, 141, 132, 1, 196, 222].

D'autre part, ils sont le fruit de différentes discussions et collaborations avec des collègues de l'équipe ICONES et du département XLIM-SIC ainsi qu'avec des chercheurs partenaires.

Le plan de ce mémoire se structure selon les axes décrits ci-dessus. Nous profiterons du premier chapitre pour rappeler certaines transformées X-LET afin d'illustrer différentes recherches actuelles sur la représentation adaptée d'images. Ensuite ce chapitre se focalisera sur nos apports en ce qui concerne la prise en compte de la géométrie des images. Nous décrirons notre transformée Ridgelet construite à partir de la géométrie discrète et nous analyserons en détail ses différences et avantages par rapport à l'approche de référence de Stanford. Puis nous décrirons notre proposition de modélisation des données sous forme de graphes pour intégrer la topologie et la couleur dans la transformée. Nous verrons que cette approche, à travers sa souplesse, ouvre de nombreuses perspectives.

Le second chapitre présentera en détail l'apport des formalismes que sont les quaternions et les algèbres de Clifford dans la manipulation des images couleur. Ce mode de codage, qui n'a été étudié que très récemment, permet de manipuler les couleurs en utilisant des opérations algébriques. Dans ce cadre nous décrirons tout d'abord l'utilisation du formalisme des quaternions associé aux images couleur introduit par Sangwine il y a une dizaine d'années. Ensuite, nous détaillerons comment poursuivre cette approche dans le cadre plus général des algèbres de Clifford (dites aussi algèbres géométriques). Ce formalisme est illustré à travers la description d'une nouvelle approche spatiale de détection de ruptures couleur généralisant celle définie avec les quaternions. Enfin, dans la dernière partie de ce chapitre, nous discuterons de nouvelles définitions de la transformée de Fourier, adaptées aux images couleur.

Dans le troisième chapitre, nous aborderons nos travaux de recherche et d'encadrement les plus récents concernant les décompositions multiéchelles. Ils se situent à l'intersection des deux chapitres précédents. L'objectif est à nouveau de définir de nouvelles transformée en ondelettes qui vont s'efforcer de capter au mieux l'information 2-D cette fois, non pas à travers une lecture géométrique mais plutôt à travers l'extension rigoureuse 2-D d'outils venant du signal, notamment le signal analytique et le concept de phase. A travers nos travaux concernant la transformée en ondelettes quaternionique (QWT), nous montrerons que ces outils, qui restent confinés dans une petite communauté, peuvent fournir des réponses applicatives grâce à la qualité de la représentation, par exemple dans le cas d'un codeur QWT.

Pour finir ce mémoire, je propose de décrire mes travaux menés dans le cadre de la sécurité des données Multimédia. L'objectif initial était d'étudier les possibilités offertes en termes de Tatouage d'image par l'utilisation des décompositions en ondelettes et la prise en compte de la dimension Couleur. Ensuite nous montrerons comment ce travail s'est poursuivi à travers différentes évolutions. Tout d'abord nous décrirons nos propositions afin de rendre l'algorithme robuste à certaines attaques géométriques, puis nous étudierons l'intégration de différentes familles de codes afin d'augmenter la robustesse de notre schéma.

Enfin, je conclurai ce mémoire en reprenant les différentes propositions de recherche faites pour les années à venir.

Chapitre 2

Prise en compte de la géométrie dans les transformées en ondelettes

Comme nous l’avons abordé dans l’introduction, la question première qui a guidé mon travail a été de définir une représentation multiéchelle (conservant ce principe de microscope mathématique) qui prenne en compte la spécificité des images, comme par exemple la géométrie, et si possible intégrer une information vectorielle associée à la couleur.

Ce chapitre décrit deux premiers axes de recherche tentant de répondre à cette problématique, axes que j’ai initié, tout d’abord durant les années qui ont suivi mon recrutement, en collaboration avec E. Andrès (professeur à XLIM-SIC en géométrie discrète) avec par exemple un travail sur le 3-D dans le cadre de la thèse de D. Helbert, puis qui se sont poursuivis sous un autre angle dans le cadre de la thèse de G. Lebrun (en co-encadrement avec Stéphane Pateux d’OrangeLabs).

Rappelons le contexte. Nous considérons comme données à analyser des informations vectorielles définies sur une grille discrète donc une application de $\mathbb{Z} \times \mathbb{Z}$ (ou d’un sous-domaine) à valeurs dans $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ (ou d’un sous-domaine) :

$$\begin{aligned} A : [i, j] &\mapsto (A_1, A_2, A_3) \\ \mathbb{Z} \times \mathbb{Z} &\rightarrow (\mathbb{R}, \mathbb{R}, \mathbb{R}) \end{aligned}$$

Dans un premier temps, nous pouvons tout simplement construire la transformée 2-D suivant une approche séparable ce qui a pour conséquence des translations et des dilatations indépendantes suivant x et y . Toutefois, on ne peut pas se satisfaire de cette analyse, qui, comme il l’est mentionné depuis de nombreuses années, ne prend pas en compte les caractéristiques de l’image. Cette prise en compte peut se traduire par exemple par une notion d’analyse directionnelle dans le plan fréquentiel ou alors par une dimension géométrique dans la transformée. C’est ce second aspect qui a fait l’objet de mes premiers travaux et directions d’études que je vais présenter dans ce chapitre.

En effet, ces premiers travaux proposent d’intégrer la géométrie de l’image dans la décomposition en ondelettes grâce à des outils plutôt liés à la communauté informatique graphique. Notons que cette transversalité est le fruit de la structuration du laboratoire XLIM-SIC et de mes échanges avec mes collègues informaticiens. Deux axes ont été déployés dans ce cadre :

- une transformée Ridgelet définie à travers la géométrie discrète,

- une transformée géodésique basée sur des manipulations de graphes.

En préambule, nous profiterons de cette section pour introduire différentes notations que nous utiliserons dans la suite de ce mémoire et donc pour faire de brefs rappels des différentes analyses multirésolution. Puis, après avoir mis en évidence les limites de la transformée en ondelettes 2-D, nous introduirons les concepts présents dans la transformée Ridgelet et notamment sa forte connexion avec la transformée de Radon. Ensuite nous étudierons les éléments de base permettant de définir un calcul de la transformée de Radon discrète, élément fondamental de la transformée Ridgelet discrète, ce qui nous emmènera à la transformée de Radon analytique que j'ai développée. J'illustrerai la potentialité de ce travail avec le débruitage d'images, ainsi que différentes perspectives traitées ou à venir.

Dans la seconde partie de ce chapitre, nous introduirons la seconde approche géométrique de la transformée en ondelettes proposée dans le cadre de mes directions de recherche. Ce travail est né de la constatation de la rigidité de la transformée Ridgelet. L'idée a été alors de construire une décomposition la plus adaptative possible. Ce travail été notamment réalisé dans le cadre de la thèse de G. Lebrun. Il se propose de redéfinir l'analyse par ondelette à travers le concept de distance géodésique, et nécessitera pour cela de revoir différentes notions théoriques et pratiques associées aux variétés. Ce travail ouvre encore aujourd'hui de nombreuses questions et perspectives que nous discuterons.

2.1 Rappel : la transformée de Fourier, la transformée en ondelettes et les évolutions en "Let"

2.1.1 La transformée de Fourier

Soit s un signal défini sur $L^2(\mathbb{R})$ (signal d'énergie finie). La transformée de Fourier de s , qu'on note $\hat{s}$, est définie par :

$$\hat{s}(\xi) = \int_{-\infty}^{+\infty} s(x) e^{-2j\pi\xi x} dx \quad (2.1)$$

avec $\xi \in \mathbb{R}$ la fréquence. La transformée de Fourier inverse est définie par :

$$s(x) = \int_{-\infty}^{+\infty} \hat{s}(\xi) e^{2j\pi x \xi} d\xi \quad (2.2)$$

La transformée de Fourier peut se généraliser sur $\mathbb{R}^n$. Dans le cas 2-D, la transformée de Fourier de $s \in L^2(\mathbb{R}^2)$ à deux variables x_1 et x_2 est :

$$\hat{s}(\xi_1, \xi_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s(x_1, x_2) e^{-2j\pi(\xi_1 x_1 + \xi_2 x_2)} dx_1 dx_2 \quad (2.3)$$

La transformée de Fourier discrète (TFD) de s , fonction de l'espace des signaux discret de longueur finie N , est définie par :

$$\hat{s}[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} s[n] e^{-j\frac{2\pi kn}{N}} \text{ avec } n, k \in \mathbb{Z} \text{ pour } k = 0 \dots N-1 \quad (2.4)$$

Et sa formule d'inversion discrète par :

$$s[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{s}[k] e^{-j\frac{2\pi nk}{N}} \text{ avec } n, k \in \mathbb{Z} \text{ pour } n = 0 \dots N-1 \quad (2.5)$$

La TFD d'une image discrète s de taille $(N_1 \times N_2)$ vaut :

$$\hat{s}[k_1, k_2] = \frac{1}{\sqrt{N_1}\sqrt{N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} s[n_1, n_2] e^{-j \frac{2\pi}{N_1 N_2} (k_1 n_1 + k_2 n_2)} \quad (2.6)$$

La limite de l'analyse de Fourier est la dissimulation de l'information temporelle. C'est pourquoi de nouvelles représentations sont apparues, transformation cherchant à représenter conjointement l'information temporelle et fréquentielle. C'est le cas des ondelettes.

2.1.2 La transformée en ondelettes

Nous introduisons maintenant l'outil, élément central de mes recherches, qui permet l'étude de signaux ou processus à comportement non stationnaire : la décomposition en ondelettes. Elle construit une représentation qui rend compte de l'évolution au cours du temps du contenu fréquentiel du signal.

2.1.2.1 Analyse multirésolution 2D

La construction directe de bases discrètes orthonormées d'ondelettes 2-D repose sur la théorie de l'analyse multirésolution (AMR). Cette analyse permet l'étude successive des approximations lissées d'un signal 2-D dans lesquelles les détails sont progressivement supprimés. L'un des éléments fondamentaux de l'analyse multirésolution 2-D est l'introduction d'une matrice de dilatation D qui définit le "processus" de lissage lors d'un changement de résolution.

L'AMR est une séquence croissante de sous-espaces fermés $\{V_l\}_{l \in \mathbb{Z}}$ de $L^2(\mathbb{R}^2)$ qui ont les propriétés suivantes [157] :

$$\left\{ \begin{array}{l} \cap_{l=-\infty}^{+\infty} V_l = \{0\}, \cup_{l=-\infty}^{+\infty} V_l \text{ est dense dans } L^2(\mathbb{R}^2) \text{ et } V_{l+1} \subset V_l \\ s(x_1, x_2) \in V_0 \Leftrightarrow s(D^{-1}\mathbf{x}) \in V_1 \text{ avec } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ s(x_1, x_2) \in V_0 \Leftrightarrow s(x_1 - k_1, x_2 - k_2) \in V_0 \\ \exists \zeta \in L^2(\mathbb{R}^2) : \{\zeta(t_1 - \mu_1, t_2 - \mu_2)\}_{\mu_1, \mu_2 \in \mathbb{Z}} \text{ est une base "orthogonale" de } V_0 \end{array} \right.$$

Nous avons alors une séquence de sous-espaces $\{V_l\}_{l \in \mathbb{Z}}$ telle que : $\dots \subset V_1 \subset V_0 \subset V_{-1} \subset \dots$

Au fur et à mesure que l'indice d'échelle l croît, le maillage devient grossier, et les projections de s dans les espaces V_l contiennent de moins en moins d'information relative à s .

D'après la définition de l'AMR, la fonction $\zeta(D^{-1}\mathbf{x}) \in V_1 \subset V_0$ peut s'exprimer par une combinaison linéaire des vecteurs de base de V_0 , à savoir $\{\zeta(\mathbf{x} - \mathbf{k})\}_{\mathbf{k} \in \mathbb{Z}^2}$. Nous avons alors la relation bi-échelles suivante :

$$|\det(D)|^{-l/2} \zeta(D^{-1}\mathbf{x}) = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} h[k_1, k_2] \zeta_0(x_1 - k_1, x_2 - k_2) \quad (2.7)$$

On voit ainsi apparaître la notion de filtre discret h . L'information (les détails) perdue entre deux approximations successives peut être obtenue par projection du signal 2-D dans un ensemble d'espaces $\{W_l\}_{l \in \mathbb{Z}}$, appelés espaces d'ondelettes. W_l est un complémentaire de V_l dans V_{l-1} ($V_l + W_l = V_{l-1}$). Les fonctions d'ondelettes sont associées au complément orthogonal de $V_1 \subset V_0$.

En général, en dimension 2, ce complément n'est pas généré par une unique fonction d'ondelette et ses translatées. Toutefois, il existe $|\det(D)| - 1$ fonctions mutuellement orthogonales $\psi^i(x_1, x_2) \in L^2(\mathbb{R}^2)$,

$i = 1, \dots, |\det(D)| - 1$ telle que $V_0 = V_1 \oplus_{i=1}^{|\det(D)|-1} W_1^i$, où les espaces W_1^i sont générés par la suite $\left\{ \psi_{l,\mathbf{k}}^i(\mathbf{x}) = \psi(D^{-l}\mathbf{x} - \mathbf{k}) \right\}_{l=1, \mathbf{k} \in \mathbb{Z}^2}$ [157].

Comme W_1 est également contenu dans V_0 , les fonctions $\psi^i(D^{-1}\mathbf{x})$ peuvent s'exprimer par une combinaison linéaire des vecteurs de base de V_0 :

$$|\det(D)|^{-l/2} \psi^i(D^{-1}\mathbf{x}) = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} g^i[k_1, k_2] \zeta_0(x_1 - k_1, x_2 - k_2) \quad (2.8)$$

Ainsi apparaît le deuxième filtre discret (ou un ensemble de filtres discrets) g présent dans la transformée discrète.

2.1.2.2 Ondelettes et bancs de filtres

Nous présentons maintenant l'algorithme de décomposition/synthèse rapide (DWT) tel qu'il a été proposé par S. Mallat [151]. La TO s'applique à des signaux échantillonnés $c_{0,k_1,k_2} = s[k_1, k_2]$ correspondant à l'approximation d'une fonction $s(\cdot, \cdot) \in V_0$ par

$$s = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} c_{0,k_1,k_2} \zeta_{0,k_1,k_2}$$

Le signal original est donc représenté sur une base de V_0 . Appliquer une TO sur L échelles revient à représenter ce signal 2-D sur une base résultant de la somme directe :

$$V_L + W_L + W_{L-1} + \dots + W_1$$

L'algorithme de transformée est itératif et consiste à remplacer la représentation d'une composante sur V_L par une représentation sur $V_{L+1} + W_{L+1}$. On peut alors calculer rapidement les approximations et détails d'un signal à l'échelle L à partir de l'approximation à l'échelle $L - 1$. On peut en effet montrer à partir des équations (2.7) et (2.8) que :

$$\begin{aligned} c_{l,\mathbf{k}} &= \langle s, \zeta_{l,\mathbf{k}} \rangle = \sum_{\mathbf{n} \in \mathbb{Z}^2} h[\mathbf{n} - D\mathbf{k}] c_{l-1,\mathbf{n}} \\ \text{et } d_{l,\mathbf{k}}^i &= \langle s, \psi_{l,\mathbf{k}}^i \rangle = \sum_{\mathbf{n} \in \mathbb{Z}^2} g^i[\mathbf{n} - D\mathbf{k}] c_{l-1,\mathbf{n}} \end{aligned} \quad (2.9)$$

avec $\{c_{l,\mathbf{k}}\}_{\mathbf{k} \in \mathbb{Z}^2}$ approximation et $\{d_{l,\mathbf{k}}^i\}_{\mathbf{k} \in \mathbb{Z}^2, 1 \leq i < |\det(D)|}$ coefficients d'ondelettes à l'échelle l .

Si nous notons $\downarrow_D$ l'opération de décimation suivant D , nous pouvons écrire l'équation (2.9) sous la forme :

$$\begin{cases} c_l = \downarrow_D (h^- \otimes \otimes c_{l-1}) \\ d_l^i = \downarrow_D (g^{i-} \otimes \otimes c_{l-1}) \end{cases}$$

avec s^- retournement tel que $s^-[k_1, k_2] = s[-k_1, -k_2]$ et $\otimes$ produit de convolution

On constate que le calcul de l'approximation passe-bas et des coefficients d'ondelettes à l'échelle l se résume à la convolution (filtrage) des coefficients de l'approximation passe-bas à l'échelle $l - 1$ suivie d'une opération de décimation suivant D . En itérant ce processus, il est ainsi possible à partir de la séquence $\{c_{0,\mathbf{k}}\}$, avec $\mathbf{k} \in [0, N - 1]^2$, d'obtenir tous les coefficients d'ondelettes $\{d_{l,\mathbf{k}}^i\}$ avec $\mathbf{k} \in [0, \frac{N}{2^l} - 1]^2$, pour

$1 \leq l \leq L$, $1 \leq i < |\det(D)|$ avec L le nombre d'échelles de décompositions. Ce processus récursif constitue l'algorithme de TO rapide.

En plus de sa rapidité, l'algorithme présente le grand intérêt de disposer d'un algorithme rapide de reconstruction. Ce dernier reprend lui-aussi des notions de l'AMR, à savoir que la superposition des détails et approximations à une échelle l produit l'approximation à l'échelle $l-1$. En suivant le même raisonnement que pour la décomposition, on obtient que la reconstruction de l'approximation passe-bas à l'échelle l se résume à la convolution des coefficients de l'approximation passe-bas et des coefficients d'ondelettes à l'échelle $l+1$ précédée d'une opération d'interpolation suivant D . Comme pour l'analyse, ce processus peut se réitérer permettant ainsi de reconstruire la séquence initiale à partir de tous les coefficients d'ondelettes et de la dernière approximation passe-bas.

Nous venons de voir que l'algorithme de décomposition-reconstruction repose sur une application successive de filtrage et son implantation pratique se fait à l'aide d'un banc de filtres [151]. Nous présentons sur la figure 2.1 un exemple d'analyse à partir d'un banc de filtres 2-D à 4 canaux. Ce schéma d'analyse correspond à une dilatation séparable qui reste, comme nous l'avons indiqué dans l'introduction, la référence.

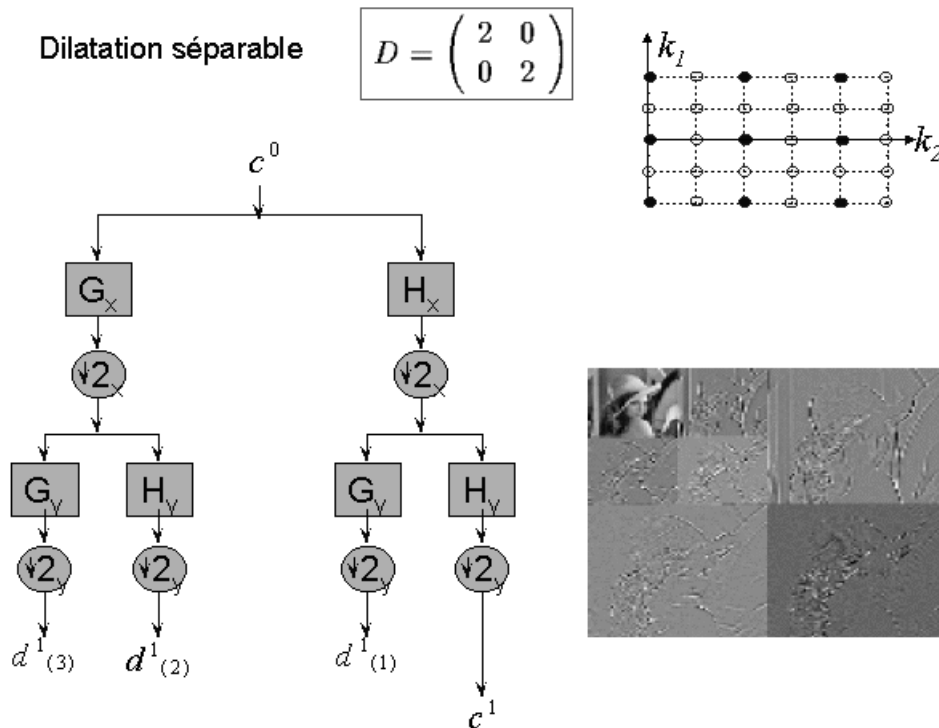


FIGURE 2.1 – Banc de filtres 2-D séparables

Nous présentons sur la figure 2.2 l'exemple le plus classique d'extension 2-D à partir d'un banc de filtres 2-D non séparables à 2 canaux. Ce schéma d'analyse correspond à une dilatation en quinconce.

Comme nous l'avons vu, les ondelettes permettent de représenter les détails perdus lors du passage d'une résolution à une résolution plus grossière. La reconstruction d'une image à partir de ses coefficients en ondelettes prend une signification intuitive évidente : l'image, à sa résolution la plus grande, est égale à la somme d'une version floue, et des détails apparaissant à des échelles différentes, c'est à dire à des résolutions différentes. De ce fait, la compression par ondelettes a été un réel progrès dans le domaine des images. En

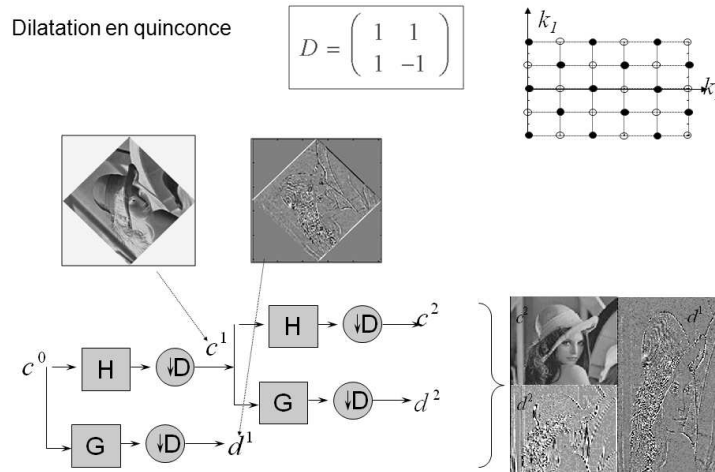


FIGURE 2.2 – Banc de filtres 2-D non-séparables suivant un schéma en quinconce

effet, la raison essentielle tient à la nature même des images. La majorité des informations auxquelles nous sommes sensibles, se trouve dans ce que l'on appelle les contours de l'image : par exemple les bords des objets, des visages,... . Ces contours sont des régions où l'intensité de l'image varie brutalement, et les "quelques" coefficients d'ondelettes correspondants vont être d'amplitude importante, y compris aux petites échelles. Or, une image contient en général relativement peu de contours, et est régulière (lentement variable) sauf au voisinage des contours. Par conséquent, beaucoup de coefficients d'ondelettes sont faibles (surtout aux petites échelles), et peuvent être négligés sans que cela entraîne de distorsion visible sur l'image.

Cependant, et c'est ce qui constitue le coeur de ce chapitre, le codage précisément du contour par les ondelettes n'est pas optimum, comme nous allons le voir dans le paragraphe suivant.

2.1.2.3 Singularités et ondelettes

La transformée en ondelettes est bien adaptée pour représenter les discontinuités ponctuelles. En effet, la fonction d'ondelette, élément de base de la décomposition, est une fonction à support temporel fini. Tous les autres éléments de la base sont fabriqués par translation et dilatation de cet élément de base et sont donc aussi à support temporel fini.

Si l'on a un signal 1-D dont l'énergie est concentrée en un point (discontinuité ponctuelle), du fait de la localisation temporelle de la base, à chaque échelle nous n'aurons que quelques coefficients d'ondelettes significatifs. Notons que concrètement le nombre de coefficients significatifs va dépendre de la base d'ondelettes utilisée.

Pour un signal de longueur N le nombre maximal d'échelle est $\log_2(N)$, donc pour une singularité la transformée en ondelettes fournit (en simplifiant) $\mathcal{O}(\log_2(N))$ coefficients significatifs. Si l'on compare à la capacité nécessaire pour stocker le signal de départ $\mathcal{O}(N)$, on peut conclure que les ondelettes sont bien adaptées pour coder les singularités. Nous illustrons ce propos sur la figure 2.3.

De cette illustration nous en déduisons que la transformée en ondelettes 2-D simplement obtenue par produit va coder correctement les singularités. Toutefois, cette construction n'est pas très flexible et ne permet pas de généraliser la propriété de coder "optimalement" les ruptures 2-D que sont les contours.

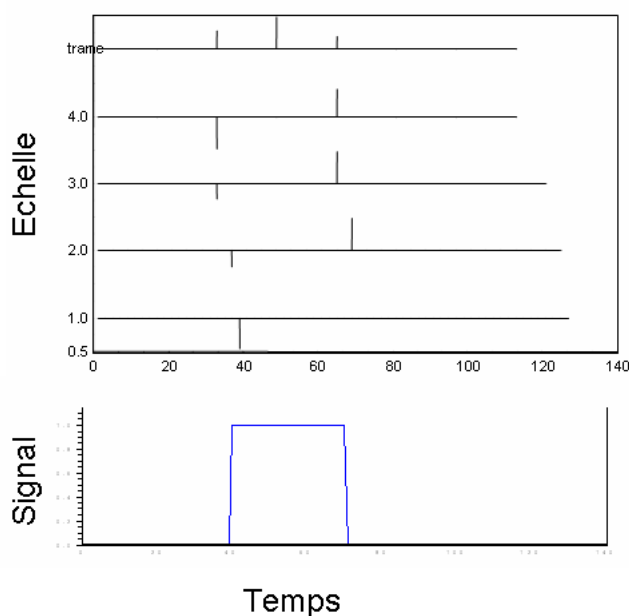


FIGURE 2.3 – Décomposition d'une discontinuité 1-D

La figure 2.4 illustre la décomposition suivant une base d'ondelettes séparables d'un objet 2-D élémentaire. On constate que les ondelettes détectent bien les points isolés de contour mais pas le contour dans sa globalité (par exemple son organisation). Si l'on reprend la propriété de codage des singularités des ondelettes on peut dire approximativement que à chaque échelle l le nombre de coefficients significatifs permettant de stocker l'unique contour est proportionnel à 2^{-l} fois la longueur du contour, ce qui n'est pas pleinement satisfaisant. De plus du fait de la construction de la base par produit tensoriel, les coefficients associés au même contour sont dispersés sur différents plans.

Pour conclure, on constate que du fait de leur construction les ondelettes 2-D sont plutôt désignées pour coder les singularités dans les directions horizontales et verticales et n'expriment pas les connexions qui existent entre les points d'un même contour. Ce problème est indépendant de la base d'ondelettes choisie (régularité de la fonction analysante).

Les images naturelles vont donc être difficiles à représenter à l'aide d'une base d'ondelettes, à la fois à cause des contours et des textures. Pour gagner en efficacité pour les images complexes, il est nécessaire soit d'utiliser des représentations rigides mais prenant mieux en compte les particularités d'une image, soit des représentations adaptatives.

Puisqu'une grande partie de mes travaux va s'inscrire dans ces transformées multirésolution pour l'image, je propose un rapide bestiaire des transformées X-LETs, certainement incomplet étant donné la richesse des propositions de ces dernières années. Notons que pour cet état des lieux nous nous appuyons sur le tour d'horizon déjà fait durant la thèse de G. Lebrun. Très récemment un panorama des différentes décompositions a été fait dans [114].

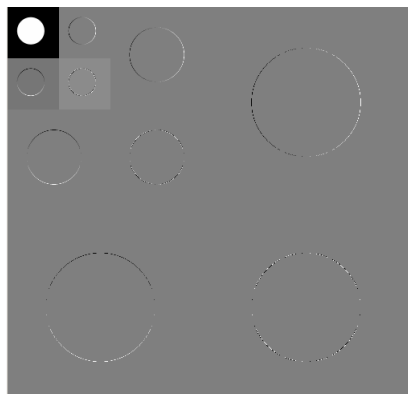


FIGURE 2.4 – Coefficients d'ondelettes d'une image représentant un cercle. Chaque image de détails est normalisée.

2.1.3 X-lets et autres

Toutes ces transformées récentes ont pour objectif de prendre en compte le "motif" formé par les singularités du signal. Cette prise en compte se fait au travers de l'ensemble du dictionnaire de fonctions d'analyse. Nous verrons que certaines d'entre elles (les transformées adaptatives) s'accompagnent de méthodes de détection explicite de ce motif avant de calculer la transformée en ondelettes.

Nous débutons notre description des transformées géométriques par les transformées basées sur un dictionnaire de fonctions définies indépendamment du contenu de l'image. Notons que nous n'abordons pas les évolutions directes du banc de filtres comme les paquets d'ondelettes ou les filtres 2-D non séparables, stratégies très proches des bancs de filtres présentés dans la section précédente.

2.1.3.1 Transformée géométrique à fonctions d'analyses fixes

Le principe de ces transformées consiste à projeter le signal à analyser dans une base de fonctions fixes, qui ne s'adaptent pas au contenu analysé.

Ridgelet La transformée Ridgelet (et son évolution la transformée Curvelet) [75], [38] correspond à la construction d'un dictionnaire de fonctions d'analyse s'appuyant sur la transformée de Radon : les fonctions d'analyse résultantes correspondent à la mise en valeur des ruptures linéaires (puis curvilignes) dans l'image. Nous ne détaillerons pas plus le principe car le premier travail que nous présenterons se place dans cette famille de bases.

Transformée Contourlet Cette transformée proposée par Do *et al.* [69] peut se voir comme une variante de la transformée Curvelet en utilisant à la place de la transformée Ridgelet un banc de filtres directionnels proposé par Bamberger *et al.* [15] ((b) figure 2.5), chaque échelle étant définie à partir d'une pyramide gaussienne de Burt *et al.* [36] ((a) figure 2.5).

La redondance de cette transformée est moins importante que celle de la transformée Curvelet. Alors que la transformée Curvelet se place sur des applications de type débruitage, cette transformée se place plus sur des applications de compression (en particulier avec la version CRISP-contourlets [139] qui est

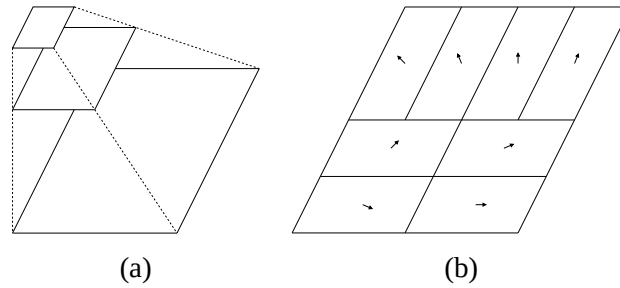


FIGURE 2.5 – (a) Pyramide gaussienne. (b) Traitement des différentes projections par la transformée en Ondelettes.

à échantillonnage critique). Il existe également une extension de la transformée pour l'étude des données organisées selon 3 dimensions avec les Surfacelets [140].

Transformée Beamlet La transformée Beamlet de Huo et Donoho [73] propose une extraction de l'ensemble des "beams" qui correspondent à tous les segments possibles entre deux pixels de l'image traitée (soit n^2 pour une image à n pixels). Les "beamlets" (figure 2.6(a)) correspondent à une limitation de l'ensemble des "beams" par le biais d'une segmentation dyadique de l'image à plusieurs échelles (figure 2.6(b)). Le principe est de ne choisir que les "beams" dont les extrémités appartiennent à la limite d'une des zones dyadiques.

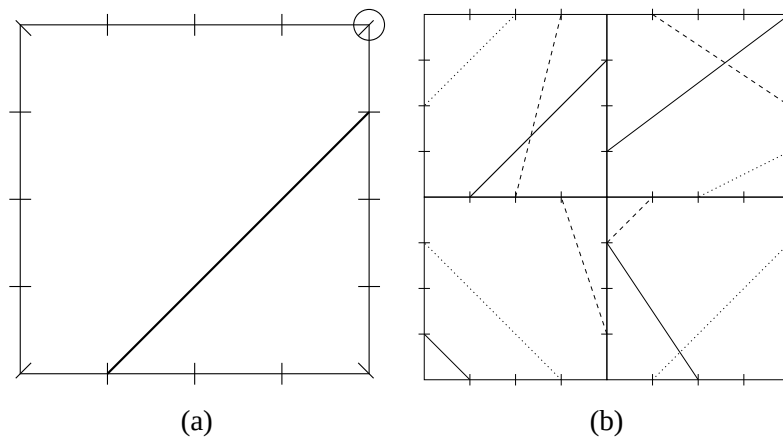


FIGURE 2.6 – (a) La beamlet est codée par deux sites du contour que l'on peut indexer dans le sens horaire en partant du coin nord-est. (b) Un ensemble de quelques beamlets.

Les coefficients de la transformée Beamlet correspondent à la somme des valeurs des pixels le long du segment. Cette transformée permet également la mise en valeur des ruptures linéaires dans l'image à partir d'une analyse multi-échelle.

Ondelettes - Triangulation Nous présentons une dernière transformée en ondelettes que nous pouvons considérer comme une transition entre une transformée adaptative et une transformée à bases fixes car il n'y a pas à proprement parler d'identification préliminaire des singularités mais le dictionnaire cherche à s'adapter aux caractéristiques de l'image. La transformée [115] est basée sur la triangulation d'une image niveau de gris considérée pour le coup comme une surface dans un espace à trois dimensions. Nous pouvons

noter un article récent utilisant cette approche pour développer un algorithme de compression rivalisant avec JPEG2000 [184].

L'information décrivant le dictionnaire de fonctions d'analyse de la transformée réside dans la topologie de l'ensemble d'échantillons qui est mise en valeur par la triangulation. L'avantage de cette nouvelle formulation est que la triangulation permet à la transformée de s'adapter aux singularités : nous avons par exemple une plus grande concentration de triangles pour les parties plus riches en détails de la surface (figure 2.7(e)).

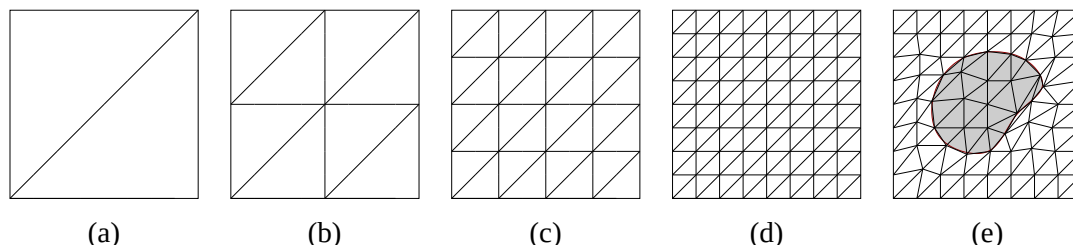


FIGURE 2.7 – Triangulation du signal. (a-d) Construction progressive (multi-échelle) d'une zone uniforme. (e) Triangulation pour un signal régulier par morceaux.

Cette approche très particulière de la transformée en ondelettes part de l'échelle la plus grossière pour redescendre vers les coefficients de détails. Elle s'appuie sur un schéma de subdivision de la triangulation de base. Chaque arête de chaque triangle est subdivisée par l'ajout d'un sommet (figure 2.7 (b-d)). Le coefficient d'ondelette résultant de cette subdivision correspond à la distance entre le milieu de l'arête et le sommet réel qui lui correspond dans la surface étudiée.

Cette transformée peut se voir comme une transition vers les transformées en ondelettes adaptatives parce qu'elle s'adapte aux données du signal (ce que les auteurs soulignent). Notons que des approches récentes [184] suggèrent d'inclure la distance géodésique dans le processus de maillage. Le deuxième volet de notre travail présenté dans ce chapitre s'approche par certains aspects de cette stratégie.

Nous proposons maintenant la description de plusieurs transformées en ondelettes adaptatives.

2.1.3.2 Transformées adaptatives

Ces transformées comme nous l'avons fait remarquer dans le paragraphe sur la transformée par triangulation, s'appuient sur un ensemble de paramètres permettant la construction du dictionnaire de fonctions d'analyse et de synthèse. Elles se retrouvent par conséquent systématiquement accompagnées d'une ou plusieurs méthodes destinées à construire l'ensemble de ces paramètres.

Transformée Wedgelet Les wedgelets [74] correspondent à des fonctions constantes par morceaux ("Horizon Class of functions"). La première phase de la transformation correspond à l'identification des "morceaux" de fonctions. Cette procédure itérative s'appuie sur un Quadtree. On cherche dans chaque bloc du Quadtree une rupture *edgel* coupant la zone en deux morceaux de fonctions constantes. Si il y a plus de deux morceaux, on sépare dyadiquement le bloc et on continue la récursion. Si la zone ne contient qu'un morceau, on indique une absence d'arêtes et cela constitue la condition d'arrêt de la récursion. Si le bloc contient une singularité suffisamment simple de forme pour être décrite par un *edgel*, on arrête également la récursion avec codage de la forme pour l'échelle en cours. Globalement la transformée s'obtient en cherchant à minimiser la différence entre l'ensemble des données traitées et sa projection sur l'ensemble des

edgels. Après la constitution de ce dictionnaire les coefficients de wedgelets vont alors définir la valeur de chaque morceau constant (figure 2.8(b)).

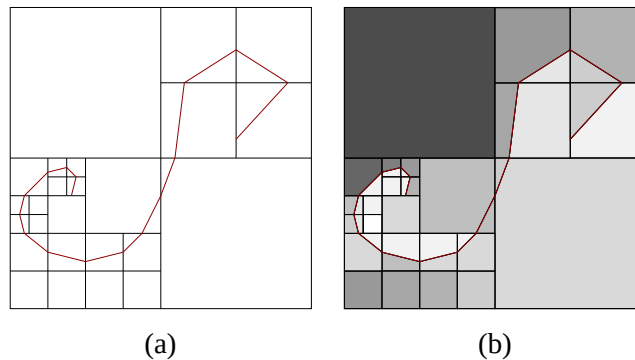


FIGURE 2.8 – (a) Dictionnaire d'*edgels* trouvé pour l'image en cours. (b) Transformée wedgelets : zones uniformes mises en évidence.

Cette transformée adaptative est particulière puisqu'elle est conçue pour les simples fonctions constantes par morceaux. Néanmoins, en combinaison avec la transformée en ondelettes 2-D elle permet une amélioration de la compression des images réelles [248] (stratégie classique de prise en charge séparée des textures et des zones uniformes de l'image).

La transformée Wedgelet illustre bien le principe des transformées adaptatives : en construisant un dictionnaire nous cherchons à adapter au mieux les fonctions d'analyse à notre signal.

Dans un autre registre, plutôt que de traiter séparément zones uniformes et textures, des approches proposent de modifier la transformée pour que le filtre d'analyse (et de synthèse) s'adapte aux singularités de l'image. Nous allons dans un premier temps décrire des transformées qui font le choix de déformer le signal pour l'adapter à la transformée en ondelettes 2-D, c'est le cas de la transformée Bandelette.

Transformée Bandelette première génération Les bandelettes de première génération [131], [130] s'appuient sur une description des contours (figure 2.9(a)) de l'image pour localiser, les singularités. Ensuite plutôt que d'adapter les fonctions d'analyse aux singularités de l'image, la transformée adapte des zones de l'image à la transformée en ondelettes séparables 2D (avec quelques adaptations pour permettre la prise en compte de la forme non rectangulaire des zones). Afin de construire ces zones à partir des contours détectés dans l'image, l'algorithme de Grassfire proposé par Blum H. [26] (figure 2.9(b)) est utilisé. Ensuite, toujours à partir de l'information des contours, des déformations sont associées à chaque zone pour permettre son adaptation aux familles d'ondelettes séparables 2D (figure 2.9(c)). Après la prise en compte des pixels des zones déformées (correspondant aux coefficients de bandelettes) il reste un ensemble de pixels à coder. Ces pixels sont pris en charge avec l'estimation de l'image initiale uniquement à partir des coefficients de bandelettes. La reconstruction parfaite est assurée par le codage du résidu entre image initiale et image estimée.

Notons que durant le travail de thèse de G. Lebrun que j'ai encadré [132], nous avons mené une étude algorithmique pour proposer des solutions numériques efficaces [133].

Malgré l'utilisation du Grassfire, la transformée souffre d'une légère redondance qui ne l'empêche pas d'obtenir de bons résultats en compression. Afin de rester sur une approche de déformation par zones du signal, et toujours à partir de la prise en compte des contours de l'image, une deuxième génération de

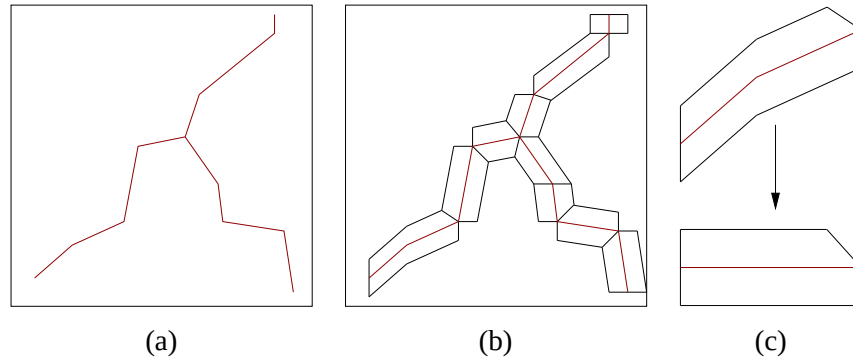


FIGURE 2.9 – (a) Détection des contours. (b) Construction des zones des pixels "bandelettes" par grassfire. (c) Déformation de la zone.

transformée en Bandelettes a été proposée par G. Peyre et al.[182].

Transformée bandelettes deuxième génération La transformée en Bandelettes de seconde génération [182] s'appuie sur une segmentation par quadtree (figure 2.10(a)) des différentes bandes de fréquences de la transformée en ondelettes un peu à la manière de la transformée en Wedgelets. Les singularités détectées dans chaque bloc du Quadtree sont prises en charge par le ré-ordonnancement préalable des coefficients, puis la transformation par ondelettes du signal 1D obtenu. Le passage de (b) vers (c) de la figure 2.10 illustre la déformation puis le parcours adapté des coefficients de la zone. Le réordonnancement s'appuie sur les lignes de flux parallèles à la singularité détectée dans la feuille dont la forme peut être plus ou moins complexe. Ce réordonnancement s'inspire de la transformée de Alpert [6].

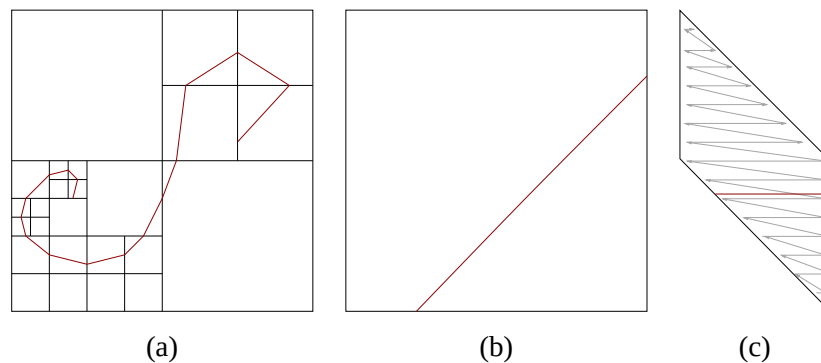


FIGURE 2.10 – Bandelettes de deuxième génération. (a) Description des singularités dans le domaines des coefficients d'ondelettes séparables. (b) Choix d'une feuille à traiter. (c) Déformation, puis ré-ordonnancement.

Cette fois, le nombre de coefficients de transformée correspond exactement au nombre d'échantillons du signal étudié. La concentration de l'énergie plus importante que pour la transformée en ondelettes permet encore une fois d'améliorer les résultats de compression.

Cette transformée a été utilisée dans le cadre de la thèse de A. Maalouf [141] que j'ai co-encadrée avec B. Augereau sous la direction de C. Fernandez. Dans ce cadre, deux modèles ont été construits pour la problématique de traitement des séquences vidéo : le premier a pour objectif le Mosaicing des séquences vidéo et le second l'Inpainting. Les deux modèles se basent sur la segmentation en quadtree de la transformée

en bandelettes afin d'étudier les similarités entre les différentes zones des trames de la séquence. Les carrés dyadiques qui ont des propriétés géométriques similaires sont collés l'un à côté de l'autre pour former une mosaïque finale ou pour combler les zones manquantes.

Nous présentons un exemple d'Inpainting réalisé dans le cadre de la thèse de A. Maalouf. La figure 2.11 présente la séquence vidéo à modifier. L'utilisateur initialise un Snake sur la zone qu'il veut faire disparaître[145], ensuite cette zone est segmentée et suivie à travers la séquence (figure 2.12). Enfin la zone est comblée par copie de blocs à propriété géométrique similaire (figure 2.13).

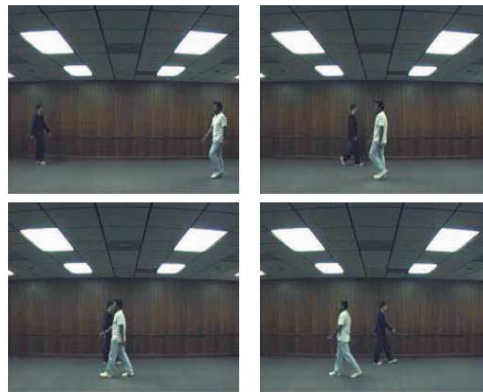


FIGURE 2.11 – Séquence vidéo d'exemple

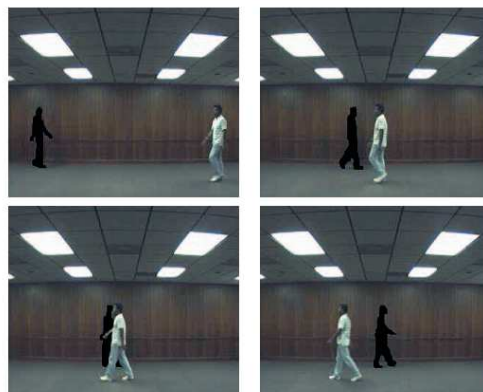


FIGURE 2.12 – Sélection de la zone à combler par intervention de l'utilisateur et d'un Snake Fovéal[141]

Ce travail a notamment été publié dans la revue *IEEE Transactions on Geosciences and Remote Sensing* [147] dans le cadre de la suppression de nuages pour des images satellitaires.

Ondelettes - Champ de déformation Le Guen *et al.* [129], [128] s'appuient quant à eux sur une déformation de l'image à partir d'un maillage quadrangulaire (figure 2.14(a)). Le but de cette déformation est, dans la même philosophie que pour les bandelettes, d'adapter l'image à la transformée en ondelettes 2D séparables. Cette approche de l'adaptation de la transformation par déformation préalable de l'image a un

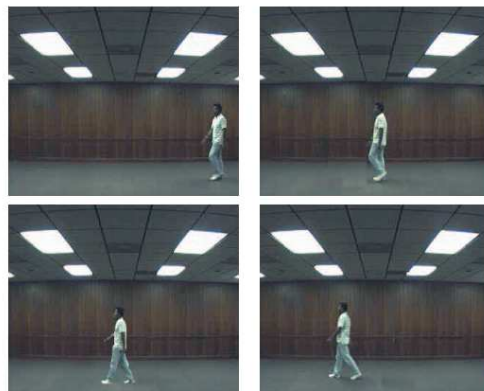


FIGURE 2.13 – Inpainting de la séquence vidéo par caractérisation de la géométrie par Bandelette

avantage par rapport aux transformées bandelettes : il n’y a pas de segmentation à gérer donc il y a beaucoup moins de problèmes de superposition du support des fonctions d’analyse avec un bord.

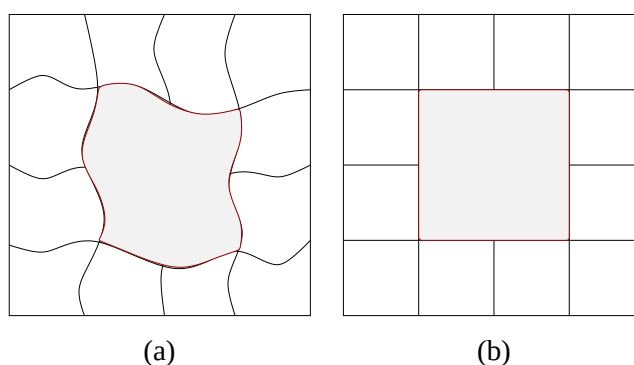


FIGURE 2.14 – Champ de déformation. (a) Maillage quadrangulaire adapté aux caractéristiques de l’image. (b) Rectification du maillage quadrangulaire synonyme d’une déformation adaptée de l’image.

Le choix des paramètres caractéristiques du maillage quadrangulaire (paramètres qui peuvent être assimilés aux paramètres du dictionnaire de fonctions d’analyse de la transformée adaptative) s’appuie sur une estimation du coût de description de l’image après déformation (figure 2.14(b)). Un algorithme itératif propose des champs jusqu’à obtention de celui qui minimisera une métrique liée aux résultats de compression avec la transformée 2D séparable sur l’image déformée.

Nous avons donc vu trois propositions sur la déformation préalable réversible du signal afin d’obtenir un signal fictif plus adapté à la transformée en ondelettes séparables. Nous allons maintenant présenter des propositions s’appuyant sur des champs permettant de diriger la convolution et ainsi déformer le filtrage plutôt que le signal.

Transformée grouplets Avec cette transformée, S. Mallat [153] propose d’utiliser un champ d’association afin de prendre en compte les singularités de l’image. La construction du champ d’association consiste à choisir pour chaque élément pair, un élément impair (choix restreint à un voisinage donné de l’élément pair) minimisant une mesure associée à l’association. Ces champs d’association permettront tout au long de la

transformation, de diriger les calculs selon les directions de "régularité".

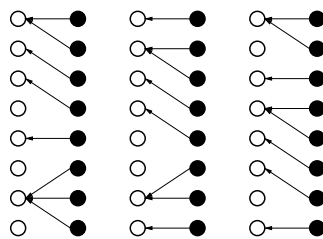


FIGURE 2.15 – Champ associant chaque élément impair à un élément pair en minimisant un critère.

Une évolution de cette transformée [183] permet également de prendre en charge des champs d'association plus complexes tels que des champs de vecteurs associés à une analyse du flux géométrique de l'image.

De nouveau, dans le cadre de la thèse de A. Maalouf [141], les Grouplettes ont été utilisées pour définir une méthode d'inpainting. La transformée en Grouplettes a permis de recomposer correctement la géométrie multiéchelle des structures de l'image et l'information à l'intérieur de la région des données manquantes est synthétisée par une propagation de cette géométrie. Nous présentons sur la figure 2.16 un exemple de suppression de nuages sur des images aériennes à partir de l'utilisation des Grouplettes [149].

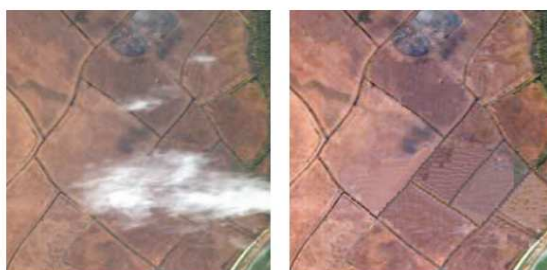


FIGURE 2.16 – Image contaminée par une ombre d'un nuage, et Inpainting de l'image par caractérisation de la géométrie par Grouplettes

Les transformées géométriques adaptatives présentées ci-dessus adaptent les données du signal afin de permettre au final une application de la transformée en ondelettes selon le formalisme des bancs de filtres. Une autre approche exploite les libertés qu'offrent les ondelettes dites de seconde génération (*i.e.* le lifting scheme [58]) afin de permettre des réécritures plus simples de transformées définies en se limitant aux outils des bancs de filtres. Comme nous le verrons, l'un des axes développé au laboratoire XLIM-SIC, que nous présenterons dans la seconde partie de ce chapitre, se positionne clairement dans cette stratégie.

Ondelettes orientées Ces transformées exploitent les orientations plus ou moins locales que l'on peut identifier dans l'image, suivant que la détection des orientations se fait pour une zone dyadique de l'image, ou pour un élément précis. La transformée proposée par Ding *et al.* [67] recherche par exemple les directions par zones dyadiques. Une fois cette direction identifiée, une interpolation est effectuée, associée au schéma de calcul Lifting (le schéma Lifting sera présenté plus tard dans ce chapitre), afin de permettre une prédiction à partir des coefficients voisins au-dessus et en-dessous de la direction pour obtenir une précision

de localisation simulant l'exploitation d'un signal associé à une fréquence d'échantillonnage quatre fois supérieure.

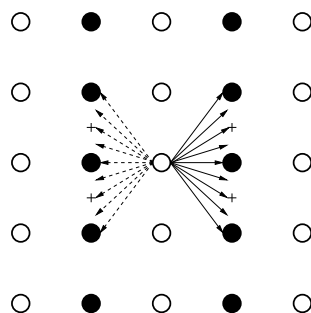


FIGURE 2.17 – Directions exploitées sur maillage séparé en pair-impair.

Une autre version de transformée orientée a été proposée par Chapellier V. et al. [50]. Elle diffère de la transformée précédente par le nombre de directions prises en charge, à savoir deux directions : horizontale ou verticale. Cette limitation s'accompagne en contre-partie d'une plus grande précision au niveau spatial : la direction n'est pas définie par zone dyadique mais par élément (figure 2.18). Cette version a aussi l'avantage de ne pas nécessiter d'interpolation. Enfin, elle s'appuie sur le champ de direction préalablement détecté pour calculer les étapes de prédiction et de mise à jour.

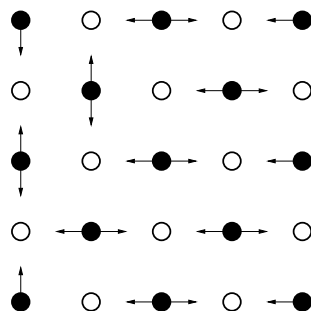


FIGURE 2.18 – Directions exploitées sur treillis quinconces : l'information de direction pour la prédiction est directement exploitée pour la mise à jour.

Ondelettes - Appariement Une alternative intéressante en terme de lifting a été proposée par Choi et Baraniuk [53] utilisant le concept des variétés. Dans ce cadre, les données à décomposer sont vues comme un nuage de point et la construction de couples de points nécessaire pour le calcul de l'échelle suivante est effectuée avec comme critère la proximité entre les éléments appariés (et non plus simplement le fait qu'ils sont voisins dans la grille temporelle).

Cet appariement est intéressant pour deux raisons :

- les données que l'on souhaite traiter ne sont plus échantillonnées régulièrement (on étudie un nuage de points de dimension n),
- la minimisation associée à l'appariement garantit la conservation de la forme globale du nuage de points.

De par l'organisation du nuage de points, il n'y a plus de distinction, pour l'échantillon traité, entre sa valeur et ses coordonnées, pas plus qu'il n'y a d'ordre établi pour le parcours des éléments.

Dans ce cas aussi, comme nous le verrons, nous trouvons des similitudes avec le travail présenté dans la seconde partie de chapitre, à savoir que la transformée ne respecte plus la grille discrète mais essaye de s'adapter à l'organisation des données.

Après ce point bibliographique permettant de situer le contexte de mon travail personnel et d'encadrement, je propose de développer plus spécifiquement les travaux menés dans ce cadre des transformées géométriques s'adaptant à l'information 2-D. Comme nous l'avons dit, nous avons déployé deux approches complémentaires, l'une à bases rigides dans le cadre des Ridgelet, l'autre adaptative à partir du concept de distance géodésique. Dans les deux cas, l'une des originalités est d'utiliser des outils de la géométrie discrète et des graphes.

2.2 Transformée géométrique, une illustration numérique : la transformée Ridgelet discrète

Après les rappels précédents, cette section va décrire le premier axe de nos contributions, à savoir un outil numérique définissant une représentation ayant pour objectif de coder efficacement les ruptures 2-D de type contours : une transformée Ridgelet numérique utilisant la géométrie discrète. A ce titre, elle s'inscrit dans la classe des analyses multirésolution géométriques. Ce travail a été mené principalement par E. Andrès et moi-même durant les années 2000-2004 en ce qui concerne le 2-D [41] et D. Helbert sous notre direction pour l'extension au 3-D [106]. Etudions pour commencer le principe de cette transformée.

2.2.1 Principe de la transformée Ridgelet

Le concept de la transformée Ridgelet est introduit et développé dans la thèse d'Emmanuel Candès [38]. Une fonction Ridge est une fonction qui est constante sur un hyperplan. La transformée Ridgelet continue d'un signal 2-D va donc correspondre à la projection de notre signal dans une base de fonctions constantes suivant des droites. La construction de cette base est similaire à celle proposée dans le cadre des ondelettes : une fonction mère constante suivant une droite est définie et les éléments de la base sont obtenus par translation, dilatation et rotation de cet élément de base.

2.2.1.1 Définition de la transformée Ridgelet

Soit le signal $s \in L^2(\mathbb{R}^2)$ la transformée Ridgelet est définie par :

$$r(a, b, \theta) = \int_{\mathbb{R}^2} \psi_{a,b,\theta}(\mathbf{x}) s(\mathbf{x}) d\mathbf{x} \quad (2.10)$$

avec $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ et $\psi_{a,b,\theta}(\mathbf{x}) \in \mathbb{R}$ la fonction Ridgelet définie à partir d'une fonction ondelette ψ :

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

avec $b \in \mathbb{R}$ le paramètre de translation, $a \in \mathbb{R}$ le paramètre de dilatation et $\theta \in [0, 2\pi[$ le paramètre de direction.

La fonction Ridgelet $\psi_{a,b,\theta} \in L^2(\mathbb{R}^2)$ est constante le long des lignes $x_1 \cos \theta + x_2 \sin \theta = b$ et est une fonction ondelette perpendiculairement à ces lignes.

Nous constatons que dans le cadre de la transformée Ridgelet, les paramètres de translation d'une transformée ondelette 2-D sont remplacés par les paramètres de la ligne (b, θ) . Ce concept d'analyse suivant une droite existe dans une autre transformée : la transformée de Radon. C'est pourquoi, nous pouvons définir la transformée Ridgelet en liant la transformée de Radon et la transformée ondelette. En effet, les coefficients Ridgelet r de s sont obtenus par une transformée en ondelette 1-D de toutes les projections de l'image contenues dans la transformée de Radon dont la direction θ est constante :

$$r(a, b, \theta) = \int_{\mathbb{R}} \psi_{a,b}(t) R_s(t, \theta) dt \quad (2.11)$$

avec R_s la transformée de Radon de s .

2.2.1.2 La transformée de Radon

La transformée de Radon a été établie en 1917 par Johann Radon [192]. Elle permet de mettre en évidence les formes linéaires d'un objet 2-D dans une image en intégrant l'intensité de l'image selon toutes les droites possibles de l'image variant d'un angle θ (figure 2.19).

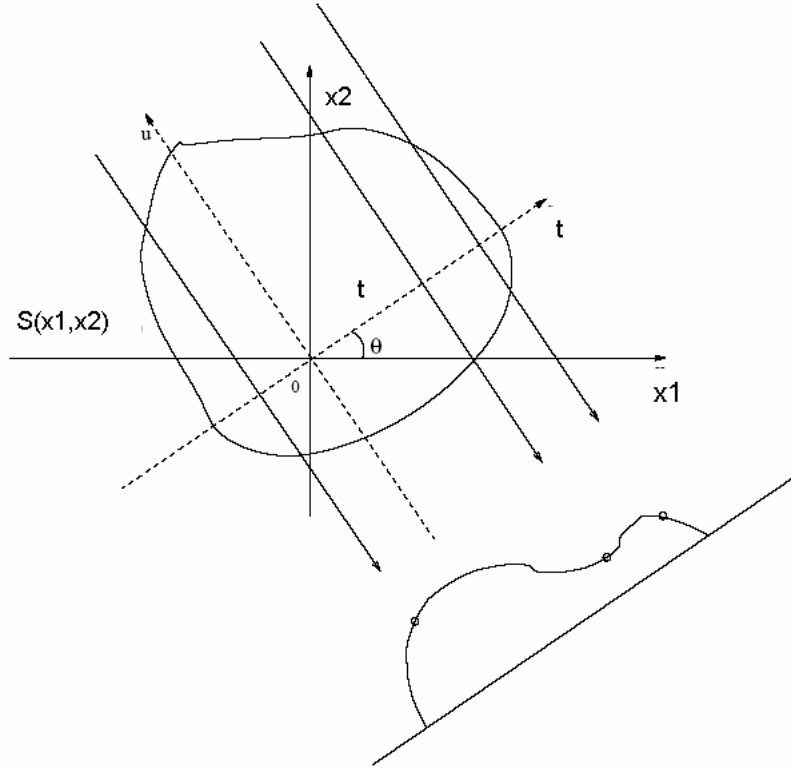


FIGURE 2.19 – Principe de l'intégration directionnelle associée à la transformée de Radon.

Transformée de Radon normale Il existe deux manières possibles de définir la transformée de Radon. Nous allons tout d'abord étudier la transformée de Radon dite normale (en référence à la manière de définir les lignes suivant lesquelles est faite l'analyse).

Soit $R_s(t, \theta)$, le coefficient de la transformée de Radon du signal $s \in L^2(\mathbb{R}^2)$, calculé selon la droite de paramètres (t, θ) , ce coefficient est défini par :

$$R_s(t, \theta) = \int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) \delta(x_1 \cos \theta + x_2 \sin \theta - t) dx_1 dx_2 \quad (2.12)$$

avec δ la masse de Dirac définie par :

$$\begin{aligned} \delta(x) &= 0 \quad \forall x \neq 0 \\ \int_{-\infty}^{\infty} \delta(x) &= 1 \end{aligned}$$

On peut aussi écrire la transformée de Radon sous la forme d'une intégrale curviligne :

$$R_s(t, \theta) = \int_{\mathbb{R}} s(t \cos \theta - b \sin \theta, t \sin \theta + b \cos \theta) db \quad (2.13)$$

D'après la définition (2.12), nous déduisons que le coefficient $R_s(t, \theta)$ de la transformée de Radon correspond à la somme suivant tous les points vérifiant $x_1 \cos \theta + x_2 \sin \theta - t = 0$ et donc éléments de la droite de direction θ . La transformée normale de Radon se caractérise suivant les deux termes (t, θ) .

Afin de comprendre l'intérêt de cette transformée dans le cadre du codage d'image, nous étudions maintenant deux transformées de signaux particuliers. Tout d'abord une singularité positionnée à une position arbitraire (x_1^*, x_2^*) définie de la façon suivante :

$$\begin{aligned} s(x_1, x_2) &= \delta(x_1 - x_1^*)\delta(x_2 - x_2^*) \implies \\ R_s(t, \theta) &= \int_{\mathbb{R}^2} \delta(x_1 - x_1^*)\delta(x_2 - x_2^*)\delta(x_1 \cos \theta + x_2 \sin \theta - t) dx_1 dx_2 \\ &= \delta(t - x_1^* \cos \theta - x_2^* \sin \theta) \end{aligned}$$

Si $\theta = \theta^*$ alors

$$R_s(t, \theta) = \int_{\mathbb{R}} \delta(\rho^* - t) db = \begin{cases} 0 & \text{si } \rho^* \neq t \\ \int \delta(0) db & \text{si } \rho^* = t \end{cases}$$

Nous constatons qu'un pic est formé à la coordonnée ρ^*, θ^* . Ceci indique qu'une rupture linéaire se traduit par une singularité dans le domaine de Radon.

Transformée Slant Stack Il existe une autre façon de définir la transformée de Radon : la transformée Slant Stack introduite par les sismographes. Cette transformée utilise la deuxième définition possible d'une droite. Soit $R_s(t, p)$, coefficient de la transformée de Radon (Slant Stack) du signal $s \in L^2(\mathbb{R}^2)$, ce coefficient est calculé par l'intégration des valeurs de s selon les lignes de pentes p et d'offset t . Le calcul est donc défini par :

$$R_s(t, p) = \int_{\mathbb{R}} s(x_1, px_1 + t) dx_1 = \int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) \delta(x_2 - px_1 - t) dx_1 dx_2 \quad (2.14)$$

Tout d'abord, étudions la transformée d'une singularité positionnée à une position arbitraire x_1^*, x_2^* définie de la façon suivante :

$$s(x_1, x_2) = \delta(x_1 - x_1^*)\delta(x_2 - x_2^*) \implies \quad (2.15)$$

$$R_s(t, p) = \int_{\mathbb{R}} \delta(x_1 - x_1^*)\delta(x_1 p + t - x_2^*) dx_1 \quad (2.16)$$

$$= \delta(t - x_2^* + px_1^*) \quad (2.17)$$

Cette équation montre qu'un point va contribuer à la création d'une ligne infinie dans le domaine de Radon.

Nous étudions ensuite la transformée d'une discontinuité linéaire paramétrée arbitrairement t^*, p^* définie de la façon suivante :

$$s(x_1, x_2) = \delta(x_2 - p^* x_1 - t^*) \implies \quad (2.18)$$

$$R_s(t, p) = \int_{\mathbb{R}} \delta(x_1(p - p^*) + t - t^*) dx_1 \quad (2.19)$$

$$= \frac{1}{|p - p^*|} \text{ si } p \neq p^* \quad (2.20)$$

Si $p = p^*$ alors

$$R_s(t, p) = \begin{cases} 0 & \text{si } t^* \neq t \\ \int \delta(0) dx_1 & \text{si } t^* = t \end{cases} \quad (2.21)$$

Nous constatons que le résultat est infini au point de coordonnées t^*, p^* . Puisque les coefficients finis sont dans ce cas négligés nous pouvons dire que la transformée de Radon d'une ligne est un point dans le domaine des paramètres, et la position du pic correspond aux paramètres de la ligne : un pic est formé aux coordonnées t^*, θ^* .

Par la suite nous utiliserons la transformée Slant Stack ou la transformée de Radon normale selon la simplicité apportée dans les expressions.

Théorème de Projection de Radon Pour calculer la transformée de Radon, nous pouvons utiliser le théorème de la projection de Radon (Fourier Slice Theorem). Celui-ci indique que la transformée de Radon peut être obtenue en appliquant une transformée de Fourier inverse 1-D le long des lignes radiales passant par l'origine dans le domaine de Fourier 2-D de l'image. Nous exprimons ce théorème à partir de la première définition de Radon car plus simple à manipuler :

Soit $\hat{R}_s(., \theta)$ la transformée de Fourier 1-D calculée sur les projections de Radon avec θ fixe et t variant :

$$\hat{R}_s(w, \theta) = \int_{\mathbb{R}} R_s(t, \theta) e^{-2j\pi w t} dt$$

Si l'on substitue dans l'équation précédente les coefficients de Radon par leurs calculs :

$$\begin{aligned} \hat{R}_s(w, \theta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) \delta(t - x_2 \sin \theta - x_1 \cos \theta) dx_1 dx_2 e^{-2j\pi w t} dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) e^{-2j\pi(x_2 \sin \theta + x_1 \cos \theta)w} dx_1 dx_2 \end{aligned}$$

Si nous rappelons la transformée de Fourier de s :

$$\hat{s}(\xi_1, \xi_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s(x_1, x_2) e^{-2j\pi(\xi_1 x_1 + \xi_2 x_2)} dx_1 dx_2 \quad (2.22)$$

nous constatons que $\hat{R}_s(w, \theta)$ correspond à $\hat{s}(., .)$ évalué le long de la ligne d'angle $\theta + \frac{\pi}{2}$ et d'offset 0. Donc

$$R_s(t, \theta) = \int_{\mathbb{R}} \hat{R}_s(w, \theta) e^{2j\pi w t} dw = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) e^{-2j\pi(x_2 \sin \theta + x_1 \cos \theta)w} dx_1 dx_2 \right] e^{2j\pi w t} dw$$

Ce théorème permet comme nous le verrons d'obtenir la formule d'inversion de la transformée de Radon à partir de la transformée de Fourier 2-D et 1-D. De plus, très récemment, nous avons abordé une généralisation de ce théorème qui fait le lien entre la transformée de Radon et la transformée de Riesz que nous étudierons dans le troisième chapitre.

2.2.2 L'aspect discret de la transformée Ridgelet

2.2.2.1 Principe

Comme nous avons vu, une méthode simple pour calculer une transformée Ridgelet est de calculer d'abord une transformée de Radon $R_s(., .)$, puis d'appliquer une transformée en ondelettes 1-D sur les projections $R_s(., \theta)$ pour chaque direction θ . La transformée Ridgelet discrète va suivre cette même stratégie.

La décomposition discrète en ondelettes est facile à implanter avec un banc de filtres et elle peut être associée avec une représentation orthogonale discrète, stable et inversible. La discrétisation de la transformée de Radon est beaucoup plus difficile. La majorité des méthodes proposées dans la littérature a été développée pour la tomographie ou pour approximer la formule continue [33, 88, 92, 164, 186, 195, 239, 30]. Mais aucune de celles-ci n'est destinée spécifiquement à être inversible et ne peut pas être utilisée pour calculer la transformée Ridgelet discrète. En effet soit l'algorithme est défini pour de l'analyse et dans ce cas seule la décomposition est caractérisée, soit nous sommes dans une problématique de projection tomographique et dans ce cas seul l'aspect reconstruction est important. La notion "d'aller-retour" entre les deux domaines n'était pas une priorité jusqu'à la transformée Ridgelet.

Quelques algorithmes ont été proposés quant au calcul de la transformée Ridgelet discrète. Certains utilisent le théorème de projection de Radon. Ils consistent à substituer les coefficients de Fourier présents à l'origine sur une grille carrée par des coefficients de Fourier de $\hat{s}$ sur une grille pseudo-polaire. Par exemple, dans ce cadre, Starck et al. [232] ont suggéré de calculer la transformée de Radon discrète suivant une stratégie d'interpolation simple consistant à prendre le plus proche voisin de la grille carrée. Ensuite, une transformée de Fourier inverse est appliquée sur chaque ligne de coefficients de Fourier pseudo-polaire passant par l'origine. Dans [44, 41], nous avons proposé d'implanter la transformée de Radon en utilisant la stratégie de Fourier en définissant des droites discrètes dans le spectre de Fourier 2-D suivant le formalisme de la géométrie discrète. L'idée est de représenter chaque direction par une droite discrète, qui forme une "bonne" approximation de la droite continue euclidienne et qui possède une symétrie centrale. Il sera préconisé l'utilisation des droites discrètes analytiques introduites par Jean-Pierre Réveillès en 1991 [194]. Ceci constitue notre premier apport que nous allons décrire.

D'autres auteurs prônent d'utiliser la transformée de Radon finie. Par exemple, Do et Vetterli [68] ont proposé une méthode d'implantation de transformée Ridgelet discrète orthogonale basée sur l'utilisation de cette transformée de Radon finie (FRAT). Cette transformée de Radon intègre suivant des lignes qui sont définies non pas géométriquement mais algébriquement. Elle est définie comme une somme de pixels sur un ensemble de lignes sur Z_p^2 , produit cartésien de deux ensembles d'entiers modulo p , p étant premier. La projection $r_k[l]$ résulte de la somme des intensités selon $L_{k,l}$, l'ensemble des points couverts par une ligne sur Z_p^2 , défini par :

$$L_{k,l} = \{(n, m) \mid m = kn + l \pmod{p}\}, n \in Z_p \quad 0 \leq k < p.$$

La transformée de Radon finie n'est pas redondante et elle est inversible. Les résultats de cette transformée sur le débruitage d'images illustrées dans [68] montrent qu'elle est plus efficace qu'une transformée ondelette pour les bords rectilignes des objets présents dans une image. Toutefois, en utilisant l'opération modulo dans la définition des lignes, l'image est considérée comme une période d'une image périodique et un phénomène de "réplication" des lignes est présent (nous retrouverons par la suite cette notion de "wrap-around"). De plus un point dans l'espace transformée ne correspond pas à une ligne dans l'espace d'origine donc cette transformée n'est pas géométriquement fidèle à la notion de Ridgelet.

Enfin, Averbuch et al. [12] ont proposé une transformée de Fourier pseudo-polaire basée sur une transformée chirp-Z particulière (base construite à partir de fonctions modulées en fréquence). La transformée de Radon associée, nommée Fast Slant Stack, est algébriquement exacte et géométriquement fidèle à la notion de Ridgelet. Il nous semble important de nous attarder sur cette dernière stratégie qui constitue la référence dans la littérature. Ceci est nécessaire pour mettre en perspective et comparer le travail mené à XLIM-SIC dans ce cadre.

Pour bien comprendre ces travaux, il est nécessaire de rapporter brièvement les travaux menés dans le cadre de la transformée de Radon discrète car bien que développés dans un cadre différent, ils ont alimenté notre réflexion dans le cadre Ridgelet. Cette étape de synthèse sur la transformée de Radon discrète a constitué une part importante du travail mené dans le cadre de la transformée Ridgelet (ce lien entre la transformée Ridgelet et les travaux sur Radon discret est très peu fait dans la littérature) et a trouvé, comme nous le verrons une implication inattendue dans nos derniers travaux sur les transformées monogéniques.

2.2.2.2 La transformée de Radon discrète

En fonction du but recherché, des artéfacts acceptés ou de la simplicité voulue de l'algorithme, il existe différentes manières de définir la transformée de Radon discrète. La solution la plus simple est d'échantillonner les deux paramètres de la transformée Slant Stack :

$$\begin{aligned} t_d &= k_1 \Delta_t \text{ avec } k_1 \in \mathbb{Z} \\ p_d &= k_2 \Delta_p \text{ avec } k_2 \in \mathbb{Z} \end{aligned}$$

avec Δ_t et Δ_p les pas d'échantillonnage.

Puisque nous calculons une transformée discrète, l'image s étudiée est bien sûr elle aussi échantillonnée :

$$s[n_1, n_2] = s(n_1 \Delta_{x_1}, n_2 \Delta_{x_2}) \text{ avec } n_1, n_2 \in \mathbb{Z}$$

Une façon simple de définir la transformée de Radon discrète est de substituer dans l'équation (2.14) l'intégrale par une somme :

$$R_s[t_d, p_d] = \int_{\mathbb{R}} s(x_1, p_d x_1 + t_d) dx_1 \approx \Delta_{x_1} \sum_{n_1=0}^{N-1} s[n_1, p_d n_1 + t_d] \quad (2.23)$$

Nous constatons que la transformée de Radon est alors une image discrète. Par simplicité nous ne gardons pas par la suite les indices d : $R_s[t_d, p_d] \leftarrow R_s[t, p]$. De plus nous considérons que les pas d'échantillonnage spatiaux sont $\Delta_{x_1} = 1$ et $\Delta_{x_2} = 1$ et enfin que $n_1, n_2 \in [0, N - 1]$.

Bien que nous ayons discrétisé le calcul il reste un problème dans l'équation (2.23). Le calcul nécessite des échantillons non définis dans l'image numérique (sur une grille cartésienne). En effet, la valeur $p n_1 + t$ doit coïncider avec les différents échantillons $n_2 \Delta_{x_2}$, ce qui est rarement le cas. Afin de résoudre ce problème, il faut mettre en place une procédure d'interpolation. Il existe différentes stratégies :

- Interpolation suivant le plus proche voisin dans la direction x_2 . Dans ce cas la transformée de Radon se définit par :

$$R_s[t, p] = \sum_{n_1=0}^{N-1} s[n_1, \lceil p n_1 + t \rceil]$$

avec $\lceil x \rceil$ l'entier le plus proche de x . Notons que le point $s[n_1, \lceil p n_1 + t \rceil]$ peut être à l'extérieur du domaine de définition de l'image, et dans ce cas il peut être considéré égal à 0. Cette stratégie est utilisée par Starck et al pour implanter la transformée Ridgelet [232].

- Interpolation linéaire dans la direction x_2 .

$$R_s[t, p] = \sum_{n_1=0}^{N-1} (1 - w) s[n_1, \lfloor p n_1 + t \rfloor] + w s[n_1, \lfloor p n_1 + t \rfloor + 1]$$

avec $\lfloor \cdot \rfloor$ la partie entière et $w = (pn_1 + t) - \lfloor pn_1 + t \rfloor$.

On peut améliorer le schéma d'interpolation en augmentant l'ordre de l'interpolation, mais la contrepartie est une augmentation de la complexité de l'algorithme.

- Interpolation par un Sinus Cardinal. Dans ce cas, en utilisant le principe du théorème d'échantillonnage, on peut considérer que la fonction continue $\{s(x_1, x_2)\}_{x_1, x_2 \in \mathbb{R}}$ peut être définie à partir de ces échantillons discrets $\{s[n_1, n_2]\}_{n_1, n_2 \in \mathbb{Z}}$ par une interpolation suivant un sinus cardinal :

$$s(x_1, x_2) \approx \sum_{n_1} \sum_{n_2} s[n_1, n_2] \frac{\sin(\pi(x_1 - n_1))}{\pi(x_1 - n_1)} \frac{\sin(\pi(x_2 - n_2))}{\pi(x_2 - n_2)}$$

On peut utiliser alors cette relation pour déterminer les échantillons manquants dans le calcul de Radon. Nous retrouvons cette idée d'interpolation dans le cadre de la méthode Fast Slant Stack [12].

Nous constatons que la définition d'une opération inverse à ces calculs discrets n'est pas immédiate du fait de la procédure d'interpolation.

2.2.2.3 Inversion de la transformée Radon discrète

Calcul direct Dans la majorité des méthodes de reconstruction d'une image à partir des projections de Radon, c'est le théorème de projection qui est utilisé. Nous rappelons ici les équations définissant ce théorème :

$$\begin{aligned} \hat{s}(v \cos \theta, v \sin \theta) &= \int_{-\infty}^{+\infty} R_s(t, \theta) e^{-2j\pi tv} dt \\ s(x_1, x_2) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \hat{S}(\xi_1, \xi_2) e^{-2j\pi(\xi_1 x_1 + \xi_2 x_2)} d\xi_1 d\xi_2 \end{aligned} \quad (2.24)$$

On peut donc utiliser directement cette expression dans le domaine discret. La procédure est la suivante :

- Nous calculons la transformée de Fourier 1-D pour chaque paramètre angulaire θ ,
- nous considérons ensuite que nous avons des échantillons définis sur une grille polaire, donc nous devons passer à une grille cartésienne,
- nous obtenons enfin l'image originale en calculant la transformée de Fourier inverse 2-D.

Bien évidemment l'étape problématique dans la transformée discrète est le passage de la grille polaire à la grille cartésienne. Nous devons définir une opération adjointe à celle proposée dans le cadre du calcul de la transformée de Radon. Toutefois comme dans le domaine spatial ce passage nécessite des échantillons non définis dans le spectre numérique. Afin de résoudre ce problème, les solutions proposées utilisent de nouveau une procédure d'interpolation dans le domaine fréquentiel (nous ne détaillons pas à nouveau les procédures d'interpolation).

Il existe une autre manière de définir un algorithme de reconstruction appelée la rétroprojection filtrée. Cette inversion est déduite de l'équation (2.24) en introduisant les coordonnées polaires :

$$\begin{aligned} s(x_1, x_2) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \hat{s}(\xi_1, \xi_2) e^{2j\pi(\xi_1 x_1 + \xi_2 x_2)} d\xi_1 d\xi_2 \\ &= \int_0^{2\pi} \int_0^{+\infty} \hat{s}(v \cos \theta, v \sin \theta) e^{2j\pi v(\cos \theta x_1 + \sin \theta x_2)} v dv d\theta \\ &= \int_0^\pi \check{R}_s(\cos \theta x_1 + \sin \theta x_2, \theta) d\theta \end{aligned}$$

avec

$$\check{R}_s(t, \theta) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} R_s(\tilde{t}, \theta) e^{-2j\pi\tilde{t}v} d\tilde{t} \right] e^{2j\pi vt} |v| dv$$

Nous constatons que la reconstruction peut se décomposer en deux parties :

- Une partie rétroprojection qui consiste à passer du domaine (t, θ) au domaine (x_1, x_2) (opération adjointe) :

$$s(x_1, x_2) = \int_0^\pi \check{R}_s(\cos \theta x_1 + \sin \theta x_2, \theta) d\theta = \int_{-\infty}^{+\infty} \int_0^\pi \check{R}_s(\theta, t) \delta(t - \cos \theta x_1 - \sin \theta x_2) d\theta dt$$

- Une partie filtrage qui correspond à la définition de $\check{R}_s$ avec un filtrage passe-haut avec le filtre $|v|$ dans le domaine fréquentiel.

L'implantation numérique de cette méthode doit résoudre les deux problèmes suivants :

- le choix du filtre passe-haut car un simple filtre $|v|$ amplifie le bruit.
- le calcul de la rétroprojection pose les mêmes problèmes que le calcul de la projection et sont résolus par interpolation.

Cette dernière stratégie n'a pas été utilisée dans le cadre de la transformée Ridgelet car elle est entièrement dédiée à l'inversion et ne permet pas de faire le lien entre le domaine spatial et le domaine transformée.

Utilisation de l'algèbre linéaire Différents algorithmes de reconstruction utilisent l'algèbre linéaire. Si l'on note $\mathcal{R}$ l'opérateur associé à la transformée de Radon, on peut exprimer la transformée suivant une représentation matricielle :

$$\begin{array}{ccccc} R_s(t, \theta) & = & \mathcal{R} & s(x_1, x_2) \\ \downarrow & & \downarrow & & \downarrow \\ b & = & A & & x \end{array}$$

La matrice A n'est pas carrée et bien évidemment elle ne peut s'inverser simplement. La formulation sous forme d'un système d'équations donne accès aux nombreuses méthodes itératives de résolution de systèmes. En général, dans les méthodes itératives, on n'utilise pas la matrice inverse, mais la matrice transposée qui correspond en terme d'opérateur à la transformée de Radon adjointe qui passe du domaine (t, θ) au domaine spatial.

$$\begin{array}{ccccc} \tilde{s}(x_1, x_2) & = & \tilde{\mathcal{R}} & R_s(t, \theta) \\ \downarrow & & \downarrow & & \downarrow \\ \tilde{b} & = & A^T & & x \end{array}$$

Dans ce cas, on cherche en général à résoudre le système

$$A^T b = (A^T A) x$$

Il est plus facile de définir la transformée adjointe plutôt que la transformée inverse. La méthode Fast Slant Stack utilise ce concept.

2.2.2.4 Transformée de Radon dans le cadre de la Fast Slant Stack

Comme indiqué précédemment, nous proposons une description détaillée de la Fast Slant Stack qui constitue l'une des approches dans le cadre des Ridgelet les plus abouties.

Tous les développements présentés correspondent notamment aux articles [215, 12]. Ces travaux ont eu par la suite différentes évolutions, nous faisons ici le choix de présenter les fondements. Shkolnisky, Averbuch et Donoho ont établi 5 principes à respecter par la transformée de Radon discrète ([215, 12]) : validité algébrique ; fidélité géométrique ; rapidité ; inversion ; parallèles avec la théorie continue. La méthode proposée par Shkolnisky et al. utilise le théorème de projection de Fourier. Le choix est motivé entre autre par l'existence des algorithmes rapides de calcul de la transformée de Fourier directe et inverse.

Tout d'abord précisons que du fait de la périodisation de l'image lors du calcul de transformée de Fourier, il apparaît alors après calcul des différentes projections un effet dit de "wraparound" qui remet en cause la fidélité géométrique de la transformée. Pour conserver une fidélité géométrique à la transformée, on réalise alors une immersion de l'image initiale dans une image de taille $(2n + 1) \times n$ (avec n taille de l'image d'origine) construite par une concaténation symétrique de zéros. La figure 2.20 illustre ce problème. Nous constatons sur la figure 2.20b, qui correspond à un calcul sans concaténation de 0, qu'un point est associé dans le domaine de Radon à une ligne mais qui se réplique.

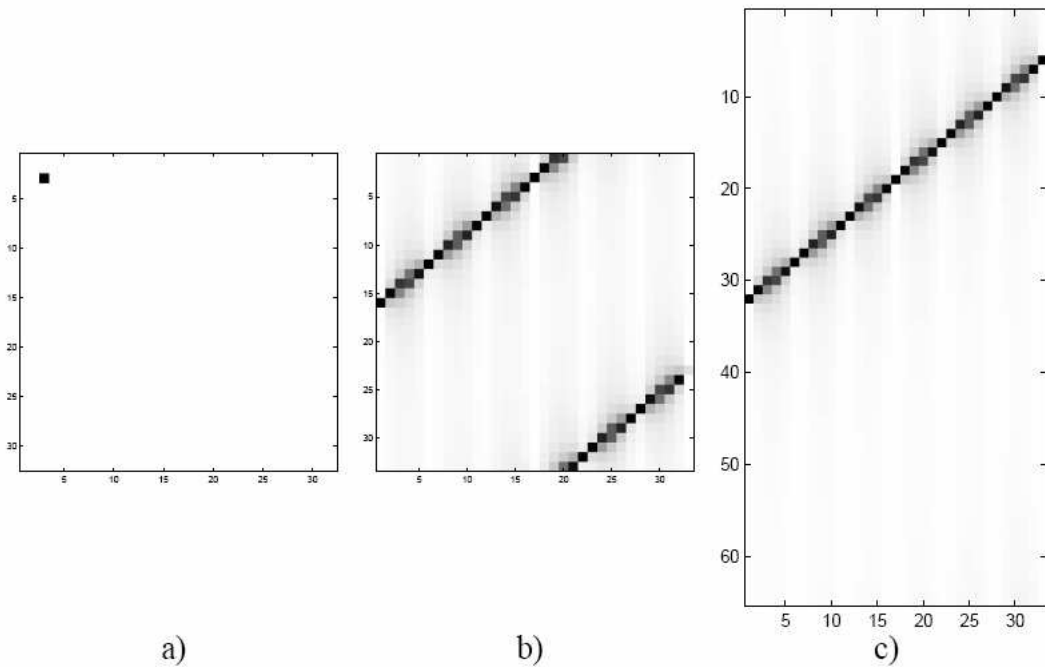


FIGURE 2.20 – Résultat expérimental de l'effet "wraparound" : a) image originale où $I[3, 3] = 1$ et 0 ailleurs ; b) transformée Radon sans concaténation de zéros (effet "wraparound") ; c) transformée de Radon avec concaténation de zéros.

Ensuite, afin de pouvoir calculer la transformée Radon, il suffit de connaître les échantillons de la transformée Fourier 2-D sur une grille non cartésienne. Cette grille s'appelle grille pseudo polaire et ses points sont décrits par les relations suivantes :

$$\begin{aligned} \xi_{l,k}^1 &= (-k \tan \theta, k), & \theta &\in \Theta_2 \\ \xi_{l,k}^2 &= (k, -k \cot \theta), & \theta &\in \Theta_1 \end{aligned}, \quad -n \leq k \leq n \quad (2.25)$$

ou, en utilisant les relations suivantes qui décrivent la discrétisation de l'angle θ :

$$\begin{aligned}\Theta_2 &= \{ \theta \mid \theta = \arctan(2l/n); \quad l \in \mathbb{Z}, -n/2 \leq l < n/2 \} \\ \Theta_1 &= \{ \theta \mid \theta = \pi/4 + \arctan(2l/n); \quad l \in \mathbb{Z}, -n/2 \leq l < n/2 \}\end{aligned}\quad (2.26)$$

nous obtenons

$$\begin{aligned}\xi_{l,k}^1 &= (-k\frac{2l}{n}, k) \quad \theta \in \Theta_2, \\ \xi_{l,k}^2 &= (k, -k\frac{2l}{n}) \quad \theta \in \Theta_1,\end{aligned} \quad , \quad -n \leq k \leq n, \quad -\frac{n}{2} \leq l \leq \frac{n}{2} \quad (2.27)$$

Nous constatons dans les expressions précédentes que nous considérons directement le cas où des 0 sont ajoutés à l'image afin d'éviter les phénomènes de wraparound, ce qui explique $-n \leq k \leq n$.

Soit $\mathcal{L}_1$ l'ensemble de lignes dans le domaine Fourier duales de celles fondamentalement horizontales dans le domaine spatial et soit $\mathcal{L}_2$ l'ensemble de lignes duales aux lignes fondamentalement verticales. L'interprétation des relations (2.27) est la suivante :

- $\xi_{l,k}^1$ est l'intersection de l -ième ligne de $\mathcal{L}_1$ avec la k -ième ligne horizontale de la grille cartésienne ;
- $\xi_{l,k}^2$ est l'intersection de l -ième ligne de $\mathcal{L}_2$ avec la k -ième ligne verticale.

Nous illustrons cette grille sur la figure 2.21.

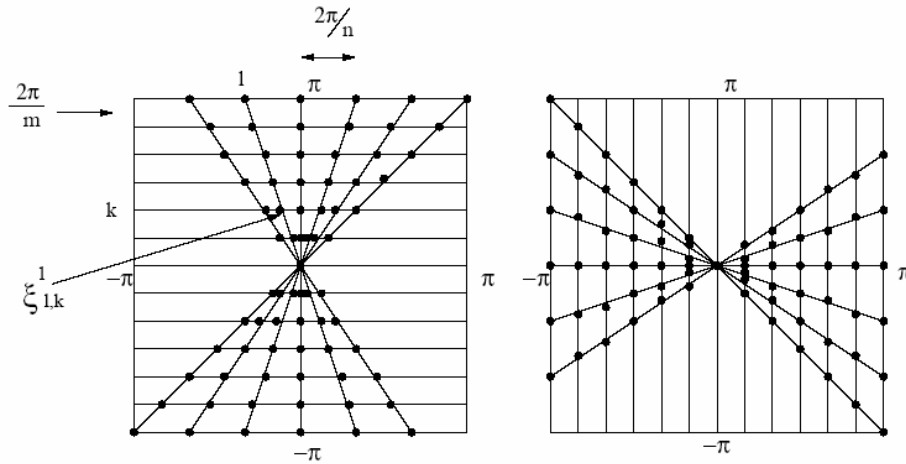


FIGURE 2.21 – Les échantillons pseudo polaires $\xi_{l,k}^1$ et $\xi_{l,k}^2$.

A la différence de la grille polaire, illustrée dans la figure 2.23, dans la grille pseudo-polaire la distance Δd entre deux points, P et Q situés sur le même rayon, n'est pas constante quand les points décrivent un mouvement autour de l'origine : $d(P, Q) \neq d(P', Q')$ comme nous le constatons sur la figure 2.22. La déviation maximale est $\Delta d_{\max}/\Delta d_{\min} = \sqrt{2}$. De plus, deux angles formés par trois rayons consécutifs ne sont pas égaux ($\angle POR \neq \angle POP'$). La déviation maximale d'angle est $\Delta \theta_{\max}/\Delta \theta_{\min} \leq 2$.

Lors de cette conversion de grille, nous souhaitons obtenir les valeurs $\hat{I}(\xi_{l,k}^1)$ et $\hat{I}(\xi_{l,k}^2)$ à partir des échantillons $I[n_1, n_2]$ issus de la TF. Ceci n'est pas possible en une seule étape (une transformée de Fourier "directement pseudo-polaire") du fait de la forme de la grille pseudo-polaire et du sur-échantillonnage du domaine Fourier dans cette grille. Donc après transformée de Fourier bidimensionnelle calculée sur l'image, le passage de la grille cartésienne vers la grille pseudo-polaire se fait par une interpolation suivant un polynôme trigonométrique. Les $n + 1$ échantillons d'une ligne $\{\hat{I}(\xi_{l,k}^1)\}_{-n/2 \leq l \leq n/2}$, pour k fixé, sont obtenus en utilisant les n échantillons de la même ligne dans la grille cartésienne, comme cela est schématisé sur la figure 2.24.

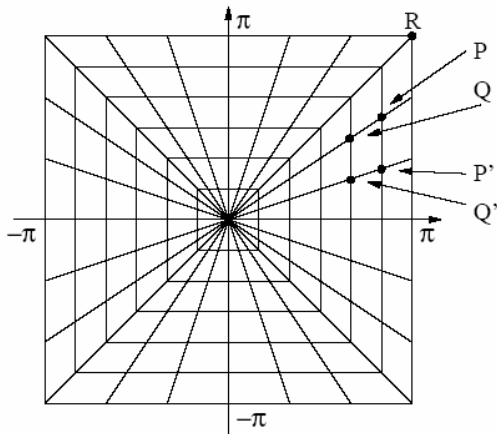


FIGURE 2.22 – Grille pseudo polaire pour $n = 6$.

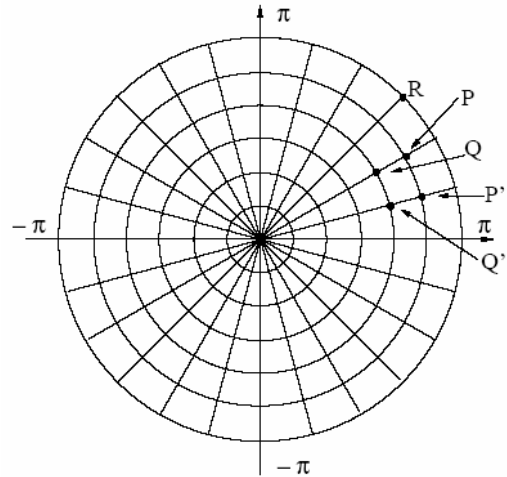


FIGURE 2.23 – Grille polaire pour $n = 6$.

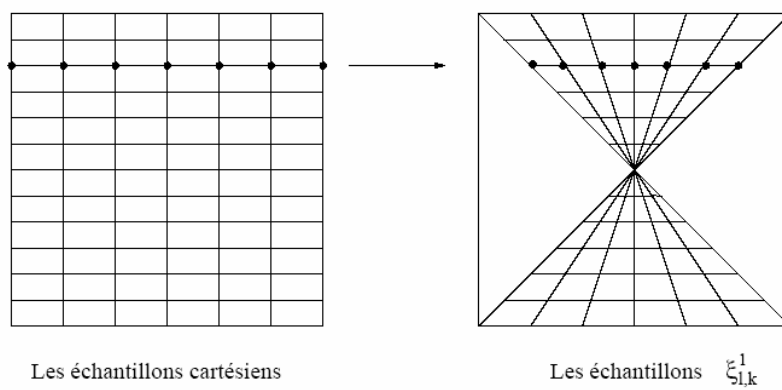


FIGURE 2.24 – Conversion de la grille cartésienne vers la grille pseudo-polaire (Les échantillons pseudo polaires $\xi_{l,k}$).

Etudions précisément le traitement d'une ligne k_0 . Soit

$$T_{k_0}^1(l) = \sum_{v=-\frac{n}{2}}^{\frac{n}{2}} c_{k_0}^1(v) e^{-\frac{2\pi i}{n} lv}, \quad -\frac{n}{2} \leq l \leq \frac{n}{2}$$

le polynôme trigonométrique qui, dans le cas des lignes fondamentalement horizontales ($s = 1$), correspond à la restriction de $\hat{I}$ à la ligne k_0 dans la grille cartésienne. Ce polynôme trigonométrique doit alors vérifier que $\hat{I}(\xi_1, k_0)_{-\frac{n}{2} \leq \xi_1 \leq \frac{n}{2}-1}$ (coefficients éléments de la grille cartésienne) soient ses valeurs pour $-\frac{n}{2} \leq l < \frac{n}{2}$. On déduit que les coefficients du polynôme sont issus de la transformée de Fourier inverse du spectre de l'image selon les lignes. Ensuite l'évaluation du polynôme aux positions αl , avec $\alpha = 2k_0/n$ nous donne les échantillons fréquentiels sur la ligne k_0 de la grille pseudo-polaire. Cette évaluation correspond au calcul d'une transformée de Fourier fractionnaire. Généralement, pour un vecteur X de $n + 1$ composantes et $\alpha \in \mathbf{R}$, la transformée de Fourier fractionnaire est définie comme suit :

$$\mathcal{F}_{n+1}^\alpha(X)(k) = \sum_{t=-n/2}^{n/2} X(t) e^{\frac{-2\pi j \alpha k t}{n+1}}; \quad k \in \mathbf{Z}; \quad -n/2 \leq k \leq n/2 \quad (2.28)$$

En termes d'opérateurs, la conversion entre les grilles se fait avec l'opérateur :

$$G_{n,k} = \mathcal{F}_M^\alpha \circ \mathcal{F}_n^{-1} \quad (2.29)$$

où $\mathcal{F}_n^{-1}$ est la transformée de Fourier discrète inverse unidimensionnelle calculée sur un signal de longueur n et $\mathcal{F}_M^\alpha$ est la transformée de Fourier fractionnaire de facteur α . Cet opérateur $\mathcal{F}_M^\alpha$ accepte une séquence de longueur n , concatène symétriquement des 0 afin d'atteindre la longueur $M = 2n + 1$, applique dessus une transformée de Fourier fractionnaire et retourne la partie centrale de longueur $n + 1$. A l'aide de l'opérateur $G_{n,k}$, on peut écrire l'opérateur S de ré-échantillonnage des lignes à partir de la grille cartésienne vers la grille pseudopolaire. Dans ce cas, la transformée de Fourier pseudopolaire s'écrit :

$$P = S \circ \mathcal{F}_2 \quad (2.30)$$

L'opérateur P est plus précisément la réunion de deux opérateurs, $P = P_1 \cup P_2$, P_1 associé aux échantillons $\theta \in \Theta_2$ et P_2 associé aux échantillons $\theta \in \Theta_1$.

Par conséquent, on peut encore écrire :

$$R_I(\theta, k) = \begin{cases} \mathcal{F}_1^{-1} \circ P_1 \circ E(I) & \text{pour } \theta \in \Theta_2 \\ \mathcal{F}_1^{-1} \circ P_2 \circ E(I) & \text{pour } \theta \in \Theta_1 \end{cases} \quad (2.31)$$

où E est l'opérateur qui étend l'image par concaténation de zéros.

Pour résumer, la transformée *fast slant stack* se décline suivant :

1. l'image I de dimensions $n \times n$ est étendue à l'image EI aux dimensions $(2n+1) \times n$ par concaténation des zéros ;
2. la transformée de Fourier bidimensionnelle de EI est calculée, la taille du tableau reste inchangée ;
3. à chaque ligne de $\mathcal{F}_2(EI)$ on applique la transformée de Fourier inverse unidimensionnelle, la taille du tableau reste inchangée ;

4. au résultat $\mathcal{F}_1^{-1} \circ \mathcal{F}_2(EI)$ on applique la transformée de Fourier fractionnelle, $\mathcal{F}_{2n+1}^\alpha$ avec $\alpha = 2k/n$, sur chaque ligne et on obtient une image de taille $(2n+1) \times (n+1)$;
5. à chaque colonne du résultat du pas précédent on applique une transformée de Fourier unidimensionnelle inverse et on obtient une image $(2n+1) \times (n+1)$.

2.2.2.5 La reconstruction dans le cadre de la Fast Slant Stack

La reconstruction des images consiste à définir la transformée inverse de *fast slant stack*. Compte tenu de l'enchaînement résumé dans le paragraphe précédent, ceci signifie trouver la transformée décrite par :

$$FSS^{-1}(R_I) = I = U \circ \mathcal{F}_2^{-1} \circ S^{-1} \circ \mathcal{F}_1(R_I) \quad (2.32)$$

avec U l'opération qui enlève les 0 ($U \circ E(I) = I$). Parmi les opérateurs présents dans la relation (2.32), il y a un opérateur problématique : il s'agit de S^{-1} , c'est-à-dire le passage de la grille pseudo-polaire à la grille cartésienne. Le problème de l'inversion de S consiste en la recherche de l'inverse de la transformée fractionnaire $\mathcal{F}_{2n+1}^\alpha$. Pour les autres opérateurs on connaît l'existence des algorithmes rapides de calcul.

Une approche naïve consiste en l'interpolation des échantillons du polynôme $T_k(\alpha l)$ afin de trouver les échantillons $T_k(l)$. Cette approche est mal conditionnée. En effet, pour des petites valeurs de α , de minuscules perturbations des échantillons $T(\alpha l)$ proches de l'origine peuvent causer des grandes variations loin de l'origine. Ceci est compréhensible si l'on tient compte du taux de ré-échantillonnage induit par la projection de Radon. En effet, ce taux à une distance $k \approx n$ est d'un échantillon cartésien pour un échantillon pseudo-polaire tandis qu'aux distances $k \approx 0$ le taux de ré-échantillonnage est d'un échantillon cartésien pour n échantillons pseudo-polaires.

On peut tout d'abord définir une opération adjointe (ce qui nous permettra d'appliquer une méthode de résolution de systèmes comme la méthode du gradient). En sachant que $\text{adj}\mathcal{F}_n^\alpha = \mathcal{F}_n^{-\alpha}$, on en déduit que :

$$\text{adj}S = \text{adj}(\mathcal{F}_m^\alpha \circ \mathcal{F}_n^{-1}) = \mathcal{F}_n \circ \tilde{\mathcal{F}}_m^{-\alpha}$$

Cet opérateur $\tilde{\mathcal{F}}_M^{-\alpha}$ accepte une séquence de longueur $n+1$, concatène symétriquement des 0 afin d'atteindre la longueur $M = 2n+1$, applique dessus une transformée de Fourier fractionnaire de paramètre $-\alpha$ et retourne la partie centrale de longueur n .

Notons que dans ce cas, l'opération adjointe n'est pas l'inverse du fait de la transformée de Fourier Fractionnaire.

Dans ce cas on peut définir la transformée de Radon adjointe :

$$\text{adj}(R_I) = U \circ \mathcal{F}_2^{-1} \circ \mathcal{F}_n \circ \tilde{\mathcal{F}}^{-\alpha} \circ \mathcal{F}_1(R_I)$$

Averbuch et al. [12] proposent d'utiliser la méthode du gradient conjugué qui est une technique itérative permettant de résoudre les systèmes linéaires définis positifs et symétriques. On utilise cette méthode pour le calcul par approximation des échantillons cartésiens à partir des échantillons pseudo-polaires.

Supposons l'existence des vecteurs Radon $R_\theta I$ de l'image I . La reconstruction de l'image proposée par Stanford fonctionne ainsi :

1. Calcul de la transformée de Fourier unidimensionnelle de $F = \mathcal{F}_1 R_\theta I$.
2. Calcul par l'algorithme du gradient conjugué des échantillons cartésiens à partir des échantillons pseudo-polaires F ;

3. Calcul de la transformée de Fourier bidimensionnelle inverse à partir des échantillons cartésiens précédemment obtenus.

Averbuch et al. ont montré l'efficacité de cet algorithme de reconstruction.

Nous avons rappelé ce qui constitue l'un des algorithmes de référence dans le domaine de la transformée Ridgelet. Nous proposons de décrire maintenant notre travail sur la transformée Ridgelet qui, comme nous le verrons, est complémentaire dans la définition du problème.

2.2.3 La stratégie de Poitiers : la transformée Ridgelet analytique ou DART

Dans [44, 41], j'ai proposé en collaboration avec E. Andres de définir la transformée de Radon en utilisant la stratégie de Fourier. Cette stratégie diffère de celle présentée précédemment. En effet, on se place bien dans le domaine de Fourier mais le problème n'est pas formulé en terme d'interpolation des coefficients de Fourier pour définir des échantillons sur une grille polaire mais la question posée est : comment définir des droites sur une grille discrète ?

Cette solution nécessite une adaptation du théorème de projection (une approximation) puisque les droites ne sont plus définies de manière continue. Mais les résultats obtenus en terme de projection sont sensiblement identiques à ceux de l'algorithme Fast Slant Stack et la méthode est associée à un algorithme de transformée inverse non itératif rapide qui permet d'obtenir une reconstruction exacte.

Le calcul proposé de la transformée de Radon est l'extraction des coefficients de Fourier le long des lignes L_θ passant par l'origine. Ces lignes L_θ sont définies dans le domaine de Fourier avec l'aide de la théorie de la géométrie analytique discrète [7, 194]. Nous précisons cela dans la section suivante. La transformée Ridgelet proposée est appelée : Discrete Analytical Ridgelet Transform, ou DART.

2.2.3.1 Définitions des droites discrètes 2-D

Afin de définir des notions de droites discrètes, nous devons préciser les notions de topologie à partir de notion de voisinage élémentaire. Il faut définir d'abord les notions de voisinage élémentaire, puis de connexité. Pour cela, on considère l'image discrète comme un graphe. Les nœuds du graphe sont les points discrets (les pixels de l'image). Les arcs sont définis par les relations de voisinage entre les points. Par exemple, sur une trame carrée, on peut considérer que deux points sont voisins si une et une seule de leurs coordonnées diffère d'une unité. Chaque point a ainsi 4 voisins, et on parle de 4-connexité. Si l'on ajoute dans le graphe les arcs reliant les points aussi en diagonale (dont les deux coordonnées diffèrent d'une unité), chaque point a alors 8 voisins, et on parle de 8-connexité.

Une fois que sont définis les voisinages élémentaires, les notions de connexité s'en déduisent grâce à l'interprétation en termes de théorie des graphes. Un chemin est ainsi défini comme une suite de sommets du graphe telle que deux points consécutifs de la suite soient joints par un arc du graphe, où les arcs sont définis en fonction de la connexité choisie. On définit

- un 4-chemin comme une suite de points $\{P_i(x_i, y_i)\}_{1 \leq i \leq n}$ tels que $|x_i - x_{i+1}| + |y_i - y_{i+1}| \leq 1$ pour $\forall i, 1 \leq i < n$
- un 8-chemin comme une suite de points $\{P_i(x_i, y_i)\}_{1 \leq i \leq n}$ tels que $\max(|x_i - x_{i+1}|, |y_i - y_{i+1}|) \leq 1$ pour $\forall i, 1 \leq i < n$.

Avec ces notions, nous pourrions caractériser la topologie des droites discrètes construites.

L'idée proposée dans la transformée de Radon analytique est de représenter chaque direction par une droite discrète. Pour cela, nous désirons une droite discrète qui forme une "bonne" approximation de la droite continue euclidienne. De plus, nous voulons une symétrie centrale puisque ces droites sont définies dans le domaine de Fourier. Ceci exclut les droites classiques de Bresenham [31]. Nous utilisons alors les droites discrètes analytiques introduites par Jean-Pierre Réveillès en 1991 [194] suivant :

Définition 1 Soit $L(q,p,c,w)$ la droite discrète 2-D de Réveillès définie par

$$0 \leq qx_1 + px_2 + c < w$$

avec

- $(q, p) \in \mathbb{Z}$ le coefficient de la droite
- $c \in \mathbb{Z}$ le coefficient de translation
- $w \in \mathbb{Z}^{*+}$ l'épaisseur arithmétique

On peut trouver une définition équivalente de type

$$\left\lfloor \frac{qx_1 + px_2 + c}{w} \right\rfloor = 0$$

On constate que l'on ne définit pas une unique ligne mais une famille de lignes dépendant d'un paramètre w qui contrôle l'épaisseur et la connectivité des lignes.

Comme nous l'avons déjà fait remarquer, nous avons besoin pour la transformée de Radon d'une symétrie centrale. En effet, le coefficient de Fourier central correspond à la composante continue et les coefficients symétriques par rapport au centre sont conjugués. Ainsi, cet ensemble doit être composé d'un nombre impair de coefficients de Fourier. La symétrie est facilement obtenue en définissant une variante des droites analytiques discrètes. Soit les deux points $O(0, 0)$ et $P(p, q)$, la droite discrète reliant ces deux points est définie par :

$$L_{(p,q)}^\omega = \left\{ (x_1, x_2) \in \mathbb{Z}^2 \mid |qx_1 - px_2| \leq \frac{\omega}{2} \right\} \quad (2.33)$$

$[p, q] \in \mathbb{Z}^2$ correspondent à la direction θ de la droite (direction de la projection de Radon) et ω , fonction de p et q , est l'épaisseur arithmétique. Nous pouvons facilement voir que cette ligne $L_{[p,q]}^\omega$ présente une symétrie centrale.

L'épaisseur arithmétique contrôle la connexité des droites discrètes : par exemple si $\omega_L < \max(|p|, |q|)$, alors la droite discrète n'est pas connectée.

Trois différents types de droites discrètes fermées ont été étudiés dans le cadre de la transformée de Radon :

- les droites discrètes naïves fermées : $\omega = \max(|p|, |q|)$. Ces droites (figure 2.25(a)) sont les plus fines droites discrètes fermées connectées. Elles sont 8-connexes,
- les droites discrètes Pythagoriciennes fermées : $\omega = \sqrt{p^2 + q^2}$. Ces lignes (figure 2.25(b)) sont 4-connexes et offrent une redondance moyenne, entre celle des droites naïves et supercouvertures.
- les droites supercouvertures : $\omega = |p| + |q|$ [8, 7]. Ces droites couvrent la droite Euclidienne qu'elles approximent (figure 2.25(c)) et elles sont les plus fines droites 4-connexes. Ce sont les droites discrètes les plus épaisses considérées dans les applications de la transformée de Radon.

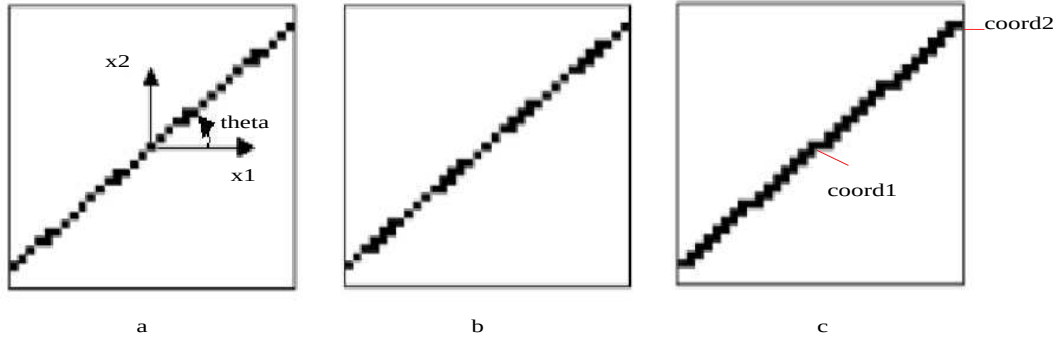


FIGURE 2.25 – Les droites analytiques discrètes 2-D : (a) naïve fermée ($\omega = \max(|p|, |q|)$), (b) Pythagoricienne fermée ($\omega = \sqrt{p^2 + q^2}$) et (c) supercouverture ($\omega = |p| + |q|$)

2.2.3.2 Stratégie de calcul de la transformée de Radon discrète

Nous utilisons donc le domaine de Fourier pour calculer la transformée de Radon : les coefficients de Fourier de l'image sont extraits le long des lignes discrètes $L_{[p,q]}^\omega$ (les points extraits sont ordonnés suivant l'ordre naturel) :

$$P_{[p,q]}^\omega = \bigcup_{k \in \mathbb{Z}^+} \hat{s}(\xi_1^k, \xi_2^k) \text{ avec } |q\xi_1^k - p\xi_2^k| \leq \frac{\omega}{2}$$

et on calcule la transformée de Fourier 1-D inverse pour chaque $P_{[p,q]}^\omega$ à chaque valeur de la direction $[p, q]$.

L'échantillonnage des coefficients de Fourier extraits dépend des droites analytiques discrètes utilisées pour calculer la transformée de Radon discrète. La transformée de Fourier inverse 1-D est calculée sur des coefficients de Fourier non équidistants et nous avons besoin d'une approximation des transformées de Fourier pour des données non équidistantes ou pour la transformée de Fourier polaire [33, 92, 164, 187, 239]. En 2-D, aucune de ces méthodes n'est conçue pour être inversible et les approches proposées dans [13, 12, 215] utilisent des méthodes itératives. La transformée Ridgelet discrète devait être rapide, simple et associée à une reconstruction exacte. Nous avons donc considéré que les droites analytiques discrètes 2-D sont des bonnes approximations des droites Euclidiennes pour chaque direction des droites et nous utilisons une transformée de Fourier discrète classique définie pour des coefficients équidistants.

Pour assurer la réversibilité de la transformée de Radon discrète, il faut que les droites discrètes couvrent complètement la grille de Fourier. Pour cela, les droites sont définies pour relier le pixel bordant le spectre fréquentiel et son symétrique par rapport au centre de l'image.

Proposition 1 Soit une grille carrée $\Omega_N^2 = [-N, N] \times [-N, N]$.

Soit l'ensemble des directions (p_m, q_m) avec, pour $0 \leq m \leq 2N$, $(p_m, q_m) = (N, m - N)$ et pour $2N + 1 \leq m \leq 4N - 2$, $(p_m, q_m) = (m - 3N + 1, N)$.

L'ensemble de toutes les droites fermées définies par $|p_m\xi_1 + q_m\xi_2| \leq \omega_m/2$ avec $\omega_m \geq \sup(|p_m|, |q_m|)$ fournit une couverture complète de la grille Ω_N^2 .

La preuve de cette proposition est évidente grâce à un résultat bien connu en géométrie analytique discrète qui énonce qu'une droite discrète fermée dans la direction (p, q) est connectée si et seulement si

$$\omega \geq \sup(|p|, |q|).$$

Cette définition implique de la redondance dans les coefficients des projections, c'est-à-dire qu'un coefficient de Fourier peut être élément de plusieurs droites discrètes. Les droites naïves fermées couvrent le domaine de Fourier avec peu de redondance (≈ 2) (voir figure 2.26 (a)), les droites pythagoriciennes fermées avec une redondance moyenne (≈ 2.3) et les droites supercouvertures avec beaucoup de redondance (≈ 3) (voir figure 2.26 (b)).

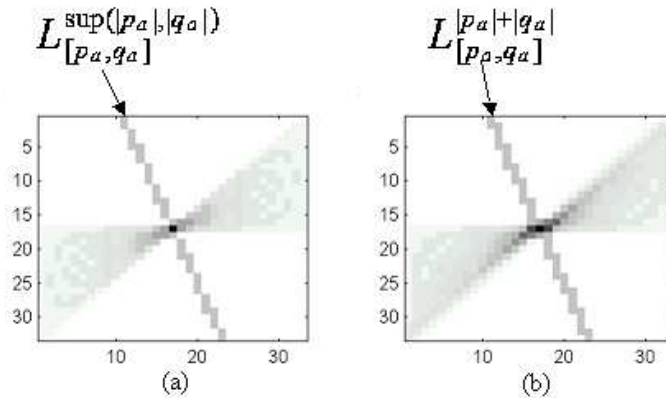


FIGURE 2.26 – Couverture de domaine de Fourier (a) avec des droites naïves fermées et (b) des droites supercouvertures (les niveaux de gris représentent le nombre de fois qu'un pixel est couvert par des droites)

2.2.3.3 Inversion de la transformée de Radon analytique discrète

Nous illustrons maintenant rapidement la stratégie d'inversion de notre transformée de Radon :

1. On calcule la transformée de Fourier 1-D pour chaque ensemble $R^\omega s([p, q], \cdot)$ afin d'obtenir $P_{[p, q]}^\omega$
2. Pour toutes les directions $[p, q]$, on substitue les valeurs échantillonnées de $\hat{s}$ quand les points sont éléments des lignes $L_{[p, q]}^\omega$ par les valeurs de $\hat{s}$ définies sur la grille carrée :

$$\hat{s}^{[p, q]}(\xi_1^k, \xi_2^k) = P_{[p, q]}^\omega(k) \text{ tel que } |q\xi_1^k - p\xi_2^k| \leq \frac{\omega}{2}$$

pour $0 < k < \text{longueur de } L_{[p, q]}^\omega$.

Notons que du fait de la redondance, des coefficients de Fourier sont éléments de plusieurs lignes discrètes. Dans ce cas, le coefficient de Fourier est calculé par une moyenne¹.

3. Calcul d'une transformée de Fourier inverse 2-D.

La procédure précédente permet d'obtenir une reconstruction exacte. Nous proposons maintenant de comparer expérimentalement les différentes transformées.

1. Le nombre de fois où un pixel de coordonnées (ξ_1, ξ_2) est élément d'une ligne discrète dépend de la fréquence (plus important pour les basses fréquences) et dépend du type de ligne discrète.

2.2.3.4 Illustration de la transformée de Radon analytique discrète

Nous illustrons les différentes transformées de Radon telles qu'elles sont définies dans le cadre des transformées Ridgelet. Notons que nous pouvons utiliser une procédure de concaténation de zéros dans la transformée de Radon analytique afin de supprimer les effets de bords (comme indiqué lors de l'étude de l'algorithme de Stanford). Cette opération n'est pas nécessaire pour obtenir une reconstruction parfaite mais elle élimine les phénomènes de "wraparound".

Nous proposons une comparaison numérique entre les différentes transformées. Les papiers de Stanford référencent une Toolbox [232] développée par les différents collègues permettant le calcul de la transformée Ridgelet et Curvelet. De même les codes de Do et Vetterli [68] sont ouverts (<http://lcavwww.epfl.ch/~minhdo>).

La figure 2.27 illustre les différents résultats pour une image ne contenant qu'une seule valeur non nulle. D'après l'étude préalable sur Radon, nous avons vu que nous devons obtenir un segment de droite.

Nous observons que :

- Les transformées de Radon Fast-Slant-Stack et Analytique sont des lignes.
- La transformée de Radon proposée par Do et Vetterli est difficile à interpréter.
- La transformée analytique présente un effet de wrap-around si l'on n'utilise pas une procédure de concaténation de 0 (figure 2.27a), et inversement sur la figure 2.27b.

La figure 2.27c illustre l'influence du choix du type de la ligne discrète. Pour cette figure, nous avons utilisé des lignes supercouvertures dans la transformée analytique ce qui rend plus difficile l'interprétation car la taille des projections n'est pas constante.

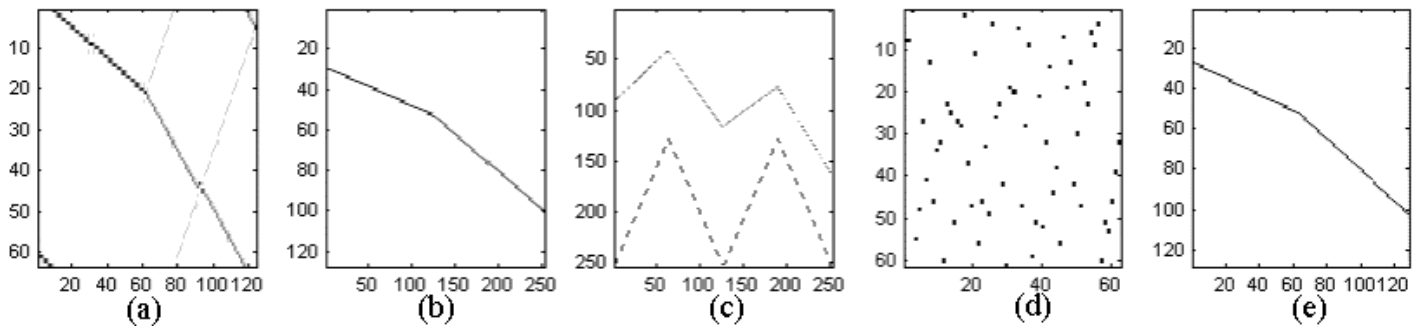


FIGURE 2.27 – La transformée d'un point : (a) avec la stratégie analytique avec des droites naïves sans ajout de 0, (b) avec la stratégie analytique avec des droites naïves avec ajout de 0, (c) avec la stratégie analytique avec des droites supercouvertures avec ajout de 0 (la ligne pointillée indique la longueur de chaque projection) (d) avec la stratégie de Do et Vetterli, (e) avec la Fast Slant Stack

A ce stade, nous constatons que notre proposition atteint ses objectifs : une transformée rapide inversible conservant l'interprétation géométrique de la transformée continue.

2.2.4 Transformée Ridgelet et applications

2.2.4.1 Calcul de la transformée DART

Pour obtenir la transformée Ridgelet, nous devons calculer la transformée ondelette 1-D le long des variables radiales dans l'espace de Radon. Le choix d'une ondelette unidimensionnelle est discuté dans

le cas 2-D par Starck et al. dans [232]. Ils indiquent que leurs expériences ont montré que des ondelettes de support compact peut entraîner différents artéfacts visuels quand on utilise un traitement non-linéaire. Notons que suivant cette idée, Flesia et al. [72] ont proposé d'utiliser des ondelettes de Meyer pour obtenir une décomposition Ridgelet basée sur des fonctions "true Ridge". Starck et al. utilisent un système redondant spécifique : ils définissent une fonction d'échelle dans le domaine fréquentiel pour assurer le lissage, et la fonction d'ondelette est la différence entre deux résolutions successives. Nous reviendrons sur ce choix de la base d'ondelettes dans le cadre du débruitage dans la section suivante.

Afin d'illustrer la fonction Ridgelet associée à une transformée DART, nous appliquons la procédure d'inversion à partir d'une matrice de coefficient ne contenant qu'un seul point non nul. La figure 2.28a illustre l'influence des artéfacts "wrap-around". Des figures 2.28b et 2.28c, on constate que la décomposition repose sur des fonctions Ridgelet correctes : la fonction $\psi_{a,b,\theta}$ est orientée selon l'angle θ , et est constante selon les lignes et c'est une fonction d'ondelette perpendiculairement à ces lignes.

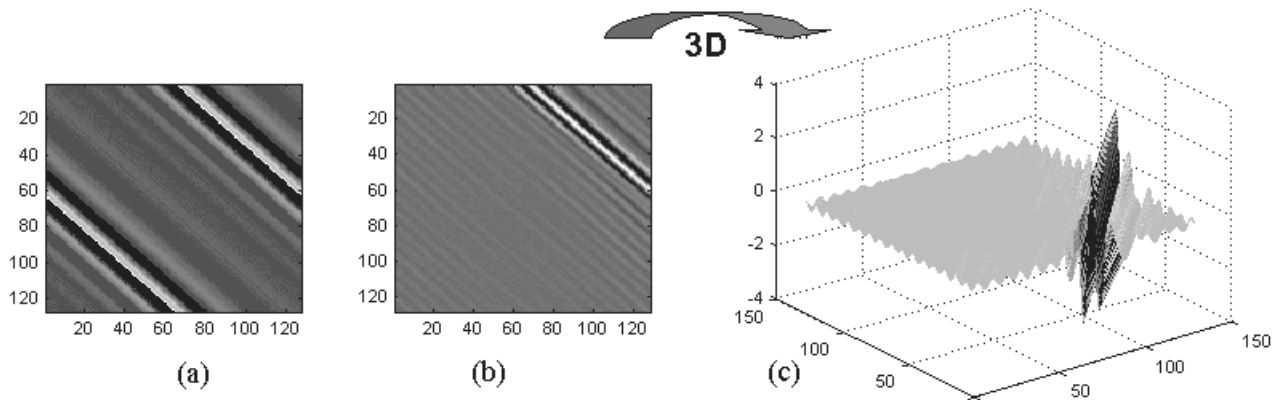


FIGURE 2.28 – La transformée Ridgelet analytique inverse d'un point : (a) utilisant les droites naïves sans zeropadding (b) utilisant les droites naïves avec la stratégie par ajout de 0 (c) vue 3-D de la fonction Ridgelet

2.2.4.2 Applications de la transformée Ridgelet 2-D dans la restauration d'images

Le travail de conception que nous avons mené dans le cadre de la transformée Ridgelet a trouvé une implication directe dans des problématiques de restauration de données.

La procédure de débruitage par transformée Ridgelet consiste en un simple seuillage des coefficients Ridgelet et par le calcul de la transformée inverse. Le seuillage est effectué à partir d'une méthode non-décimée développée dans le cadre des décompositions ondelettes [55]. La redondance de la décomposition en ondelettes réduit les artéfacts qui apparaissent après l'opération de seuillage.

Nous utilisons une opération de seuillage de type dur² avec un seuil α qui peut être défini par $\alpha = \sigma \sqrt{2 \log(N)}$ [70]. Pour l'estimation du niveau de bruit σ , nous considérons que ce niveau dépend de l'index de projection et est estimé en utilisant la valeur médiane absolue des coefficients d'ondelettes de la première échelle pour chaque projection radiale. Nous reprenons ici l'approche classique de débruitage par ondelettes développée pour le signaux 1-D.

2. mise à 0 des coefficients inférieurs à un seuil, conservation des autres

Influence des droites discrètes Nous proposons tout d'abord d'illustrer l'influence des différents types de droites dans le cadre de la DART. Pour cela nous avons généré une image artificielle (figure 2.29a) et ajouté un bruit Gaussien blanc additif très important (figure 2.29b). La mesure SNR³ de l'image bruitée est égale à 15 dB. Les figures 2.29 (c), (d) and (e) sont les résultats obtenus suivant les trois définitions des lignes discrètes : naïve, pythagoricienne et supercouverture. Nous pouvons constater qu'avec une décomposition redondante le débruitage est meilleur (figure 2.29e) que pour une représentation plus compacte (figure 2.29c). La rupture est reconstruite d'une manière plus précise et les zones uniformes sont mieux lissées. Pour une première comparaison, avec les approches classiques, nous présentons le résultat de la transformée en ondelettes classique décimée (figure 2.29f). La base utilisée est Daubechies D20. On constate une qualité visuelle médiocre dans l'image reconstruite.

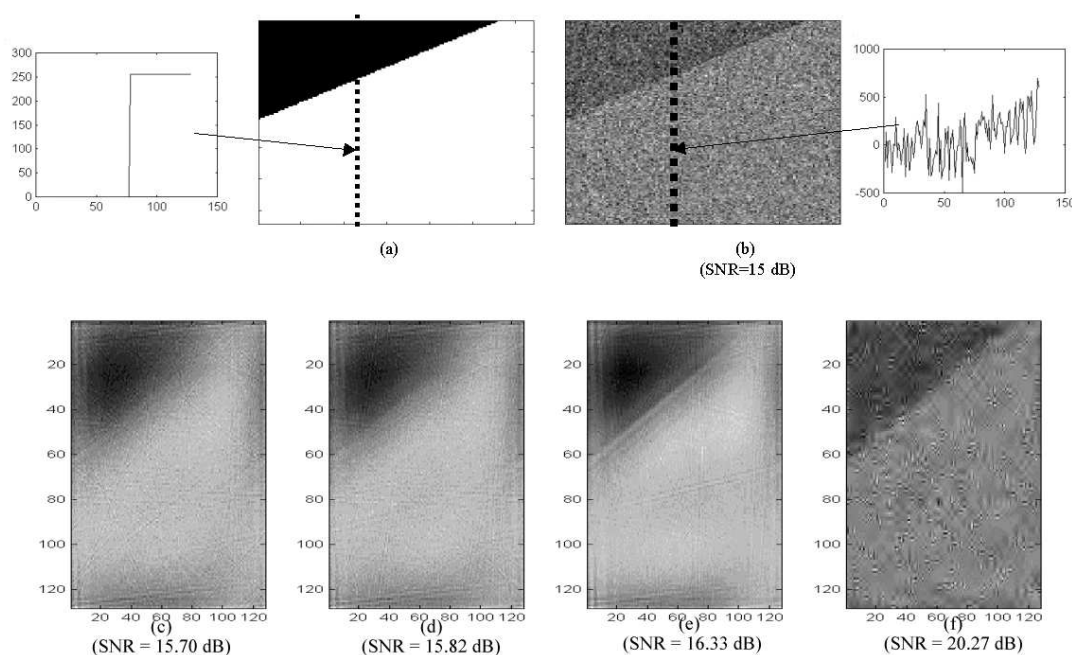


FIGURE 2.29 – (a) Image originale (b) Image bruitée (pour voir l'influence du bruit nous illustrons une colonne de l'image) (c) Débruitage avec ligne naïve (d) Débruitage avec ligne pythagoricienne (e) Débruitage avec ligne supercouverture

Pour comprendre l'influence de cette redondance, ainsi que la qualité de représentation associée à notre transformée, nous testons maintenant le problème de la reconstruction partielle, qui correspond à l'analyse du signal reconstruit à partir simplement de quelques coefficients (en général les plus importants). Nous retrouverons cette approche dans la suite de ce mémoire car elle permet, me semble-t-il, de caractériser empiriquement assez finement l'information captée par la transformée analysée.

La figure 2.30 compare la reconstruction partielle d'une image artificielle en utilisant seulement 30% des coefficients de la DART. Dans ce cas, nous constatons logiquement que la plus faible redondance associée aux droites naïves permet d'obtenir de bien meilleurs résultats.

3. pour éviter les problèmes de comparaison, nous utilisons la fonction SNR de Toolbox de Stanford.

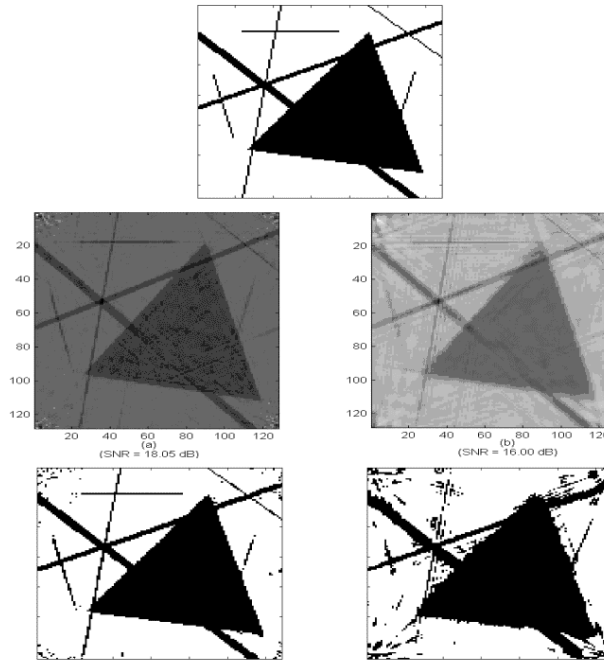


FIGURE 2.30 – Reconstruction partielle d’une image artificielle (a) naïve DART (b) supercouverture DART

Influence du choix de la base d’ondelettes Discutons maintenant de l’influence de la base d’ondelettes dans le cadre de la DART. Pour cela, différentes décompositions en ondelettes sont appliquées le long des projections angulaires. Pour cette expérimentation, nous utilisons uniquement les droites naïves. Nous utilisons les bases d’ondelettes suivantes :

1. Les ondelettes discrètes de Meyer (dont le support est compact dans le domaine de Fourier).
2. Les ondelettes Daubechies D20 (dont le support est compact dans le domaine temporel).
3. Les ondelettes Daubechies D20 utilisées dans un algorithme non-décimé.
4. Une décomposition redondante suivant la proposition de Starck et al. (dont le support est compact dans le domaine de Fourier) [232] .

La figure 2.31 considère un objet utilisé par Do et Vetterli dans leurs articles [68] et compare le débruitage de cette image par seuillage de la DART utilisant les différentes bases d’ondelettes. Nous observons que :

- les méthodes redondantes obtiennent des résultats supérieurs aux transformées décimées,
- avec une approche décimée, les ondelettes dont le support est compact dans le domaine de Fourier obtiennent un résultat supérieur aux ondelettes à support compact dans le domaine temporel.
- les deux transformées redondantes obtiennent des résultats similaires.

Nous avons mené des expérimentations à plus grande échelle et les constatations énoncées ci-dessus se sont confirmées. Nous avons donc décidé d’utiliser la transformée en ondelettes Daubechies D20 dans un algorithme non-décimé pour des problématiques de débruitage.

Comparaison des différentes transformées Ridgelet numériques La figure 2.32 considère l’objet utilisé par Do et Vetterli et compare le débruitage de cette image par le seuillage des différentes transformées

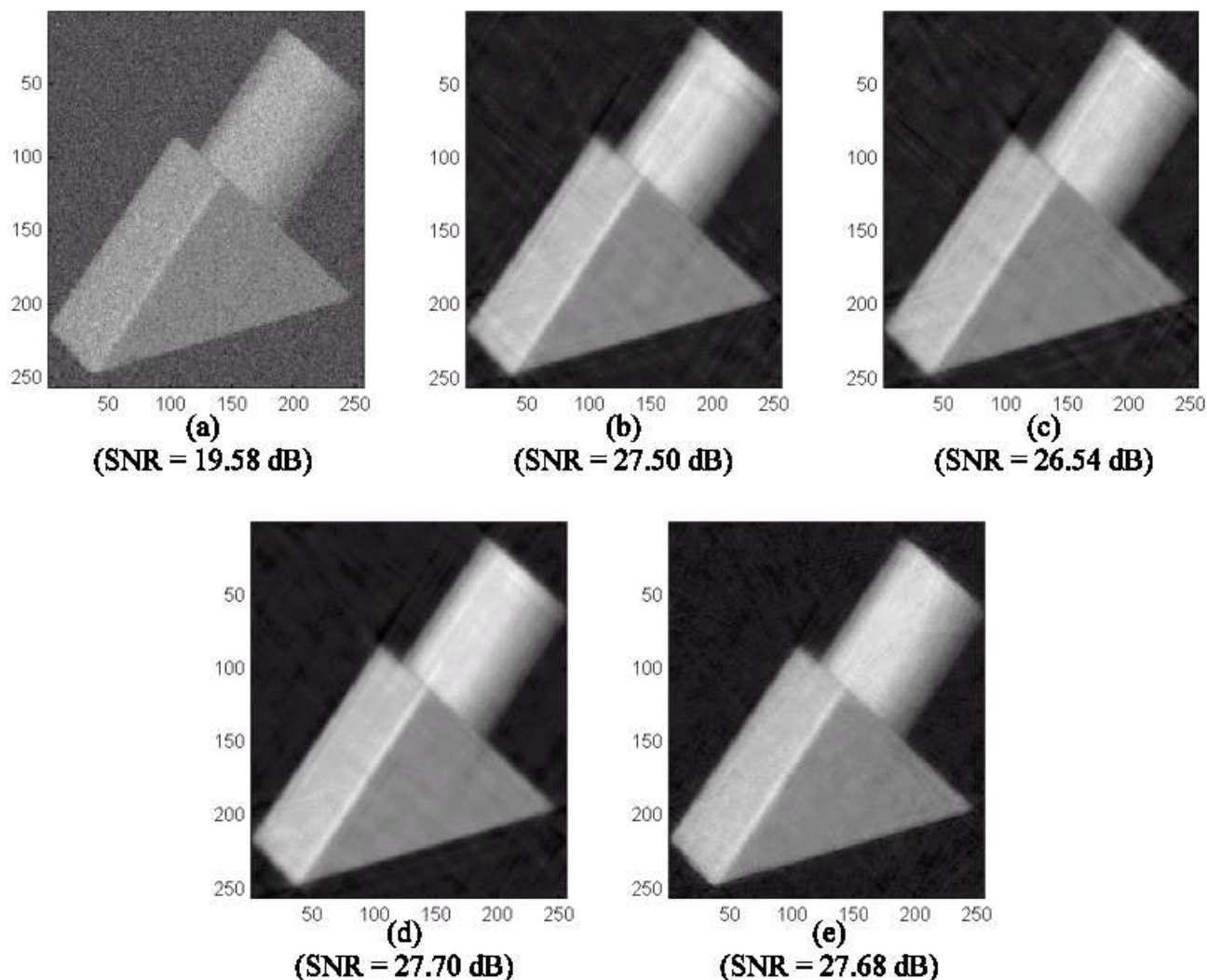


FIGURE 2.31 – (a) Image bruitée (b) résultat de la Meyer-DART (c) résultat de la D20-DART (d) résultat de la D20-DART non-décimée transform (e) résultat de la Starck-DART

Ridgelet : DART, stratégie de Lausanne, stratégie de Stanford.

Nous constatons que :

- les transformées Fast Slant Stack (Beamlab code) et analytique ne contiennent pas les nombreuses taches parasites présentes dans les reconstructions par ondelettes,
- La reconstruction de Do et al. présente de nombreux artefacts car la transformée correspond à un échantillonnage critique,
- La stratégie par ajout de 0 ne permet pas d'augmenter les résultats en terme de qualité de débruitage.

Nous avons mené d'autres expérimentations qui ont confirmé les conclusions exposées ci-dessus, comme par exemple sur la figure 2.33 avec l'image Lenna très bruitée.

Nous pouvons constater que l'image reconstruite après seuillage de la DART souffre de défauts "rectilignes" qui traversent l'image. Ceci est bien sûr dû au seuillage de certains coefficients (et constitue un phénomène équivalent au phénomène de pseudo-gibbs présent avec les ondelettes).

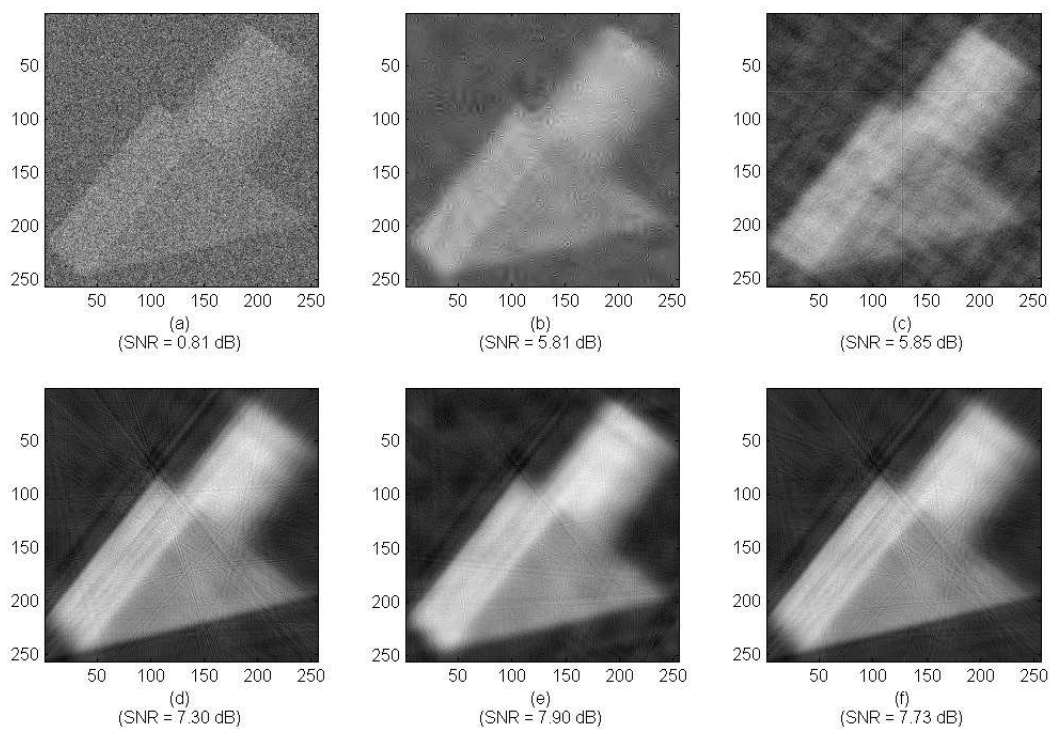


FIGURE 2.32 – (a) Image bruitée ; Image débruitée par (b) transformée en ondelette non-décimée (c) transformée de Do et Vetterli (d) transformée Fast Slant Stack et ondelettes D20 non-décimées (e) transformée analytique supercouverture et ondelettes D20 non-décimées sans zeropadding (f) transformée analytique supercouverture et ondelettes D20 non-décimées avec zeropadding.

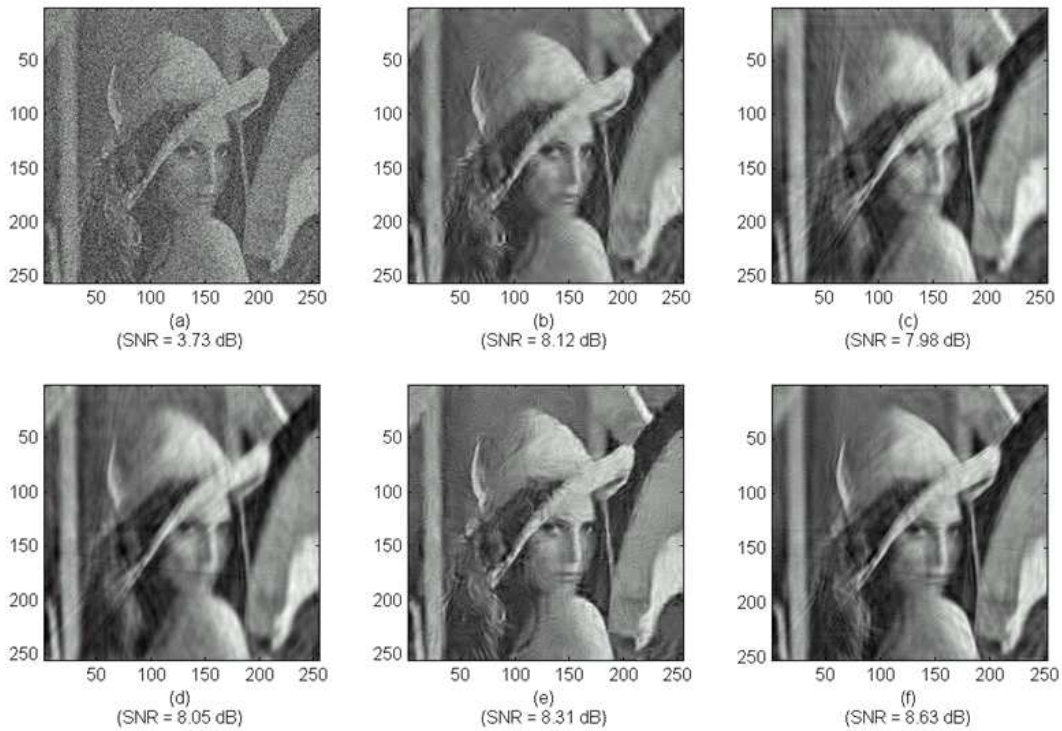


FIGURE 2.33 – (a) Image bruitée ; Image débruitée par (b) transformée en ondelette non-décimée (c) transformée de Do et Vetterli (d) transformée Fast Slant Stack et ondelettes D20 non-décimées (e) transformée analytique supercouverture et ondelettes D20 non-décimées sans ajout de zéros (f) transformée analytique supercouverture et ondelettes D20 non-décimées avec ajout de zéros.

2.2.5 Evolution de la DART

2.2.5.1 Transformée locale

Pour limiter le phénomène de défauts "rectilignes", la transformée Ridgelet peut être facilement étendue à une transformée locale en partitionnant l'image avec des fenêtres les plus régulières possibles. Ensuite chacune de ces zones ainsi définies est analysée suivant la transformée Ridgelet. Ce traitement permet d'avoir une analyse plus localisée dans le domaine spatial. La conséquence est la réduction du nombre de lignes parasites présentes lors du seuillage d'une transformée Ridgelet. Nous illustrons ceci avec l'image Lenna fortement bruitée, dans le cadre de la transformée Ridgelet analytique (Figure 2.34).

Nous observons que la procédure de fenêtrage permet ainsi la limitation des lignes parasites grâce à l'analyse locale.

2.2.5.2 Vers une transformée Curvelet

De même la transformée DART peut être étendue à la transformée Curvelet. La transformée Curvelet a pour objectif de représenter en un nombre minimal de coefficients les courbes présentes dans une image. Nous avons donc une évolution des ruptures linéaires avec les Ridgelet aux ruptures courbes avec les Cur-

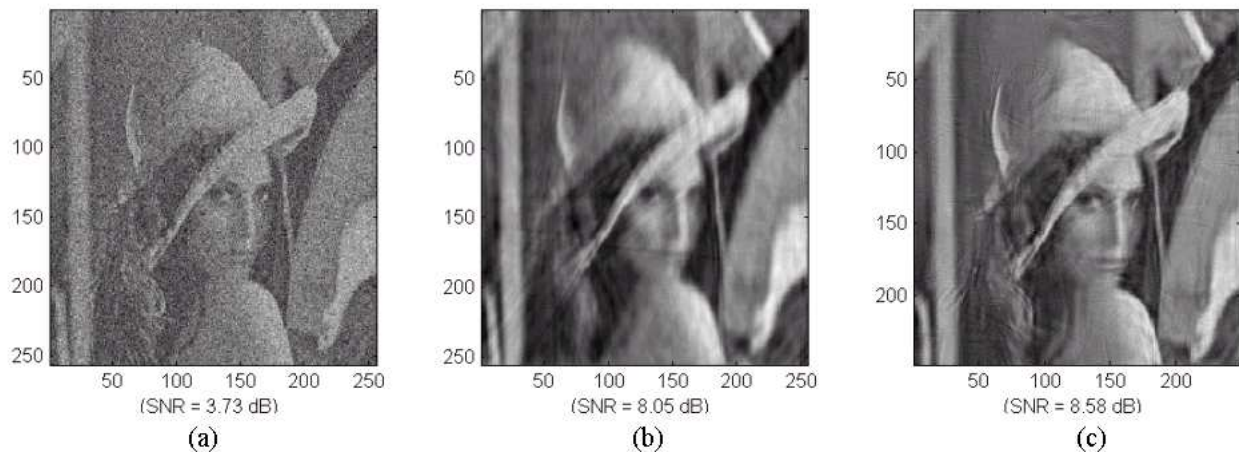


FIGURE 2.34 – (a) Image bruitée ; Image débruitée par (b) la transformée Ridgelet analytique (c) la transformée Ridgelet analytique fenêtrée sur des zones de taille 32

velet.

Il existe dans la littérature deux stratégies en ce qui concerne la réalisation numérique de cette transformée. La première stratégie de calcul repose directement sur l'utilisation de la transformée Ridgelet [232], et bien évidemment c'est dans cet axe que nous allons nous situer. Une seconde génération a été proposée récemment dans [251] qui repose sur des manipulations élégantes dans le domaine de Fourier. Nous ne traiterons pas ce point dans ce mémoire.

La transformée DART peut facilement s'étendre à la transformée Curvelet :

- décomposition de l'image en différentes bandes de fréquence,
- chaque bande de fréquence est fenêtrée en des zones carrées,
- chaque zone est analysée par la transformée Ridgelet DART.

Pour définir les différentes sous-bandes fréquentielles, Starck et al. proposent d'utiliser l'algorithme à trous avec une définition des plans d'ondelettes par simple soustraction de deux résolutions successives. Nous nous plaçons dans la même stratégie à savoir une décomposition non-séparable redondante, cependant nous remplaçons l'algorithme à trous par un banc de filtres en quinconce non-décimé afin de préserver l'orthogonalité entre les différentes bandes de fréquences, et avoir une meilleure conservation du niveau de bruit à travers les échelles dans le cas du bruit blanc Gaussien.

Comme pour la transformée Ridgelet, la procédure de débruitage consiste en un simple seuillage des coefficients Curvelet. Pour estimer le seuil, nous considérons que notre transformée DART-Curvelet conserve l'énergie (grâce à l'utilisation du schéma en quinconce), et donc le niveau de bruit est estimé avant la transformée Curvelet et intégré ensuite dans la procédure de seuillage.

Dans notre exemple, un bruit blanc Gaussien avec un écart-type de 60 est ajouté à l'image Lenna (figure 2.35a). Différentes méthodes sont testées pour débruiter l'image : seuillage d'une transformée en ondelettes non-décimée (figure 2.35b), seuillage de la DART-supercouverture (figure 2.35c), seuillage de la DART-supercouverture locale (figure 2.35d), seuillage de DART-Curvelet basée sur un schéma quinconce et des droites de supercouverture (figure 2.35 (e)).

Nous observons que la transformée DART-Curvelet et la transformée locale obtiennent les meilleurs

résultats, ce qui confirme que ces deux approches sont plus précises dans la description des informations 2-D.

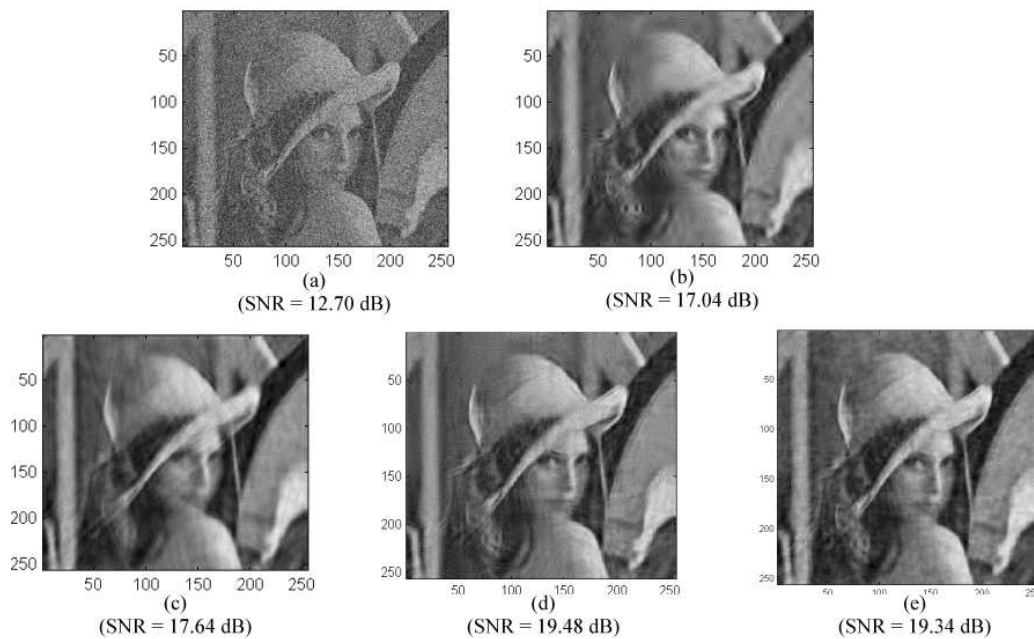


FIGURE 2.35 – (a) image bruitée (b) résultat d’une transformée en ondelettes non-décimée, (c) résultat de la DART-supercoverture, (d) résultat de la DART-supercoverture locale, (e) résultat de la DART-Curvelet basée sur un schéma quinconce et des droites de supercoverture

Afin de proposer une comparaison avec la transformée Curvelet de Stanford, nous avons repris les exemples débruitage proposés dans [232] : un bruit gaussien avec un écart-type de 20 est ajouté à l’image Lenna de taille 512 par 512. Les restaurations par les transformées DART locale et DART-Curvelet sont présentées figure 2.36. Les résultats de l’approche de Stanford sont visibles sur le site Web <http://www-stat.stanford.edu/~jstarck> et correspondent à l’article [232]. On ne constate pas de différences entre les deux stratégies.

Ces différentes expérimentations, nous proposons une mesure objective à partir du PSNR⁴. Nous indiquons les valeurs PSNR associées avec l’algorithme de Stanford [232]. De plus nous testons différents paramétrages de la transformée DART locale (taille des fenêtres).

	Bloc 64	Bloc 32	Bloc 16	Curvelet
DART	31.42	31.38	30.61	31.23
Stanford	30.79	30.97	30.87	31.95

Pour conclure, ces mesures confirment les impressions visuelles à savoir que les deux stratégies (DART et Stanford) obtiennent des résultats similaires. Ensuite, de nouveau nous n’avons pas noté de différences notables dans notre approche entre la transformée locale et la version Curvelet. Enfin le paramétrage de la transformée locale est un compromis, à savoir des tailles de fenêtres suffisamment petites pour assurer une

4. la mesure PSNR est calculée avec la fonction Matlab Stanford.

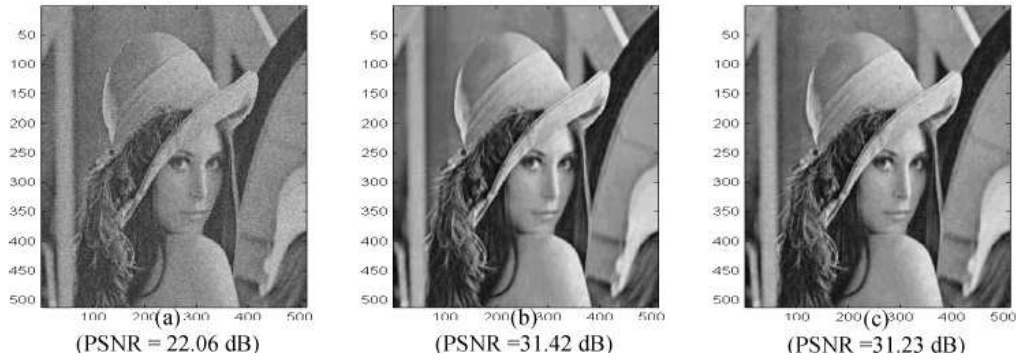


FIGURE 2.36 – (a) Image bruitée suivant le même protocole que [232] (b) restaurée par la transformée Dart locale (c) restaurée par la transformée DART-Curvelet.

analyse locale et suffisamment grandes pour assurer une restauration correcte par la DART. Notons que, comme nous le verrons par la suite, ce travail ouvre encore de nombreuses perspectives et surtout a des connexions inattendues avec le travail récent mené autour des ondelettes monogéniques.

La simplicité de calcul de la transformée analytique a permis son extension aux données 3-D [105, 106].

2.2.5.3 DART 3-D

Cette extension au cas 3-D s'est fait à travers l'utilisation de la généralisation des outils issus de la géométrie discrète. Ce travail a été fait notamment par D. Helbert sous ma co-direction avec E. Andrès. Nous allons donc définir une droite discrète 3-D dans le domaine de Fourier, comme nous l'illustrons sur la figure 2.37 [106].

Dans le domaine cartésien, la droite analytique discrète 3-D $L_{(p,q,r)}^{\omega_1, \omega_2, \omega_3}$ reliant les points cartésiens $O(0, 0, 0)$ et $Q(p, q, r)$ est définie par :

$$L_{(p,q,r)}^{\omega_1, \omega_2, \omega_3} = \pi_{oxy}(p, q, r) \cap \pi_{oyz}(p, q, r) \cap \pi_{ozx}(p, q, r) \quad (2.34)$$

avec :

$$\pi_{oxy}(p, q, r) = \left\{ (x, y, z) \in \mathbb{Z}^3 \mid |qx - py| \leq \frac{\omega_1}{2} \right\} \quad (2.35)$$

$$\pi_{oyz}(p, q, r) = \left\{ (x, y, z) \in \mathbb{Z}^3 \mid |ry - qz| \leq \frac{\omega_2}{2} \right\} \quad (2.36)$$

$$\pi_{ozx}(p, q, r) = \left\{ (x, y, z) \in \mathbb{Z}^3 \mid |pz - rx| \leq \frac{\omega_3}{2} \right\} \quad (2.37)$$

$qx - py = 0$ est l'équation de la droite 2-D reliant $(0, 0)$ et (p, q) et résulte de la projection orthogonale de OQ sur le plan Oxy .

ω_1 est fonction de p et q , ω_2 est fonction de q et r et ω_3 est fonction de p et r . Ces valeurs représentent l'épaisseur arithmétique de la droite analytique discrète 3-D dans la direction Ox , Oy et Oz .

Si $\pi_{oxy}(p, q, r)$, $\pi_{oyz}(p, q, r)$ et $\pi_{ozx}(p, q, r)$ sont des droites 2-D supercouvertures, alors $L_{(p,q,r)}^{\omega_1, \omega_2, \omega_3}$ est la supercouverture de la droite Euclidienne 3-D OQ (figure 2.37), étudiée en géométrie analytique discrète [7].

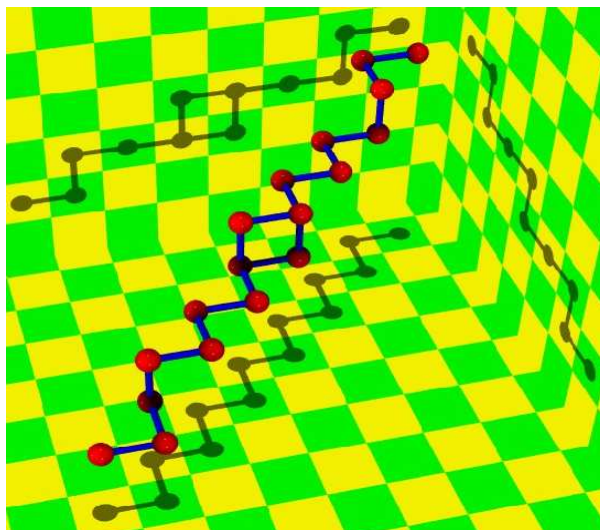


FIGURE 2.37 – Une ligne 3-D de supercouverture .

Une fois la droite 3-D définie, nous parcourons le cube, afin d’extraire toutes les droites possibles, comme nous le faisons en 2-D.

Une autre stratégie utilisant dans une première étape un plan 3-D, puis une droite 2-D parcourant ce plan a été testée. Nous ne détaillerons pas plus ces travaux similaires au 2-D en ce qui concerne le principe. Tous ces développements généralisant les principes développés précédemment dans le cadre 2-D sont décrits dans [105, 106] .

Comme pour la transformée 2-D, ce travail a été utilisé avec succès dans des problématiques de restauration de vidéos [106]. Nous reprenons dans les figures suivantes les principales expérimentations menées sur une vidéo fortement bruitée(figures 2.38 et 2.39). Nous pouvons constater la réussite de cette approche.

2.2.6 Discussion

Cette première partie a permis de poser les bases de notre travail, à travers la définition d’une transformée discrète inversible. Nous avons, à partir des concepts de la géométrie discrète proposé une représentation qui tente de prendre en compte les spécificités de l’information 2-D. Cet axe a concrétisé une collaboration entre deux thèmes du département XLIM-SIC et a donné lieu à différents co-encadrements et publications.

Toutefois, nous pouvons constater dans ce premier travail que la dimension couleur est encore absente. En effet, il est très difficile, dès la définition, d’intégrer la dimension vectorielle des données, si ce n’est à travers une redéfinition d’une transformée de Fourier adaptée à la couleur ce qui fera l’objet du chapitre suivant.

Nous avons mené des études pour la restauration d’images couleur (illustrées dans [39]). La figure 2.40 montre le débruitage de l’image Mandrill bruitée par un bruit blanc Gaussien d’un niveau égal à 77dB. Le traitement est appliqué individuellement pour chaque composante couleur de l’espace de représentation YCrCb. Nous pouvons observer que les zones homogènes sont perceptuellement bien restaurées.

Cependant, la transformée Ridgelet et ses évolutions font partie de la famille des bases fixes. Ainsi elles ne peuvent pas offrir, dans certains cas, la flexibilité pour s’adapter pleinement à la structure de l’image, notamment à sa dimension vectorielle. De cette constatation, nous avons développé un second axe d’étude



(a)

(b)

(c)

FIGURE 2.38 – 4 images extraites de la séquence : (a) vidéo originale, (b) vidéo bruitée (c) vidéo débruitée par la DART 3-D dans le domaine RGB. [106]

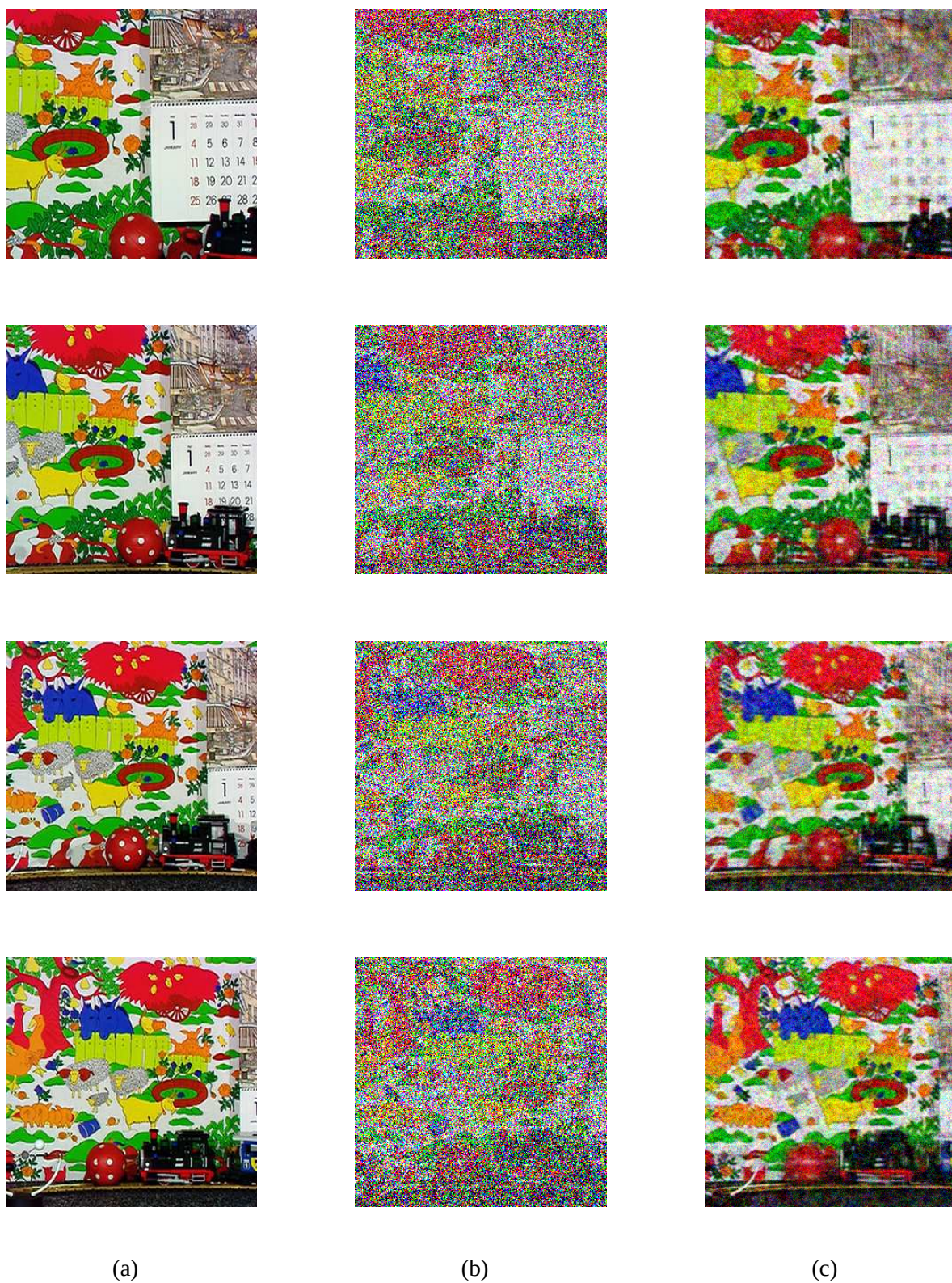


FIGURE 2.39 – 4 images extraites de la séquence : (a) vidéo originale, (b) vidéo bruitée (c) vidéo débruitée par la DART 3-D dans le domaine YCrCb. [106]

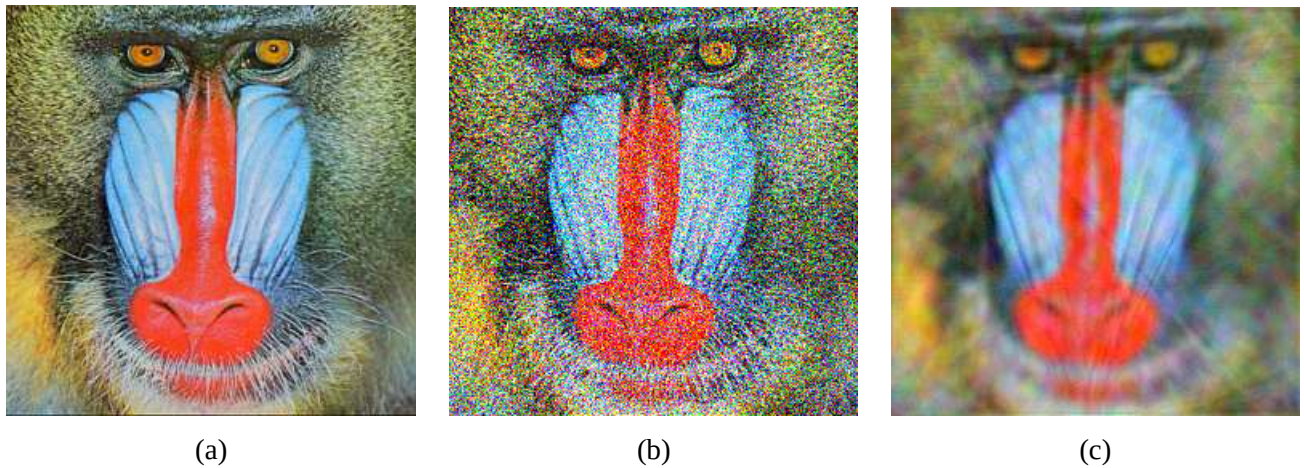


FIGURE 2.40 – Images (a) originale, (b) bruitée, (c) débruitée avec la DART

suivant une approche complémentaire : l'étude des transformées en ondelettes géométriques adaptatives. Cette étude s'est appuyée notamment sur la thèse de Guillaume Lebrun [132], thèse co-encadrée avec Stéphane Pateux de OrangeLabs, sous la direction de C. Fernandez. Cette thèse avait pour objectif la construction d'une décomposition en ondelettes adaptatives performante numériquement (à savoir que la dimension algorithmique était prépondérante) dans un cadre de restauration et de compression.

Après un travail pour définir des approches numériques dans le cadre des transformées bandelettes, nous avons proposé de construire une nouvelle approche adaptative multirésolution à travers le concept des variétés et plus précisément de la distance géodésique. Cet axe se poursuit actuellement et est décrit dans la section suivante.

2.3 Une transformée en ondelettes géodésiques

Comme nous l'avons vu, pour traiter les images, nous pouvons utiliser les transformées en ondelettes géométriques dont l'idée clé est d'adapter l'ensemble des fonctions d'ondelettes constituant la transformée (son dictionnaire) à la géométrie de l'image.

Des ensembles fixes aux ensembles adaptatifs, le gain sera une plus grande liberté et une plus grande précision dans la description des singularités de l'image. Néanmoins cette liberté se paye par la nécessité d'une bonne identification des singularités de l'image et en général le coût correspondant à la conservation de l'ensemble des paramètres décrivant le dictionnaire des fonctions d'ondelettes. Le choix entre un dictionnaire fixe pour toutes les images et un dictionnaire adapté à l'image en cours de traitement est une question de compromis. Nous allons aborder dans cette section le problème de la définition d'une transformée qui s'adapte à l'information analysée.

Nous allons également amorcer le problème de la prise en compte de la couleur. En effet, il n'existe pas à notre connaissance de transformées en ondelettes géométriques qui permettent d'intégrer la couleur durant le processus de transformation. Dans la plupart des cas, la seule alternative couleur conduit à une version séparable de la transformée. Le travail exposé dans la première partie de ce chapitre souffre de cette limite. C'est pourquoi nous avons proposé, à travers ce nouvel axe sur la transformée en ondelettes géométriques adaptatives, de prendre en compte la couleur, dans sa dimension vectorielle. Cette adaptation intervient lors de la modélisation de la géométrie de l'image.

Un exemple de traitement de l'image utilisant la couleur explicitement en tant que données vectorielles tout en prenant en compte la géométrie de l'image, est le filtrage anisotrope. Dans ce cadre nous pouvons par exemple citer les travaux de Tschumperlé et Deriche [65]. Basé sur la norme d'un gradient couleur correctement défini, chaque filtre est adapté à la configuration du voisinage du pixel en cours de traitement. Le but est dans ce cas d'obtenir la direction de variation maximale directement à partir du champ de vecteurs et ainsi d'obtenir un filtre adaptatif non séparable. Suivant une logique similaire, le filtrage bilatéral [240] (qui, nous le verrons par la suite, est le point de départ de notre proposition) est également une méthode de filtrage anisotrope. Le filtrage bilatéral est défini par une distance entre chaque pixel du voisinage au pixel central. Les deux méthodes définissent un filtre en fonction de la caractéristique du voisinage du pixel sur lequel s'applique le filtre.

Comme nous le verrons, le filtrage bilatéral et la transformée en ondelettes possèdent des points communs à travers le schéma Lifting. Le travail concernant la définition d'une nouvelle transformée en ondelettes géométriques a donc débuté à travers une réflexion autour du lissage bilatéral qui intègre une vraie adaptation aux informations 2-D.

Pour avoir une meilleure prise en compte de la dimension géométrique mais aussi couleur, notre proposition a été d'utiliser la distance géodésique. En effet, cette distance nous permet d'intégrer la topologie de l'ensemble des échantillons à travers notamment la connexité de nos échantillons. Ensuite, étant donné les liens mis en avant, nous intégrerons ce concept de distance géodésique dans le schéma Lifting afin de construire une transformée en ondelettes géométrique adaptative.

Ce travail a débuté dans le cadre de notre collaboration avec France Telecom R&D notamment avec la thèse de G. Lebrun [132] et fait toujours l'objet de recherche notamment dans sa dimension couleur.

2.3.1 Filtrage bilatéral et extension à travers la distance géodésique

2.3.1.1 Présentation du filtrage bilatéral

Le filtrage bilatéral proposé par Tomasi et al. [240], repris par Paris et al. [169] (comme alternative au filtrage anisotrope), correspond à l'adaptation d'un filtrage gaussien à partir des valeurs des éléments en pondérant la participation de chaque échantillon dans le support du filtre, noté Ω , en fonction de sa ressemblance avec l'élément central traité de coordonnée x :

$$c(x) = \frac{1}{n(x)} \int_{\Omega} s(t)h(t, x)w(s(t), s(x))dt \quad (2.38)$$

avec le dénominateur $n(x)$ défini par :

$$n(x) = \int_{\Omega} h(t, x)w(s(t), s(x))dt. \quad (2.39)$$

La fonction s correspond à notre signal et la fonction h définit le filtre classique (à savoir sa réponse impulsionnelle). La fonction w définit la pondération suivant les échantillons. Si w est constante et égale à 1, nous retrouvons simplement l'opération de convolution entre s et h . Dans la proposition initiale [240], le filtrage utilisé est un filtrage Gaussien :

$$h(t, x) = e^{-\frac{1}{2}\left(\frac{d(t, x)}{\sigma_d}\right)^2}$$

En général, la formulation utilise directement une notion de distance entre les coordonnées x de l'élément à filtrer $s(x)$ et celles t de son voisin $s(t)$, comme par exemple $d(t, x) = \|t - x\|_2$.

La fonction w permet d'incorporer le voisinage dans le filtrage et constitue l'originalité du filtrage bilatéral. Cette pondération est aussi définie à travers une fonction Gaussienne :

$$w(s(t), s(x)) = e^{-\frac{1}{2}\left(\frac{\delta(s(t), s(x))}{\sigma_r}\right)^2}$$

Nous constatons que cette pondération intègre la distance entre les valeurs des échantillons $\delta(s(t), s(x)) = \|s(t) - s(x)\|_2$. En plus de la proximité sur la grille, cela permet ainsi d'intégrer dans le filtrage la proximité en terme de valeur photométrique.

Des aspects pratiques sont discutés dans [240] dont en particulier la mise en évidence du respect des contours de l'image par ce filtrage. L'influence de w par rapport à h est étudiée afin de montrer dans quelle mesure les contours seront respectés en modifiant σ_d et σ_r .

Nous avons un filtrage adaptatif basé sur la distance entre les valeurs des échantillons. Nous montrons dans la section suivante qu'en abordant le signal à traiter suivant le concept des variétés nous retrouvons une approche similaire au filtrage bilatéral.

Nous avons proposé une introduction plus complète de la géométrie avec l'aide de l'utilisation du concept de variétés, et implicitement le contexte que l'on retrouve dans la notion de flot géométrique.

2.3.1.2 Introduction de la notion de variété

Il n'est pas aisé de donner une description simple et brève de la notion de variété car elle requiert des notions mathématiques vastes. Nous allons essayer d'extraire ici les éléments principaux. Une variété est un espace topologique qui localement se comporte comme un espace euclidien. Intuitivement, une variété

est un espace "courbe" de $\mathbb{R}^N$. Par exemple, une sphère ou un tore sont des variétés de $\mathbb{R}^3$ ou en général, toute surface est une variété de $\mathbb{R}^3$. Prenons l'exemple de la sphère. La sphère est une variété de $\mathbb{R}^3$. Si l'on observe une sphère de suffisamment près, on peut approximer la courbure de sa surface par un plan tangent. C'est exactement ce que nous faisons lorsque l'on dessine une carte géographique : on approche la courbure de la terre par un plan. Par définition, une variété différentielle est une variété qui est localement difféomorphe à un espace vectoriel de dimension fixée.

Dans la même idée, prenons les habitants vivant sur un plan (donc dans un monde à deux dimensions). Ils n'auraient pas beaucoup de mal à formuler la géométrie comme nous le faisons nous. Prenons maintenant le cas des habitants d'une tôle ondulée. Pour nous qui les regardons d'une dimension supérieure, il est évident que leur monde n'est pas plat mais ils peuvent élaborer également une géométrie (qu'est ce un point, une droite, ...). Il n'y a aucune raison de ne pas généraliser cet approche aux dimensions supérieures (exemple tiré du cours de B. Houchmandzadeh sur les variétés différentielles).

Dans le cadre des variétés différentielles, la première étape est de se donner un ensemble de points P . Nous enrichissons ensuite cet ensemble en définissant une topologie, c'est à dire essentiellement que nous sommes capables de définir le voisinage d'un point de P . Enfin, dernière étape, nous disons que P est une variété de dimension N si nous pouvons le munir d'un système de coordonnées, c'est à dire si nous pouvons repérer de façon unique chaque point de P par un n-tuplet de $\mathbb{R}^n$. Nous noterons les coordonnées d'un point par $(\mu_1; \mu_2; \dots; \mu_n)$. L'exemple le plus simple est notre espace euclidien classique à trois dimension munie des coordonnées cartésiennes (x, y, z) ; la tôle ondulée dont la surface est donnée par $z = \sin x$ est une variété à deux dimensions, chaque point pouvant être repéré par un couple (x, y) .

Dans le cadre de l'image, la variété peut être une manière d'entrevoir les signaux en associant coordonnées et valeurs des échantillons dans un ensemble que nous appelons la variété dans un espace euclidien de dimension $n_c + n_v$ (respectivement dimension de l'espace des coordonnées et de l'espace des valeurs des échantillons). Cela veut dire par exemple qu'une image en niveau de gris correspondra à une variété de dimension 2 plongée dans un espace euclidien de dimension 3. La troisième dimension peut être directement la valeur du pixel à la position ou une valeur dérivée de l'image (gradient par exemple). A notre connaissance l'un des travaux fondateurs dans l'utilisation des variétés dans le cadre du traitement d'image, notamment couleur a été l'article de Sochen [218].

L'idée est donc de considérer l'image multivariée

$$s : \mathbf{x} = (x_i)_{i=1,2} \in \mathbb{R}^2 \rightarrow s(\mathbf{x}) = (s_j(\mathbf{x}))_{j=1\dots P} \in \mathbb{R}^P$$

comme une surface paramétrée plongée dans un espace de dimension $\mathbb{R}^{2+P}$

$$\begin{aligned} S : \Omega &\mapsto M \\ \{\mu_1, \mu_2\} &\mapsto \{(x_i(\mu_1, \mu_2))_{i=1,\dots,N}, (s_j(\mu_1, \mu_2))_{j=1\dots P}\} \end{aligned}$$

avec Ω la surface bidimensionnelle, (μ_1, μ_2) les coordonnées sur cette surface et M la variété. En général on utilise l'identification suivante $\mu_1 \equiv x$ et $\mu_2 \equiv y$.

Comme nous l'avons vu, l'un des éléments important est en général la construction d'un voisinage à chaque élément de l'ensemble étudié. Cette notion de voisinage correspond en fait à l'association d'une topologie à notre ensemble. La topologie est une notion définissant l'ensemble des sous-ensembles formés

uniquement à partir des éléments de la variété et constituant nos voisinages. Nous remarquons qu'en introduisant la topologie, ou plutôt en faisant le choix de modéliser le signal par une variété, nous introduisons la question de la connexité de cet ensemble.

Si nous reprenons la définition générale, l'identification est effectuée par une fonction définie sur le voisinage appelée carte de la variété. Le but de cette carte est de permettre l'identification d'une partie de la variété à un espace euclidien. Par exemple si le point de vue est suffisamment local, la surface d'une sphère peut être considérée plane. Chaque élément dans ce plan est décrit à partir de coordonnées ayant uniquement 2 composantes. L'intérêt du nouveau système de coordonnées de l'espace des paramètres, par rapport à celui de l'espace de plongement, est qu'il ne souffre plus des problèmes liés à la forme de l'ensemble (le plus court chemin dans une variété est un segment dans l'espace des paramètres).

C'est justement ce dernier point que nous souhaitons exploiter avec les variétés. En particulier, nous nous intéressons à la distance géodésique correspondant à la longueur des arcs géodésiques de la variété. Cet arc correspond au plus court chemin entre deux éléments de la variété dont tous les points qui le composent sont inclus dans la variété.

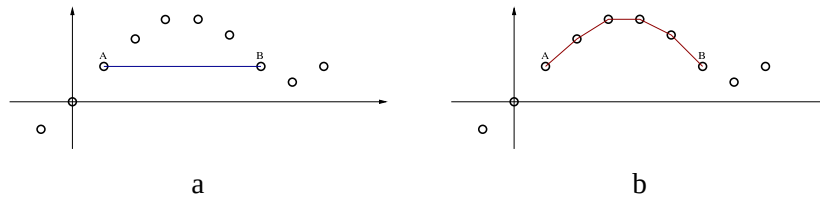


FIGURE 2.41 – Distances de deux éléments d'un sous-ensemble de $\mathbb{R}^2$: (a) distance euclidienne, (b) distance géodésique liée à la variété formée par les éléments de $\mathbb{R}^2$.

La figure 2.41 nous montre une variété plongée dans un espace 2D et représentée par un sous-ensemble fini de ses éléments. La figure illustre la différence entre la distance associée à la variété et la distance associée à l'espace de plongement. Nous pouvons donc voir à travers la figure 2.41 la modélisation associée à la distance géodésique : cette distance prend en compte la configuration globale de la variété. En particulier, cette distance pourra être infinie si les éléments de la variété n'appartiennent pas à la même composante connexe.

Nous allons maintenant chercher à exploiter cette qualité de description dans notre filtrage bilatéral.

2.3.1.3 Généralisation du filtrage bilatéral

Dans le cadre du travail mené au laboratoire XLIM-SIC, l'idée est d'intégrer la modélisation de l'image multivariée dans une décomposition en ondelettes. Pour cela, la première étape a consisté à intégrer ce formalisme dans le cadre du filtrage bilatéral à travers une modélisation du signal s'appuyant sur le concept des variétés. Nous considérons donc l'image s comme une surface plongée dans un espace de dimension $\mathbb{R}^{2+P}$

$$\{x, y, (s_j(x, y))_{j=1 \dots P}\}$$

Nous proposons de redéfinir le filtrage bilatéral avec ce nouveau point de vue. Pour cela, nous introduisons les éléments $S(\mathbf{x}) = (\mathbf{x}, s(\mathbf{x}))$ avec $x \in \mathbb{R}$ associant dans un même vecteur les coordonnées et les

valeurs des échantillons étudiés et σ_v qui correspond à la caractéristique de la fonction Gaussienne appliquée à notre signal identifié comme une variété. Le filtrage bilatéral peut alors s'écrire :

$$C_v(\mathbf{x}) = \frac{1}{n_v(\mathbf{x})} \int_{\Omega} S(\mathbf{t}) g(D(S(\mathbf{t}), S(\mathbf{x})), \sigma_v) d\mathbf{t} \quad (2.40)$$

Toujours avec notre dénominateur :

$$n_v(\mathbf{x}) = \int_{\Omega} g(D(S(\mathbf{t}), S(\mathbf{x})), \sigma_v) d\mathbf{t} \quad (2.41)$$

Dans cette nouvelle formulation, nous avons associé coordonnées et valeurs des éléments, et nous voyons apparaître une nouvelle définition de la distance entre le pixel étudié et ses voisins $D(S(\mathbf{t}), S(\mathbf{x}))$. En effet, la distance que nous utilisons correspond en fait à la distance entre deux éléments de la variété. Dans ce cas, durant le filtrage, la somme des voisins n'est plus seulement pondérée par une quantité qui dépend de la distance sur la grille mais d'une distance dépendant de l'organisation du signal lui-même.

Comme nous le constatons l'élément central est le calcul de la distance entre deux éléments de la variété, distance qui doit prendre pleinement en compte l'organisation des données. Cette extension du calcul de la distance dans le cadre du filtrage bilatéral nous conduit tout naturellement à l'utilisation de la distance géodésique qui permet de prendre en compte les connexités qui existent dans le signal.

Pour conclure sur cette partie introductive, il faut noter qu'il existe différents travaux associant dans la manipulation des images un processus de "lissage" et la modélisation suivant le concept des variétés. Tout d'abord, pour des données un peu différentes, on peut trouver de nombreuses études concernant la manipulation de maillage qui utilisent cette notion de distance géodésique et de variétés [185]. Ensuite, indépendamment de nos travaux menés durant les années 2007-2008 sur l'extension du filtrage bilatéral, est parue à la fin de notre travail, une publication majeure de Sochen et al. [218] qui décrit une réflexion à travers la régularisation d'image suivant le concept du flot géométrique de Beltrami. L'utilisation du flot de Beltrami dans le cadre des images permet de faire une régularisation prenant en compte la géométrie de l'image (justement à travers le concept de variétés). A l'origine, algorithmiquement, cette régularisation est un processus itératif, mais une simplification à travers une formulation par un simple produit de convolution peut être faite. De cette formulation, on peut établir une connexion avec le filtrage bilatéral et ainsi proposer d'insérer une meilleure intégration de la géométrie dans le filtrage bilatéral. Voici en quelques mots, le cheminement proposé entre autre dans cet article très dense de Sochen et al. [230]. Nous retrouvons une généralisation similaire à celle abordée ci-dessus, à savoir l'utilisation d'une distance dans le filtrage prenant en compte l'organisation des données.

2.3.2 Notion de distance géodésique en traitement d'images

La distance géodésique trouve plusieurs applications et interprétations dans les domaines de l'imagerie et de l'informatique graphique. Nous proposons tout d'abord de reprendre la présentation faite par L. Cohen [54] qui décrit la distance géodésique à travers le chemin minimum connectant deux pixels d'une image dont les valeurs définissent les caractéristiques du milieu de propagation d'une onde lumineuse. Dans cette interprétation physique, les chemins minimaux correspondent alors au principe de Fermat de propagation de la lumière. Dans les milieux homogènes le front d'onde est régulier du fait de la propagation isotrope de ce front et dans les milieux non homogènes ce front prend des formes plus complexes en fonction, par exemple, de frontières entre deux milieux impliquant des phénomènes de réfraction.

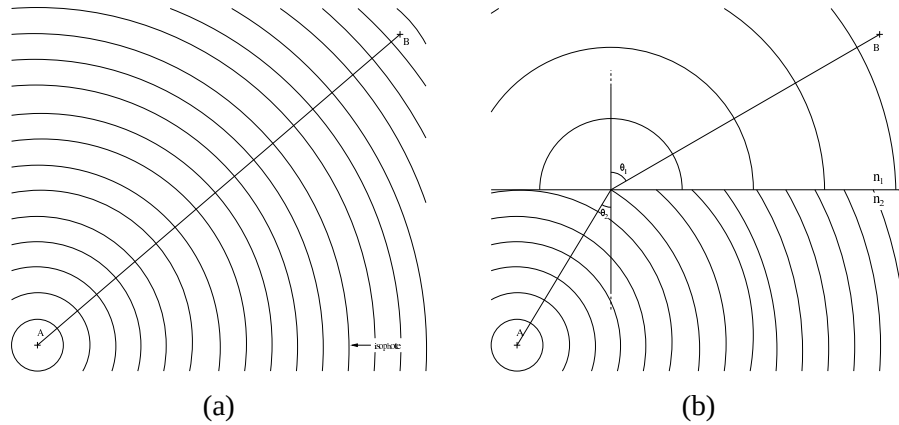


FIGURE 2.42 – Plus court chemin emprunté par la lumière pour (a) un milieu homogène et (b) deux milieux d'indice n_1 et n_2 .

Le schéma de la figure 2.42(a) montre le chemin le plus court de la lumière se propageant dans un milieu homogène. Du fait de la présence de deux milieux d'indice de réfraction différents n_0 et n_1 , le chemin de l'image 2.42(b) entre les deux points n'est plus une droite.

D'une manière générale, la distance géodésique entre les deux points p_0 et p_1 va correspondre au chemin suivi par la lumière et dépend de la minimisation de

$$D(p_0, p_1) = \frac{1}{c_l} \int_{p_0}^{p_1} nr(x) dx$$

avec c la vitesse de la lumière dans le vide et nr l'indice de réfraction [54]. L'une des applications directes est son utilisation dans le cadre des contours actifs, où le contour est considéré comme un arc géodésique dont l'énergie est mesurée en intégrant $n(s)$.

Ce point de vue nous permet d'entrevoir l'aspect recherché à travers cette distance : la mesure fournie par cette distance ne dépend pas seulement des deux éléments impliqués dans la mesure, mais bien de tout l'ensemble par lequel peut potentiellement passer le plus court chemin, et donc de l'organisation des données.

Dans le cadre des variétés, nous considérons la distance géodésique en tant que longueur des arcs géodésiques de la variété. Cet arc correspond au plus court chemin entre deux éléments de la variété dont tous les points qui le composent sont inclus dans la variété, comme nous l'avons illustré sur la figure 2.41.

La manière numérique dont est calculée la distance va suivre le même principe, à savoir que l'on débute d'un des deux points mesurés et l'on estime la quantité T pour tous les points de la forme par propagation de voisins en voisins (algorithme Fast Marching [213]). La distance est obtenue lorsque l'on atteint le deuxième point. Cette méthode, très proche de l'algorithme de recherche des plus courts chemins proposé par Dijkstra [66] que nous verrons par la suite, par sa propagation de voisins en voisins, illustre la nécessité d'associer une topologie à l'ensemble (par exemple la notion de voisins) pour lequel nous mesurons la distance géodésique.

Dans notre méthodologie, la distance géodésique doit remettre en cause la topologie classiquement associée aux signaux qui se base uniquement sur les coordonnées des échantillons pour définir les relations de voisinage. C'est pourquoi dans le travail mené au département XLIM-SIC, nous avons choisi d'utiliser une topologie définie par le graphe des k plus proches voisins de l'ensemble des pixels. Le choix de cette méthode ainsi que de l'algorithme de Dijkstra [66] pour obtenir au final la distance géodésique s'inspire de la

méthode d'identification de variété de Tenenbaum M. et al. [237]. En effet, nous posons les mêmes questions que les différentes techniques développées ces dernières années dans les problèmes d'analyse de données suivant des approches non linéaires, qui cherchent à extraire les structures dominantes dans des nuages de points. Pour illustrer ce point, nous reprenons l'une des figures de l'article présentant la méthode Isomap [237] sur la figure 2.43 qui met bien en évidence cette concordance scientifique entre les deux thématiques.

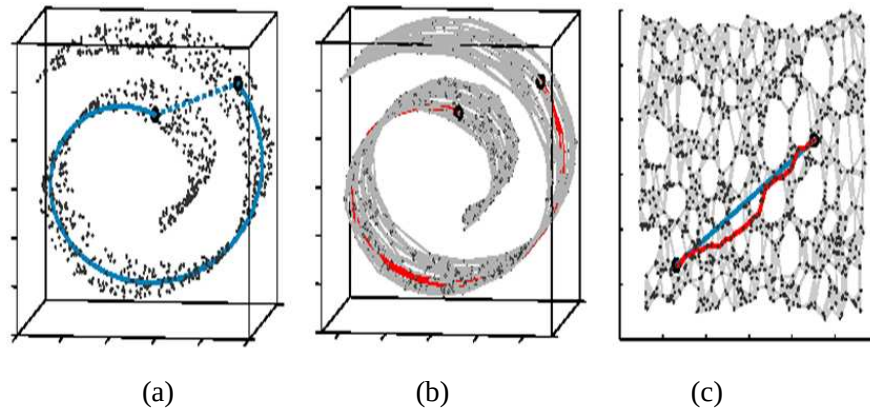


FIGURE 2.43 – Application de l'algorithme Isomap sur les données Swiss Roll (a) calcul de la distance entre deux points arbitraires, distance ne respectant pas la structure des données, (b) graphe des plus proches voisins permettant l'estimation de la distance géodésique entre ces deux points, (c) projection calculée par l'algorithme Isomap après identification de la variété.

Cette proximité avec les techniques d'analyse de données a renforcé notre intérêt dans l'approche utilisant la distance géodésique, car ces méthodes d'analyse s'adaptent tout naturellement à des données vectorielles, ce qui reste l'un de nos objectifs principaux pour la définition de notre transformée en ondelettes.

Nous présentons maintenant l'aspect numérique proposé, permettant l'identification de l'organisation géométrique de l'image.

2.3.3 Extraction de la topologie

2.3.3.1 Principe

Nous avons choisi d'utiliser une topologie définie par le graphe des k plus proches voisins de l'ensemble des pixels. Le choix de cette méthode ainsi que de l'algorithme de Dijkstra [66] pour obtenir au final la distance géodésique s'inspire de la méthode d'identification de variété de Tenenbaum et al. [237]. Isomap est une technique de réduction de dimension non-linéaire qui tient compte de l'ensemble des données. Elle va être définie en trois étapes :

- construire un graphe de voisinage, pour cela on définit un graphe G sur tous les points en connectant les points p_1 et p_2 s'ils sont soit plus proches qu'un seuil ϵ (ϵ -Isomap) ou soit si p_1 est un k plus proche voisin de p_2 (k -Isomap). La longueur des arcs est la distance entre p_1 et p_2 $d_{p_1 p_2}$.
- Calculer les distances point à point les plus courtes sur le graphe, pour cela il est utilisé une programmation dynamique (algorithme de Floyd ou Dijkstra).
- Construire la réduction en d -dimension suivant cette matrice de dissimilarité.

Dans notre approche, seules les deux premières étapes nous intéressent puisque nous cherchons la construction d'une topologie dont le but est de s'adapter aux singularités de l'image. En lien avec la distance géodésique qui en découle, cette topologie va nous permettre de prendre en compte la géométrie durant notre filtrage. Comme nous l'avons précisé, il est intéressant de noter que les données manipulées dans le cadre de ces méthodes sont de dimensions variées. Ceci va nous permettre de prendre en compte la dimension "couleur" directement dans l'algorithme de calcul du filtrage.

Pour construire cette topologie adaptative, il nous faut construire le graphe des k plus proches voisins. Nous proposons de revoir ci-dessous les choix qui ont été faits durant la thèse de G. Lebrun pour l'extraction de cette topologie adaptative [132]. Nous verrons que ces choix ont une influence sur la qualité de la représentation.

2.3.3.2 Graphe des k plus proches voisins

Pour obtenir le graphe des k plus proches voisins, on peut par exemple s'inspirer des algorithmes s'appuyant sur le concept de "distance partielle" tels que [206, 156]. Ces algorithmes s'appliquent à un nuage de points qui dans le cadre de notre signal correspond à notre variété S . Nous structurons alors S de telle manière que l'on puisse déterminer une distance minimum (distance partielle) entre l'élément dont le voisinage est en cours de construction et les éléments d'un sous-ensemble U de S . Si cette distance est supérieure à la distance du plus éloigné des voisins en cours d'étude, alors il n'est pas nécessaire de traiter les éléments de U . Après exécution, ces algorithmes fournissent un graphe orienté liant chaque éléments à ses k plus proches voisins selon un critère de distance prenant en compte l'organisation du nuage de points (et non selon la structure régulière des signaux). Nous constatons qu'à travers ce graphe, nous prenons en compte les singularités de l'image.

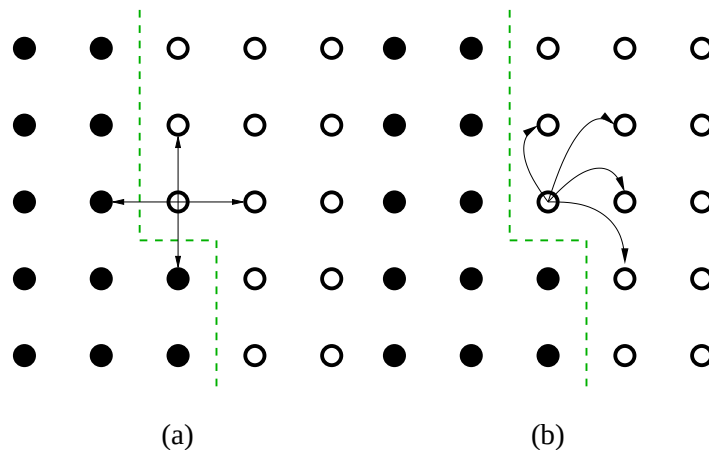


FIGURE 2.44 – Deux exemples de topologie : (a) le classique voisinage 4 connexes et (b) les 4 plus proches voisins. La ligne en pointillé représente une singularité dans le signal.

Si nous prenons l'exemple de la figure 2.44, nous remarquons que dans le cadre du voisinage 4 connexe, nous pouvons mettre en relation des échantillons qui ne sont pourtant pas liés si l'on considère le fait qu'ils sont de part et d'autre d'une singularité de l'image. À travers le graphe des k plus proches voisins, nous avons la construction d'un voisinage prenant en compte la présence ou non d'une singularité.

Toutefois, pour définir la distance géodésique à partir de ce graphe il est nécessaire de le rendre non

orienté. En effet, si nous prenons la définition classique d'une distance, nous ne vérifions pas à ce stade la condition de symétrie, à savoir $d(x, y) = d(y, x)$.

Par exemple, pour un chemin élémentaire correspondant à un arc entre deux sommets, il n'existe pas nécessairement de chemin inverse élémentaire correspondant à l'arc symétrique.

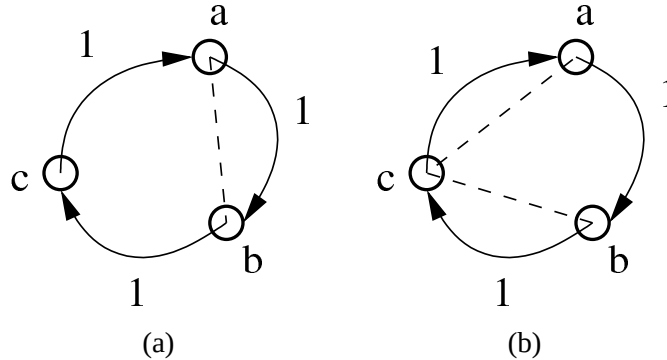


FIGURE 2.45 – Incompatibilité de la notion de distance avec les graphes orientés : (a) Le plus court chemin de a à b est de longueur 1. (b) Le plus court chemin de b à a est de longueur 2. Les plus courts chemins sont représentés par les segments en pointillés.

Le graphe de la figure 2.45 est un contre-exemple montrant que la définition d'une distance symétrique à partir d'un graphe orienté "quelconque" ne peut être satisfaisante sans ajout de contraintes sur la définition de ce graphe. Nous devons donc obtenir un graphe non orienté et se pose alors la question du passage d'un ensemble d'arcs à un ensemble d'arêtes pour le graphe dans son ensemble.

2.3.3.3 Construction du graphe non orienté

Durant la thèse de G. Lebrun, deux stratégies simples de construction d'un graphe orienté symétrique (un graphe strict, non orienté, valué) ont été testées. Les valeurs associées aux arcs sont les distances entre les éléments de S correspondant aux extrémités de l'arc. Donc, si il existe dans le graphe orienté, deux arcs de mêmes extrémités et d'orientations opposées, ils doivent avoir nécessairement la même valeur. Deux solutions simples peuvent alors être mises en place : il suffit, soit de compléter chaque arc par son arc associé d'orientation opposée avec la même valeur, soit de supprimer les arcs qui dans le graphe ne sont pas associés à un arc opposé.

Mais au-delà d'une simple question de manipulation des graphes, il est important d'étudier l'influence de ces modifications sur la modélisation du signal. Pour cela, nous proposons quelques cas de figures, ou plus précisément trois cas ont été identifiés.

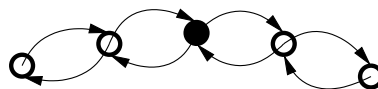


FIGURE 2.46 – Configuration de graphe orienté (2 plus proches voisins) d'un signal 1D, pour l'échantillon • dans une zone uniforme.

Dans le cas d'élément au milieu d'une zone uniforme de l'image, la symétrisation ne change pas le graphe orienté. La figure 2.46 illustre ce cas où l'on constate que le graphe (formé à partir des 2 plus

proches voisins) est déjà un graphe orienté symétrique et donc ne change pas. Ce cas de figure est fréquent dans le cadre des images naturelles puisqu'elles sont constituées en grande partie de zones uniformes.

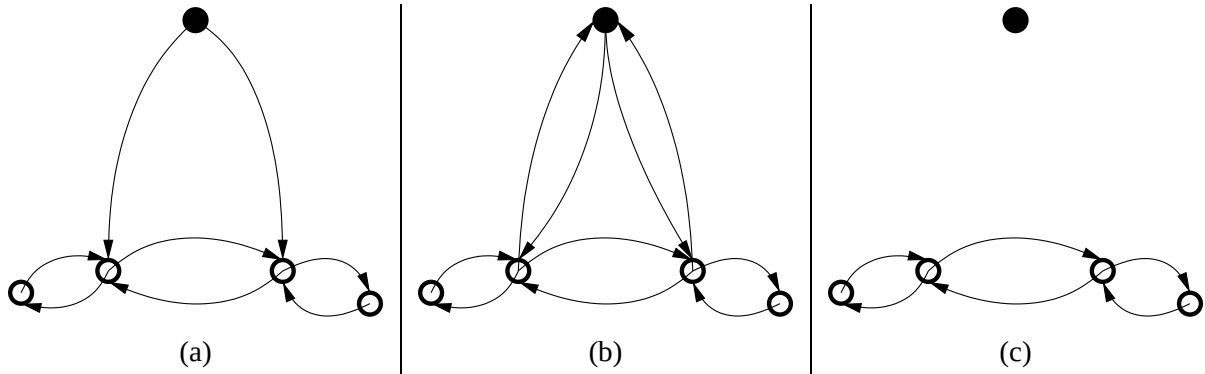


FIGURE 2.47 – Configuration de graphe orienté (a) pour l'échantillon $\bullet$ isolé dans une zone uniforme. La symétrisation (b) par ajout d'arcs ne modifie pas les composantes connexes du graphe alors que la symétrisation (c) par suppression d'arcs va isoler l'élément.

Dans le cadre d'éléments isolés (figure 2.47), nous constatons que l'utilisation de la symétrisation apporte des réponses différentes suivant que l'on utilise l'ajout ou la suppression d'arcs. Dans le cadre de la symétrisation par ajout, un élément isolé se retrouve associé aux éléments de la zone uniforme. Dans le cadre de la symétrisation par suppression d'arcs ce même élément se retrouve isolé donc associé à une distance géodésique infinie. Cette configuration se rencontre en général dans des zones bruitées de l'image (zones où tous les éléments sont différents).

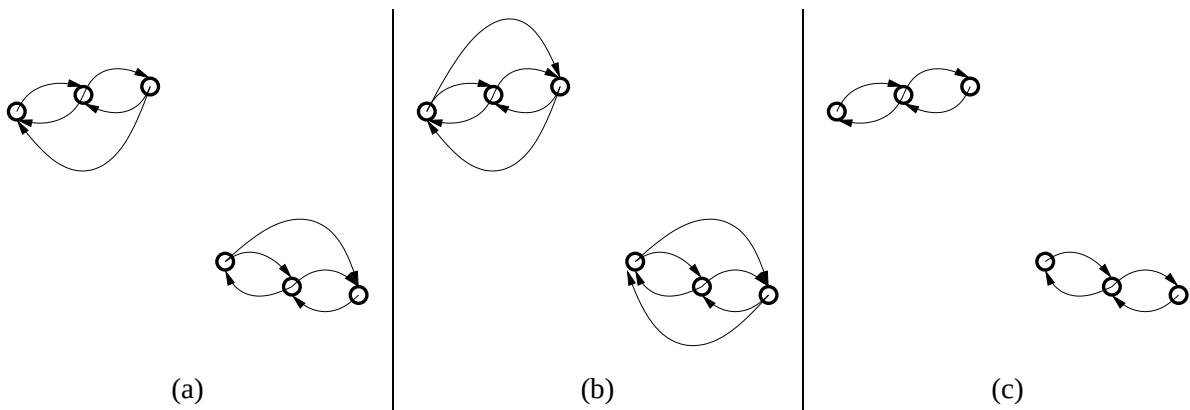


FIGURE 2.48 – Configuration de graphe orienté (a) pour deux ensembles d'échantillons dissociés. Dans les deux cas (ajout (b) ou suppression (c) d'arcs) la symétrisation change la configuration du graphe, mais ne change pas de manière significative les distances impliquées (pas d'apparition de distances infinies supplémentaires).

La figure 2.48 représente un dernier cas où deux zones uniformes se retrouvent isolées par l'application de la construction du graphe de 2 plus proches voisins. Bien évidemment, la symétrisation, qu'elle soit par ajout (figure 2.48(b)) ou par suppression (figure 2.48(c)) changera effectivement l'ensemble de départ. Pour autant, comme dans le cadre des zones uniformes il n'y a pas de changement des composantes connexes du graphe. Ce cas de figure correspond aux ruptures (contours) d'une image.

Pour conclure, nous constatons que seul le cas de l'échantillon isolé (figure 2.47) a un comportement différent selon la méthode choisie pour la construction du graphe non orienté.

Ce changement de connexité amène la construction de nouvelles composantes connexes. Illustrée sur un élément isolé, cette déconnexion peut également s'opérer sur des sous-graphes plus complexes à partir du moment où la connexité de ce sous-graphe dépend uniquement d'arcs qui n'ont pas d'équivalent dans la direction opposée lors de la suppression par symétrisation. Cette apparition de composantes connexes amène l'apparition de distances infinies entre les éléments de composantes différentes. Cette distance infinie est la conséquence la plus marquante du comportement très particulier des distances géodésiques par rapport à la distance euclidienne : les distances géodésiques ne dépendent pas seulement des deux éléments impliqués dans la distance mais bien de l'ensemble des éléments de la variété.

Notons que ce choix entre l'ajout ou la suppression d'arcs dépend de l'application ; selon que nous voulons profiter de la création de composantes connexes amplifiant l'isolement de certains échantillons ou plutôt que nous souhaitons conserver les composantes connexes initiales.

Nous avons opté pour la symétrisation du graphe par suppression d'arcs sans équivalent dans la direction opposée car nous souhaitons profiter de la remise en cause de la connexité pour mettre en valeur les ruptures de l'image. Ce choix n'a pas été remis en cause depuis.

Nous avons donc à partir de cette étape un graphe orienté, symétrique, valué (en respectant la symétrie arc par arc), strict. Il n'est plus k -régulier puisque nous supprimons les arcs sans prendre en compte la conservation d'un nombre fixe d'arcs par sommet.

Il nous faut maintenant construire le plus court chemin en s'appuyant sur la valeur des arcs et ainsi obtenir notre distance géodésique.

2.3.3.4 Recherche du plus court chemin et extraction de la distance géodésique

Nous avons défini un graphe associé à l'identification d'une variété correspondant au signal en cours de traitement. Nous nous sommes assurés que ce graphe vérifie les conditions voulues (en particulier la symétrie). Le calcul de la distance géodésique se traduit maintenant dans les faits par la recherche du plus court chemin et nous utilisons l'algorithme classique de E. Dijkstra [66].

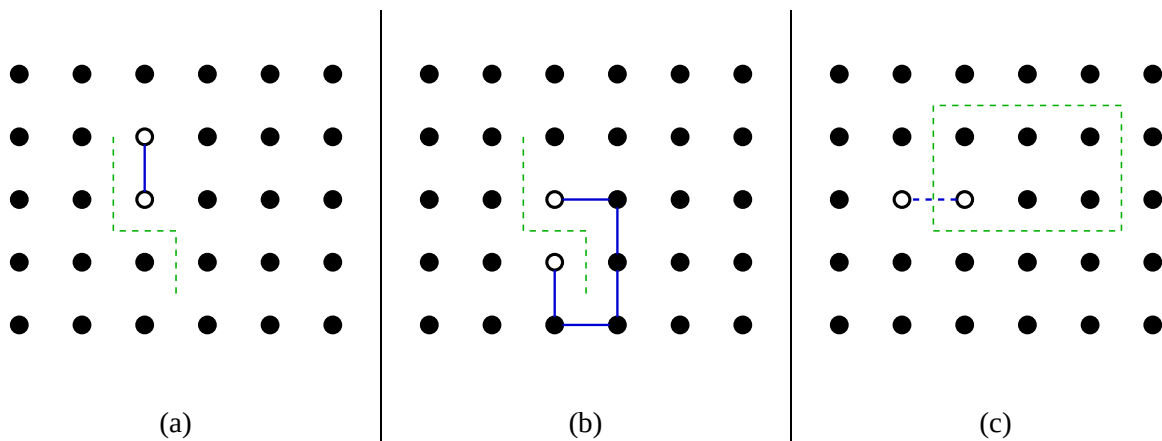


FIGURE 2.49 – Les trois configurations d'arcs géodésiques. a Zone uniforme. b Présence d'une singularité. c Présence de deux composantes connexes. Le trait pointillé vert correspond à une singularité.

Cette recherche concrétise la structuration de l'image à travers la distance géodésique. La figure 2.49

illustre les trois configurations attendues. Dans le premier cas de figure, nous sommes en présence d'une zone uniforme : la longueur du chemin correspond directement à la distance selon la grille entre les éléments. Dans le deuxième cas, le chemin doit contourner une rupture pour relier les deux éléments étudiés et donc la longueur de l'arc est grandement amplifiée. Enfin dans le dernier cas, les deux points font parti de deux composantes connexes et par conséquent, les deux éléments ne peuvent être liés par un quelconque chemin. Dans ce dernier cas la longueur du chemin sera considérée infinie.

Nous proposons maintenant quelques illustrations du comportement de la distance géodésique. Pour cela, nous construisons une représentation des distances de tous les pixels d'une image à son pixel central. Pour représenter cette distance nous quantifions les valeurs des distances en les ramenant modulo à l'intervalle $[0, 255]$.

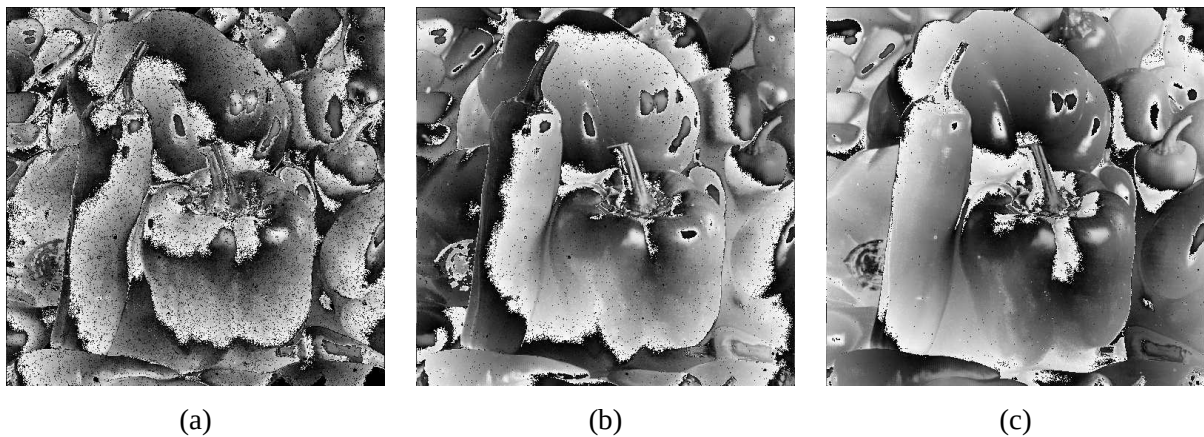


FIGURE 2.50 – Exemple des distances géodésiques des pixels au centre de l'image Peppers. Les distances sont obtenues à partir d'une topologie correspondant à un graphe k plus proches voisins. (a) Pour $k = 4$. (b) Pour $k = 8$. (c) Pour $k = 16$.

Si nous reprenons l'analyse de Cohen [54], nous pouvons assimiler les plus courts chemins à la propagation de la lumière dans un milieu non uniforme. Les ensembles de pixels associés à une même valeur sont donc les fronts d'ondes de cette propagation : le passage de 255 à 0, conséquence de l'utilisation de l'opérateur modulo, est la version la plus visible de ce front d'ondes. Nos images sont des milieux non homogènes de propagation d'une onde.

Nous voyons sur la figure 2.50 que ce front d'ondes s'avère de plus en plus régulier, et son déplacement de plus en plus "rapide" si l'on augmente le nombre de plus proches voisins utilisé dans l'extraction de la topologie : plus k augmente, plus le front d'ondes (passage blanc-noir) se propage rapidement du centre vers l'extérieur de l'image. De plus, nous constatons que :

- les zones texturées sont associées à des fronts d'ondes beaucoup moins réguliers que dans les zones uniformes,
- les frontières de l'image produisent des phénomènes de "diffraction" qui entraînent un retard sur le front d'ondes lorsqu'il passe d'un pixel à son voisin (voisin au sens de la grille) situé de l'autre côté de la frontière.
- enfin plus le nombre de voisins est important, moins les distances géodésiques tiennent compte des singularités de l'image.

A ce stade, nous avons une méthodologie permettant d'extraire d'un signal défini sur une grille régulière

une topologie adaptée au signal vu en tant que variété, et à travers cette organisation une distance géodésique. Nous avons donc tous les éléments pour mettre en oeuvre notre filtrage adaptatif.

2.3.4 Filtrage bilatéral géodésique

Nous proposons un filtrage bilatéral géodésique comme illustration de l'intégration de la distance géodésique dans un processus de filtrage.

Nous rappelons que notre filtrage bilatéral s'écrit :

$$C_v(\mathbf{x}) = \frac{1}{n_v(\mathbf{x})} \int_{\Omega} S(\mathbf{t}) g(D(S(\mathbf{t}), S(\mathbf{x})), \sigma_v) d\mathbf{t}$$

avec

$$n_v(\mathbf{x}) = \int_{\Omega} g(D(S(\mathbf{t}), S(\mathbf{x})), \sigma_v) d\mathbf{t}$$

avec $D(S(\mathbf{t}), S(\mathbf{x}))$ la distance géodésique calculée suivant la méthode présentée dans le paragraphe précédent.

Nous proposons d'appliquer ce nouveau filtre sur l'image utilisée en exemple par Tomasi and Manduchi [240]. Nous appliquons à titre de comparaison le filtrage bilatéral tel qu'il est défini à l'origine. Les filtres sont itérés 10 fois afin d'obtenir une version régularisée de l'image. L'écart-type du noyau de lissage est fixé à 10.

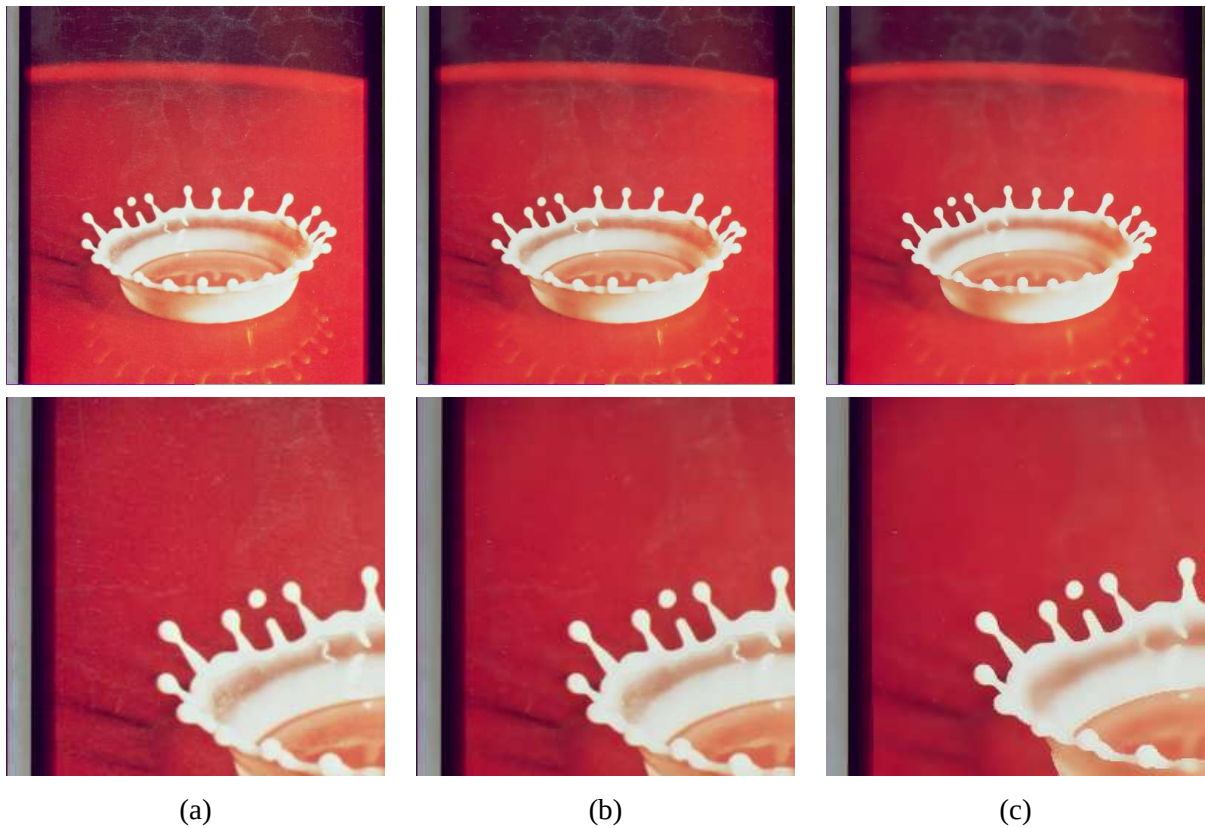


FIGURE 2.51 – (a) Image originale. (b) Application du filtrage bilatéral classique. (c) Application du filtrage bilatéral géodésique.

La figure 2.51 présente les résultats des filtrages. Bien évidemment, le filtrage bilatéral "classique" obtient un résultat tout à fait satisfaisant. Mais nous pouvons constater que la modélisation du signal à travers le concept de variétés permet de lisser l'image en assurant une conservation très nette des ruptures comme par exemple entre le rouge et le blanc (figure 2.51(c)).

Pour compléter cette première illustration concernant l'intégration de la distance géodésique dans un processus de filtrage, nous proposons une expérimentation dans une problématique de restauration de données. Il existe des travaux qui ont étudié le paramétrage du filtrage bilatéral dans le cadre d'image bruitée, par exemple l'article de Zhang et al.[252]. En ce qui concerne le filtrage bilatéral classique, σ_d (paramètre de la pondération selon la position sur la grille) a une faible influence et nous le fixons égal à σ_r pour limiter le nombre de paramètres. Le support du filtre sur la grille est de 3×3 pixels. Enfin, σ_r (paramètre de la pondération selon la valeur des éléments) doit être défini en fonction du niveau de bruit. Dans cette étude, nous proposons d'utiliser $\sigma_d = \sigma_r = 3 \times \sigma_e$ avec σ_e le niveau de bruit.

En ce qui concerne le filtrage géodésique, le compromis est de pouvoir détecter les singularités tout en limitant l'influence du bruit. Nous proposons le paramétrage suivant : $\sigma_v = 9 \times \sigma_e$, avec σ_e le niveau de bruit estimé à chaque itération du filtrage. Pour estimer le niveau de bruit nous avons repris la stratégie classique utilisée dans le cadre des ondelettes, à savoir à partir de la valeur médiane des coefficients d'ondelette de la première échelle [71].

L'expérimentation est menée sur une image corrompue par un bruit blanc Gaussien d'amplitude 30.

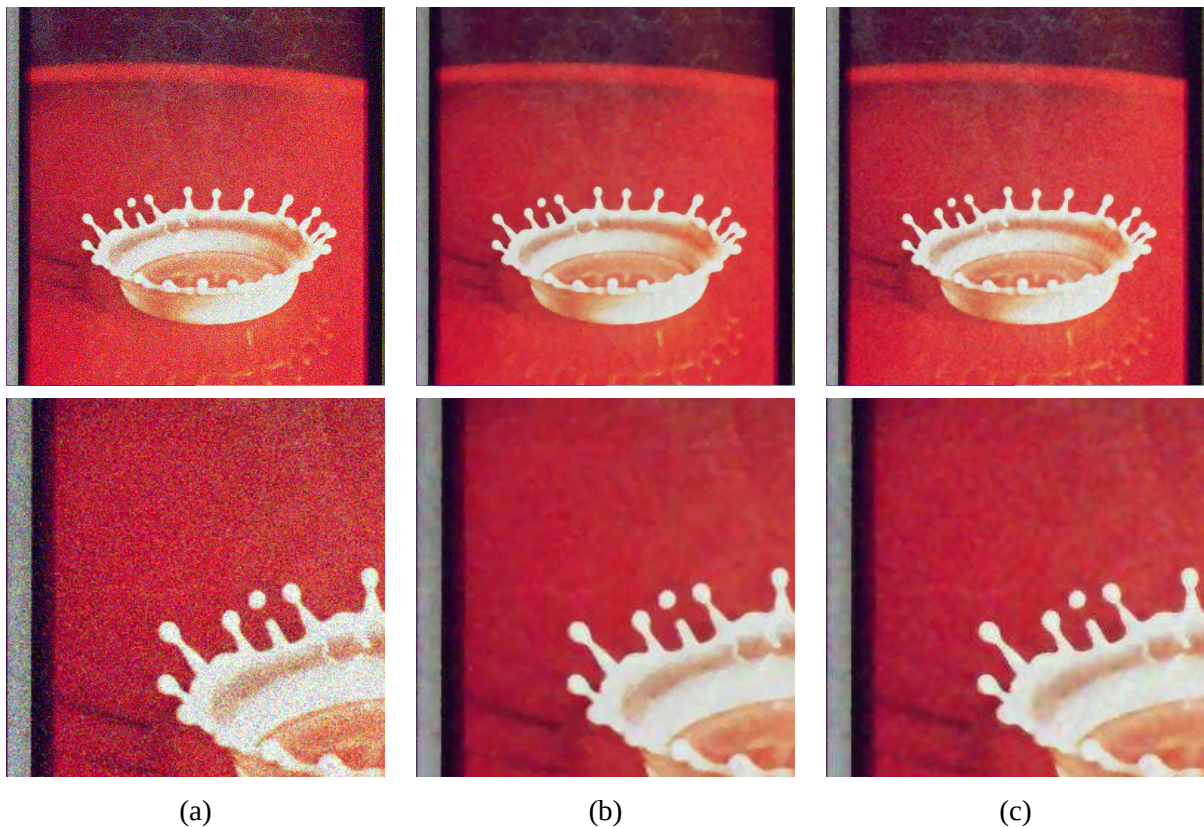


FIGURE 2.52 – (a) Image bruitée (PSNR = 19.08 dB). (b) Image restaurée par le filtrage bilatéral (PSNR = 29.31 dB). (c) Image restaurée par le filtrage bilatéral géodésique (PSNR = 30.46 dB).

La figure 2.52 présente les résultats obtenus suivant les deux filtrages. Les deux méthodes obtiennent

des résultats tout à fait satisfaisants. Cependant, le filtrage géodésique obtient un meilleur résultat en terme de mesures PSNR, et nous pouvons observer que les contours sont très bien conservés ce qui confirme les résultats de la première expérimentation. Toutefois, on peut aussi remarquer l'influence du bruit dans l'image reconstruite, à travers l'apparition de petites taches dans les zones uniformes. Ceci provient dans de nombreux cas de la création d'éléments isolés lors de l'extraction du graphe (caractéristique commentée dans le paragraphe concernant la construction du graphe non orienté). Afin d'améliorer ce travail, nous devons intervenir directement sur le graphe, nous retrouverons ce point dans le cas des ondelettes.

Dans le cadre de ce mémoire, nous ne poursuivrons pas plus les expérimentations concernant le filtrage géodésique car il ne constitue qu'une étape. Cependant d'une manière générale, nous pouvons constater que la modélisation à travers la notion de variété et le calcul d'une distance géodésique associée, peut être une stratégie afin d'intégrer dans un processus de filtrage à la fois l'information géométrique et aussi la dimension couleur. Etant donné notre objectif initial, c'est tout naturellement que nous nous sommes posés la question de l'intégration de cette distance géodésique dans un schéma de décomposition en ondelettes. Mais pour cela, nous avons besoin d'un schéma de calcul plus souple que le banc de filtres. C'est pourquoi nous avons proposé d'utiliser le Lifting Scheme. Nous suggérons en préambule dans la section suivante de faire un bref rappel de ce schéma de calcul et, à travers cette courte présentation, montrer comment une information supplémentaire comme la distance géodésique peut venir s'insérer dans la transformée.

2.3.5 Transformée en ondelettes 2-D géodésiques

Dans cette partie, nous allons définir les aspects se rapportant à la mise en oeuvre de notre schéma de lifting avec l'intégration de la distance géodésique. Tout d'abord, rappelons le principe du schéma Lifting.

2.3.5.1 Lifting scheme

Nous présentons ici une courte introduction au Lifting scheme. Ce paragraphe s'inspire du texte introductif de W. Sweldens [235] créateur de cette nouvelle technique.

Le Lifting scheme est une méthode assez récente permettant de construire des décompositions en ondelettes biorthogonales. La principale différence de cette technique avec l'algorithme classique de décomposition présenté au début de ce mémoire est qu'il n'est pas relié à la transformée de Fourier. Ceci permet alors de construire une seconde génération d'ondelettes qui ne sont pas nécessairement le résultat d'une dilatation et translation d'une fonction "mère". Les bases d'ondelettes associées à des filtres peuvent être quand même "traduites" suivant le formalisme de ces bases de seconde génération.

Cette nouvelle décomposition a pour avantages :

- une implantation associée à un coût de stockage mémoire plus faible ;
- un schéma algorithmique plus "intuitif" (par exemple la reconstruction est un simple changement de signe dans les expressions) ;
- et surtout peut être appliquée lorsque la transformée de Fourier n'est pas applicable.

C'est ce dernier point qui nous intéresse ici. En effet, l'intégration de la distance géodésique n'est pas naturelle dans un schéma ayant une justification basée notamment sur Fourier comme les bancs de filtres.

Etudions la construction de cette nouvelle décomposition à partir d'un simple exemple. Nous ne présentons pour des raisons de simplicité, que l'algorithme 1-D. Nous avons à notre disposition un signal discret s correspondant à l'échantillonnage uniforme d'une fonction continue. Nous notons $c_{0,n} = s[n]$ pour $n \in \mathbb{Z}$ le premier étage de la décomposition.

Supposons que nous voulons stocker l'information présente dans ce signal avec un nombre réduit de coefficients. Nous voulons donc approximer le signal original avec une erreur acceptable. Nous pouvons réduire le nombre de coefficients en supprimant simplement les échantillons d'indices impairs. Nous obtenons alors une nouvelle séquence $c_{1,n}$ définie par :

$$c_{1,n} = c_{0,2n} \text{ pour } n \in \mathbb{Z}.$$

D'un autre coté, nous voulons stocker l'information perdue lors de cette opération pour pouvoir définir une méthode de reconstruction, mais aussi pour pouvoir évaluer la quantité d'information perdue. Nous utilisons des coefficients $d_{l,n}$ qui vont encoder la différence et vont correspondre aux coefficients d'ondelettes. De nombreuses solutions peuvent être proposées pour la définition des coefficients $d_{l,n}$. Le choix le plus simple consiste à stocker dans $d_{l,n}$ les coefficients impairs :

$$d_{1,n} = c_{0,2n+1} \text{ pour } n \in \mathbb{Z}.$$

Cela correspond aux ondelettes Lazy. Toutefois, nous constatons rapidement que ce choix n'est pas forcément judicieux. En effet, sauf cas particulier, l'approximation du signal à partir de la séquence $c_{1,n}$ n'est pas significative et l'erreur stockée dans $d_{1,n}$ est importante. Nous allons mettre alors en place une méthode plus intéressante. La problématique est de retrouver les coefficients impairs $c_{0,2n+1}$ à partir des coefficients $c_{1,n}$. S'il existe une corrélation forte sur un voisinage proche de chaque échantillon, alors une méthode performante pour estimer $c_{0,2n+1}$ consiste à dire qu'un coefficient impair a une valeur proche de la moyenne de ces deux voisins. Puisque l'erreur de cette approximation est stockée dans les coefficients d'ondelettes $d_{1,n}$, ils sont alors définis par :

$$d_{1,n} = c_{0,2n+1} - \frac{1}{2}(c_{0,2n} + c_{0,2n+2}) \quad (2.42)$$

Si le signal est fortement corrélé, alors les coefficients $d_{1,n}$ seront faibles. Dans le cas de signaux linéaires par partie, ces coefficients d'ondelettes seront nuls ainsi la représentation sera plus compacte.

Comme pour la décomposition en ondelettes classique, les coefficients d'ondelettes encodent les détails, les hautes fréquences, alors que les coefficients $c_{1,n}$ encodent l'approximation passe-bas du signal. Toutefois le choix du calcul des coefficients $c_{1,n}$ n'est pas satisfaisant. En effet, nous appliquons un ré-échantillonnage sans pour autant réduire la bande fréquentielle du signal, ce qui a pour conséquence l'apparition d'artefacts dus au phénomène de repliement spectral. Il faut appliquer un lissage avant de ré-échantillonner le signal. Le lissage peut se définir à partir d'une seconde contrainte : la moyenne des approximations à chaque échelle reste la même, à savoir (le coefficient 2 résulte de la réduction du nombre de points par 2).

$$\sum_n c_{0,n} = 2 \sum_n c_{1,n}$$

On constate que cette condition n'est pas respectée avec le choix de $c_{1,n}$ actuel. Définissons alors le coefficient de l'approximation du signal à l par :

$$c_{1,n} = c_{0,2n} + \frac{1}{4}(d_{1,n-1} + d_{1,n}) \quad (2.43)$$

En substituant cette définition, on peut décrire la relation entre l'approximation à l'échelle 1 et l'échelle

0 :

$$\begin{aligned} c_{1,n} &= c_{0,2n} + \frac{1}{4} \left[c_{0,2n-1} - \frac{1}{2} (c_{0,2n-2} + c_{0,2n}) + c_{0,2n+1} - \frac{1}{2} (c_{0,2n} + c_{0,2n+2}) \right] \\ &= -\frac{1}{8} c_{0,2n-2} + \frac{1}{4} c_{0,2n-1} + \frac{3}{4} c_{0,2n} + \frac{1}{4} c_{0,2n+1} - \frac{1}{8} c_{0,2n+2} \end{aligned}$$

Nous constatons tout d'abord que l'approximation à l'échelle 1 résulte d'un filtrage passe-bas de l'échelle 0. De plus, on peut vérifier que :

$$\begin{aligned} \sum_n c_{1,n} &= \sum_n \left[\frac{1}{4} c_{0,2n-1} + \frac{1}{4} c_{0,2n+1} \right] + \sum_n \left[-\frac{1}{8} c_{0,2n-2} + \frac{3}{4} c_{0,2n} - \frac{1}{8} c_{0,2n+2} \right] \\ &\quad \text{Somme sur les coefficients impairs} \quad \text{Somme sur les coefficients pairs} \\ &= \frac{1}{2} \sum_{n'} c_{0,2n'+1} + \frac{1}{2} \sum_{n'} c_{0,2n'} = \frac{1}{2} \sum_n c_{0,n} \end{aligned}$$

Le schéma de lifting exposé ci-dessus correspond à la transformée en ondelettes de Cohen-Daubechies-Feauveau (2, 2) (Sweldens W. [236]) que nous noterons CDF par la suite.

Pour résumer, le calcul d'une décomposition suivant le lifting scheme se décline par :

1. *Division* : On sépare les coefficients $c_{l,n}$ de l'échelle l suivant :

$$\text{pair}_l = c_{l,2n} \text{ et } \text{impair}_l = c_{l,2n+1}$$

2. *Prédiction* : On utilise un opérateur P qui calcule la prédiction des coefficients impairs à partir des coefficients pairs. Les coefficients d'ondelettes stockent alors la différence entre cette approximation et la valeur réelle des coefficients impairs

$$d_{l+1,n} = \text{impair}_l - P(\text{pair}_l)$$

3. *Ajustement* : on fixe que la valeur moyenne de l'approximation à l'échelle l est identique à l'échelle $l+1$. On définit un opérateur U permettant de réaliser cet ajustement :

$$c_{l+1,n} = \text{pair}_l + U(d_{l+1,n})$$

Nous présentons sur la figure 2.53 le schéma de décomposition suivant ce formalisme.

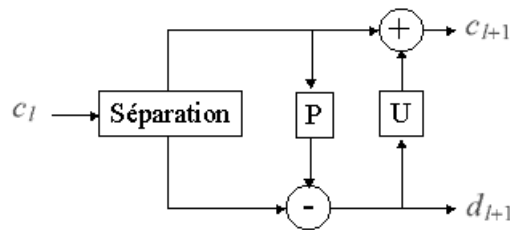


FIGURE 2.53 – Décomposition en ondelettes suivant le schéma Lifting scheme

En poursuivant avec cette définition générale, on constate que la reconstruction va se définir tout simplement par trois étapes :

1. *Opération adjointe de l'ajustement.* En connaissant d_l et c_l , on peut reconstruire les coefficients pairs de l'échelle précédente suivant :

$$\text{pair}_{l-1} = c_l - U(d_l)$$

2. *Opération adjointe de la prédiction.* En connaissant les coefficients pairs et les coefficients d'ondelettes, on peut reconstruire les coefficients impairs suivant :

$$\text{impair}_{l-1} = d_{l,n} + P(\text{pair}_{l-1})$$

3. *Fusion.* Maintenant que nous avons les coefficients pairs et impairs nous pouvons les rassembler afin de reconstruire l'approximation à l'échelle $l - 1$.

$$c_{l-1} = \text{fusion}(\text{pair}_{l-1}, \text{impair}_{l-1})$$

A cet instant discutons de l'opportunité d'utiliser ce schéma. Du fait de sa nature, cette nouvelle décomposition permet de mettre en place plus facilement des approches plus adaptatives, notamment par l'intégration de la distance géodésique.

En effet, si nous reprenons le schéma de lifting associé à la fonction d'ondelettes de Cohen-Daubechies-Feauveau, nous pouvons l'écrire de la manière suivante :

$$\begin{aligned} d_{1,n} &= c_{0,2n+1} - \frac{1}{2}(c_{0,2n} + c_{0,2n+2}) = c_{0,2n+1} - m_{0,2n+1} \\ c_{1,n} &= c_{0,2n} + \frac{1}{4}(d_{1,n-1} + d_{1,n}) \end{aligned} \quad (2.44)$$

Nous constatons que pour calculer le coefficient d'ondelette, l'échantillon $c_{0,2n+1}$ est comparé à un "échantillon virtuel", noté $m_{0,2n+1}$ correspondant à la moyenne entre l'échantillon de droite et l'échantillon de gauche de l'échantillon en cours de calcul. Cette lecture du schéma de lifting CDF illustre deux aspects intéressants :

- Tout d'abord la généralisation possible du schéma à tous les signaux nD. En effet, l'élément $m_{0,2n+1}$ étant la moyenne des échantillons les plus proches de l'échantillon en cours de calcul, il possède une interprétation immédiate quelle que soit la dimension du signal. Cette transposition devra nécessairement prendre en compte la séparation paire-impair. Par exemple la généralisation à l'image peut se faire à travers le calcul de la moyenne des 1-voisins à travers un treillis en quinconce pour la décimation du signal. Nous verrons tous ces aspects dans la section suivante.
- Le deuxième aspect est l'introduction d'une notion de pondération des éléments voisins. En effet, d'une manière générale, le calcul de l'échelle suivante peut intégrer une somme pondérée des échantillons les plus proches de l'échantillon en cours de calcul, par exemple :

$$m_{l,2n} = p_l(2n, 2n-1)c_{l,2n-1} + p_l(2n, 2n+1)c_{l,2n+1}, \quad (2.45)$$

avec $p_l(.,.)$ des poids dépendant des échantillons traités.

C'est à travers cette lecture du schéma, que les décompositions en ondelettes ont par exemple été généralisées aux maillages [209]. Ces pondérations, dont la somme doit être égale à 1, sont une possibilité pour introduire la distance géodésique. C'est ce que nous proposons dans la suite de cette section. Pour cela nous devons aussi étendre le schéma Lifting à une grille 2-D.

2.3.5.2 Décimation par treillis quinconces

L'idée est d'utiliser un schéma de Lifting qui lors de la prédiction du coefficient à l'échelle suivante utilise les différents voisins du pixel étudiés suivant toutes les directions. Pour cela nous proposons d'utiliser deux treillis quinconces entrelacés, tels qu'ils sont définis par Gouze A. et al. [96]. Ainsi, l'intérêt principal est de permettre une décimation par 2 et non par 4 de l'ensemble des échantillons, offrant ainsi une plus grande précision dans les changements d'échelle (les échelles évoluent suivant un pas $\sqrt{2}$) et également parce que cette décimation ne privilégie pas de directions particulières, si ce n'est par l'intermédiaire de la prise en compte de la géométrie comme dans le cadre de la transformée définie par Chappelier et al. [50]. En effet, le schéma associé s'appuie sur une prédiction ou une mise à jour impliquant tous les voisins de l'échantillon en cours de calcul.

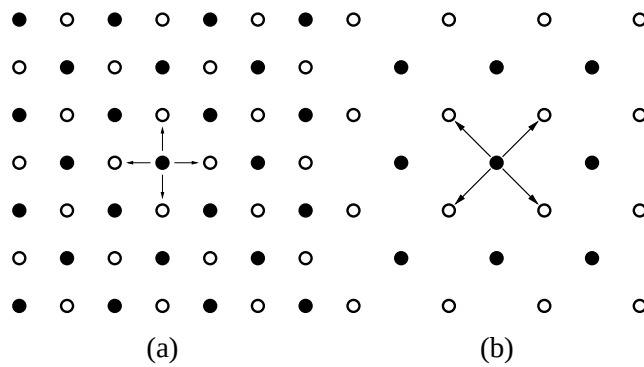


FIGURE 2.54 – Deux treillis quinconces entrelacés d'une image avec (a) les treillis d'une échelle paire et (b) les treillis d'une échelle impaire.

La figure 2.54 illustre deux treillis quinconces successifs. Nous remarquons en particulier que pour un échantillon de l'échelle paire, au centre de la figure, les plus proches voisins sont les 1-voisins définis par les flèches. Nous allons maintenant introduire la distance géodésique dans le passage d'une échelle à l'autre.

2.3.5.3 Introduction de la distance géodésique

Si nous reprenons le schéma 1-D, nous avons introduit une notion de somme pondérée dans la prédiction avec $m_{l,2n} = p_l(2n, 2n-1)c_{l,2n-1} + p_l(2n, 2n+1)c_{l,2n+1}$. Nous pouvons donc redéfinir les pondérations en fonction de la distance géodésique calculée sur les points étudiés :

$$\begin{aligned} p_l(2n, 2n-1) &= \frac{1}{DG_l(2n, 2n-1)} \left(\frac{1}{DG_l(2n, 2n-1)} + \frac{1}{DG_l(2n, 2n+1)} \right)^{-1} \\ p_l(2n, 2n+1) &= \frac{1}{DG_l(2n, 2n+1)} \left(\frac{1}{DG_l(2n, 2n-1)} + \frac{1}{DG_l(2n, 2n+1)} \right)^{-1} \end{aligned} \quad (2.46)$$

avec $DG_l(2n, 2n-1)$ la distance géodésique calculée entre les échantillons de coordonnées $2n$ et $2n-1$, à l'échelle l . La mise en œuvre de ce schéma mène à une transformée en ondelettes adaptatives puisque l'utilisation de cette distance entraîne la construction de fonctions d'analyse spécifiques à chaque position $2n$.

Nous pouvons facilement étendre ce schéma à la dimension 2 en définissant un échantillon virtuel correspondant à la valeur du barycentre des 1-voisins de l'échantillon en cours de calcul :

$$m_{l,e} = \frac{1}{k(e)} \sum_{v \in V(e)} \frac{c_v}{DG_l(e, v)} \quad (2.47)$$

avec e les coordonnées du pixel étudié, et $V(e)$ l'ensemble de vecteurs de coordonnées des 1-voisins du pixel étudié et défini par $V(e) = \{e + (1, 0), e + (0, 1), e + (-1, 0), e + (0, -1)\}$

Le dénominateur assure toujours la normalisation :

$$k(e) = \sum_{v \in V(e)} \frac{1}{DG_l(e, v)} \quad (2.48)$$

Nous obtenons un lissage et donc un changement d'échelle qui va dépendre de la structuration locale du signal étudié puisque la pondération dépend de la distance géodésique entre le pixel étudié et ses voisins $DG_l(e, v)$. Nous avons vu que dans le cadre de la distance géodésique, l'émergence de composantes connexes conduit à une distance infinie. Dans le cadre de notre schéma de calcul cela entraîne alors une absence de prise en compte de l'échantillon associé à cette distance dans l'étape de prédiction ou de mise à jour. Nous avons ainsi défini un schéma de lifting nous permettant de construire une transformée en ondelettes géométriques adaptatives à échantillonnage critique.

Notons de plus que la gestion du support borné de l'image à partir de ce schéma, se fait sans périodisation, ni symétrisation du signal. En effet, le mécanisme d'exclusion des échantillons n'appartenant pas aux mêmes composantes connexes est également le mécanisme de prise en charge du support de l'image. Tout "voisin" de l'échantillon en cours de calcul, qui se situerait en dehors du support de l'image, se verrait attribuer une distance géodésique infinie.

Il reste actuellement un dernier aspect concernant cette transformée qui constitue une question ouverte : nous devons pour chaque échelle, reconstruire le graphe des k plus proches voisins et mesurer les différentes distances géodésiques utiles à la transformée. En effet, la transformée en ondelettes, même adaptative, ne peut garantir la conservation de la topologie au fur et à mesure du changement d'échelle. La topologie identifiée à une échelle est difficile à transférer directement à l'échelle supérieure. Nous recalculons donc notre topologie à chaque échelle. La réflexion actuelle porte sur la construction d'une topologie hiérarchique. Nous étudions par exemple les liens possibles avec les techniques proposant la construction d'un maillage de plus en plus grossier. La difficulté est qu'en général ces processus se déroulent dans un sens inverse, à savoir que l'on part d'un maillage grossier que l'on raffine petit à petit.

Nous avons ainsi défini et discuté de l'ensemble des choix mis en œuvre dans notre transformée appelée transformée en ondelettes géodésiques, le but étant de servir au mieux l'objectif d'une transformée en ondelettes géométriques adaptatives. Nous proposons maintenant d'illustrer le comportement de la transformée.

2.3.6 Analyse de la transformée géodésique

Dans le cadre de la thèse de G. Lebrun [132], différentes expérimentations ont été réalisées permettant d'illustrer le comportement de cette nouvelle transformée. La première d'entre-elles correspond à une problématique de reconstruction partielle (analyse du signal reconstruit à partir simplement de quelques coefficients). Le nombre de plus proches voisins utilisés pour redéfinir la topologie de l'image a été fixé à quatre, comme nous l'avons défini dans l'équation (2.47).

Pour tester l'application de reconstruction partielle, nous trions l'ensemble des coefficients d'ondelette, toutes échelles confondues, en fonction de leur énergie, puis nous ne gardons que les n premiers coefficients (coefficients de plus grande énergie) en fonction d'un pourcentage fourni en paramètre.

Nous avons choisi d'illustrer le comportement de la transformée géodésique avec deux images bien différentes : l'image *Splash* qui est très peu texturée et l'image *Mandrill* qui est fortement texturée. De plus nous comparons les résultats de la représentation géodésique avec une transformée en ondelettes classique offrant souvent une excellente référence en termes de concentration de l'information.

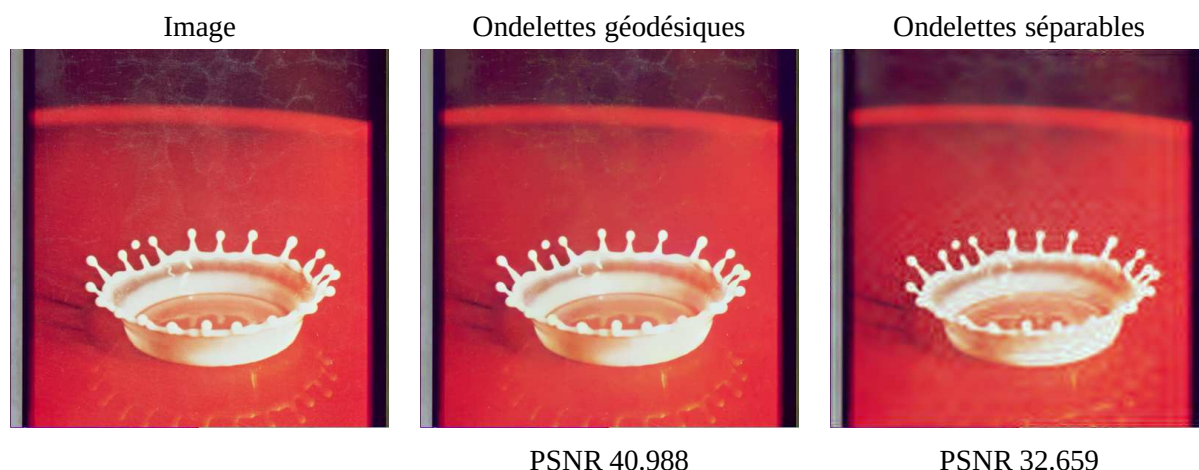


FIGURE 2.55 – Reconstruction partielle de l'image *Splash* par ondelettes géodésiques et ondelettes séparables avec 1% des coefficients d'ondelettes. Le PSNR correspondant est indiqué sous chaque reconstruction.

La reconstruction partielle de l'image *Splash* (figure 2.55) indique de très bons résultats (une amélioration de 8 dB suivant la mesure PSNR par rapport aux ondelettes classiques) et montre l'intérêt que peut offrir l'utilisation d'une distance géodésique pour des images très structurées géométriquement (présentant des objets et donc des zones uniformes avec des ruptures importantes).

Nous remarquons, notamment dans l'image reconstruite après transformée géodésique, l'absence d'effet de pseudo-Gibbs, et la conservation de certains détails comme le reflet (en bas à droite dans l'image) que les ondelettes classiques ont complètement supprimé. Toutefois, pour que la comparaison soit complètement honnête, il faut préciser que nous raisonnons ici en nombre de "coefficients" conservés et non en taille d'information. En effet, pour chaque coefficient de la transformée géodésique, l'information stockée est plus riche, car chaque pixel est associé à quatre distances. Pour envisager des applications de type codage, nous devons nous appuyer sur une géométrie moins précise. Il serait intéressant d'étudier à l'avenir des moyens de résumer la géométrie pour n'en garder que l'essentiel. Cela pourra se faire par exemple en identifiant les pixels de zones uniformes, ou d'autres modélisations de la géométrie de l'image qui pourraient mener à une distance géodésique moins précise mais suffisante pour définir des pondérations correctes dans notre schéma de calcul.

Le deuxième test de reconstruction partielle présenté sur la figure 2.56 illustre certaines limites attendues dans le cadre de ce type d'approches géométriques. Du fait de la présence de texture, la suppression de 99% des coefficients d'ondelettes ne permet plus la reconstruction de toutes les "zones uniformes" de l'image. Les zones de texture vont correspondre lors de l'extraction de la topologie à un grand nombre d'éléments

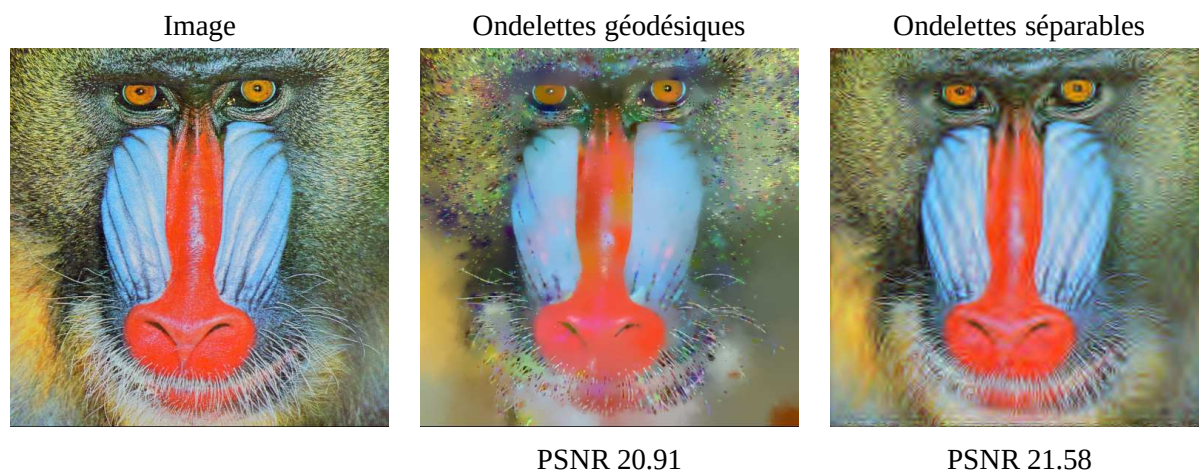


FIGURE 2.56 – Reconstruction partielle de l’image Mandrill par ondelettes géodésiques et ondelettes séparables avec 1% des coefficients d’ondelettes. Le PSNR correspondant est indiqué sous chaque reconstruction.

isolés. Si ces coefficients disparaissent, la zone est partiellement faussée. Dans le cadre des ondelettes, les premiers travaux discutant de cet aspect concernaient les maxima d’ondelettes qui n’étaient pas, de la même façon, adaptés pour le codage de texture. La solution proposée consistait dans un cadre de codage à assurer l’extraction de la géométrie à travers les maxima d’ondelettes et les textures étaient ensuite codées par une approche séparable orthogonale classique.

Dans le cadre de notre application, l’évolution proposée est la suivante :

- modifier la géométrie pour qu’elle ignore les composantes connexes qui ne font que quelques pixels ;
- puis à partir de l’image de résidu issue de ce codage, sélectionner la texture restante par simple seuillage d’une approche séparable orthogonale classique.

Nous proposons une dernière manipulation afin de comprendre le comportement de cette transformée. Nous allons tout simplement effectuer la transformée en ondelettes géodésiques sur deux images de même taille et reconstruire ces images en intervertissant les géométries. Afin de mettre en valeur l’influence de la géométrie, nous utilisons une image très simple de damier (figure 2.57). Elle fera apparaître deux composantes connexes : les cases blanches pour l’une et les cases noires pour l’autre.

Nous avons interverti la géométrie de l’image Splash et de l’image Damier. Nous constatons dans les deux cas que les coefficients d’ondelettes conservent une partie de la structuration de l’image du fait de la position des coefficients. Néanmoins, sans la géométrie, les coefficients ne suffisent plus à définir l’image. Nous voyons par exemple dans l’une des cases du damier (figure 2.57(b)) que si nous prenons la reconstruction de Splash, le contour est dilué.

Nous voyons sur l’image Splash (image 2.57 (c)) que nous avons, avec une géométrie associée aux zones uniformes, complètement perdu la précision de la localisation des contours. De par l’utilisation de la géométrie de l’image Damier, cette localisation des contours est biaisée au point de complètement mélanger les données de part et d’autre du contour.

Enfin, comme pour la transformée Ridgelet, cette représentation a été utilisée dans des applications de restauration d’image. Nous proposons, dans la section suivante, de les illustrer.

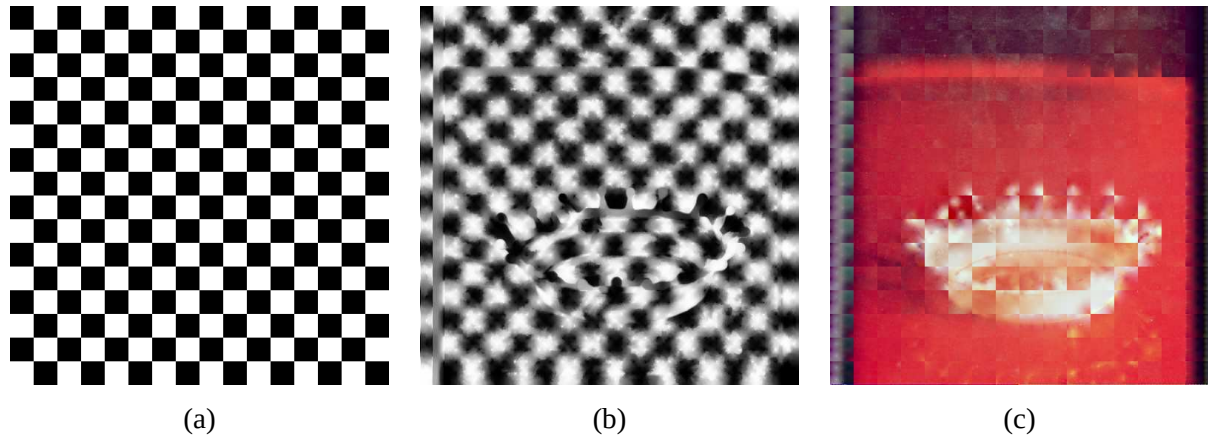


FIGURE 2.57 – Image de Damier et l’image réelle Splash. (b) Utilisation de la géométrie de l’image Splash et des coefficients d’ondelettes de l’image Damier. (c) Utilisation de la géométrie de l’image Damier et des coefficients d’ondelettes de l’image Splash.

2.3.7 Application à la restauration d’image

Nous proposons pour conclure sur cette nouvelle transformée de confronter la transformée géodésique aux applications de restauration d’image. Mais tout d’abord, comme il est bien connu qu’une transformée redondante obtient de meilleurs résultats dans le cadre d’un problème de restauration (résultat que j’avais pu analyser durant mon propre travail de thèse sur les décompositions non-décimées), il nous a semblé pertinent durant la thèse de G. Lebrun de définir un schéma géodésique redondant. Nous le rappelons rapidement dans la section suivante.

2.3.7.1 Introduction de redondance

Nous avons construit une version redondante de notre transformée géodésique. Nous nous inspirons des travaux de Lee et al. [134] dans le cadre du lifting 1D. L’idée est de ré-écrire les étapes de prédiction et de mise à jour pour prendre en compte le fait que les échantillons ne sont plus séparés par l’étape de l’ondelette paresseuse (à savoir séparation paire-impair). Nous devons prendre en compte l’augmentation du support de la fonction analysante en dilatant simplement à chaque changement d’échelle la taille du support sur lequel est calculé le Lifting : la position n d’un coefficient de la fonction devient $2^l n$ avec l l’échelle en cours.

Cette adaptation de la transformée à une version redondante induit que nous travaillons, pour une échelle donnée, sur plusieurs treillis entrelacés correspondant aux 2^l différentes versions translatées de la transformée à une échelle donnée, avec l l’échelle en cours. Nous retrouvons la caractéristique bien connue dans le cadre des décompositions non décimées. Cela signifie que les coefficients des différents treillis ne correspondent pas à la même variété. Nous devons prendre en charge cet aspect notamment lors de la construction de la topologie : un élément d’un treillis donné ne doit pas être le voisin, par le graphe des k plus proches voisins, d’un élément d’un autre treillis.

2.3.7.2 Application de la transformée géodésique au débruitage

La procédure de débruitage que nous avons proposée s'appuie sur la version redondante de la transformée que nous venons de décrire. Afin de définir le seuil, nous suivons la même approche que pour la transformée en ondelettes séparable, à savoir un seuil qui s'appuie sur l'écart-type estimé des coefficients d'ondelettes à la première échelle. Le seuillage que nous avons choisi dans le cadre de cette application est un seuillage dur.

Les résultats de la transformée géodésique sont comparés avec ceux d'une transformée en ondelettes non décimée pour avoir un critère pertinent de comparaison.

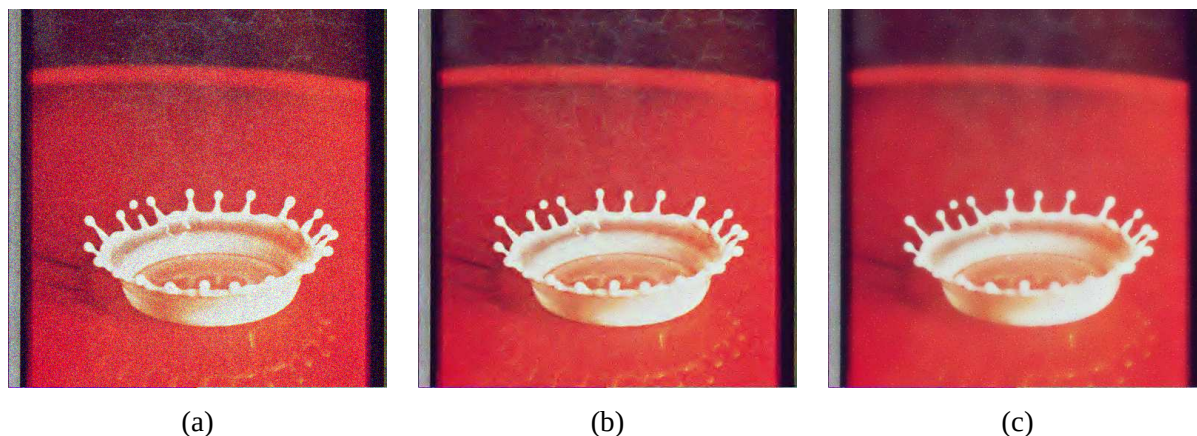


FIGURE 2.58 – Débruitage sur l'image Splash. (a) L'image bruitée. (b) Transformée en ondelettes séparables non décimée PSNR = 26.6779 dB. (c) Transformée en ondelettes géodésiques PSNR = 27.4944 dB.

Nous commentons une première expérimentation avec l'image Splash (figure 2.58). Cette image est dégradée par un bruit blanc Gaussien additif. La conservation des contours est meilleure dans le cadre de la transformée géodésique si l'on considère l'absence totale d'effets de Gibbs (frontière blanc/rouge de la goutte pas toujours nette pour le résultat avec la transformée en ondelettes non décimée). Toutefois, nous constatons que la transformée en ondelettes géodésiques fait apparaître des points sur les zones uniformes un peu à la manière d'un bruit poivre et sel. Nous poursuivons l'expérimentation avec l'image Mandrill qui est bruitée de la même manière (figure 2.59).

De nouveau, les contours sont mieux conservés avec la transformée géodésique : la frontière très marquée délimitant le nez ne souffre pas de l'effet de Gibbs). Cependant nous constatons sur l'image Mandrill reconstruite le bruit poivre et sel dans le cadre du débruitage par ondelettes géodésiques.

Ces premières expérimentations nous montrent que le seuil utilisé n'est pas adapté. En effet, nous conservons sans aucune difficulté les ruptures, mais du fait d'un seuil trop bas, certaines petites composantes connexes liées au bruit et présentes dans les zones uniformes sont conservées, provoquant l'apparition d'éléments parasites particuliers. Nous proposons d'augmenter le seuil afin de supprimer ces éléments parasites.

Nous avons augmenté le seuil (empiriquement d'un facteur 1.2) en considérant que la conservation satisfaisante des contours nous permet de supprimer un peu plus de coefficients. Nous pouvons voir sur le résultat obtenu sur l'image Splash (figure 2.60) que nous avons une amélioration très nette de la qualité visuelle puisque nous limitons très fortement les artefacts de reconstruction tout en conservant parfaitement les contours et en supprimant le bruit.

Le résultat obtenu sur l'image Mandrill (figure 2.61) avec le même seuil augmenté va lui aussi vérifier

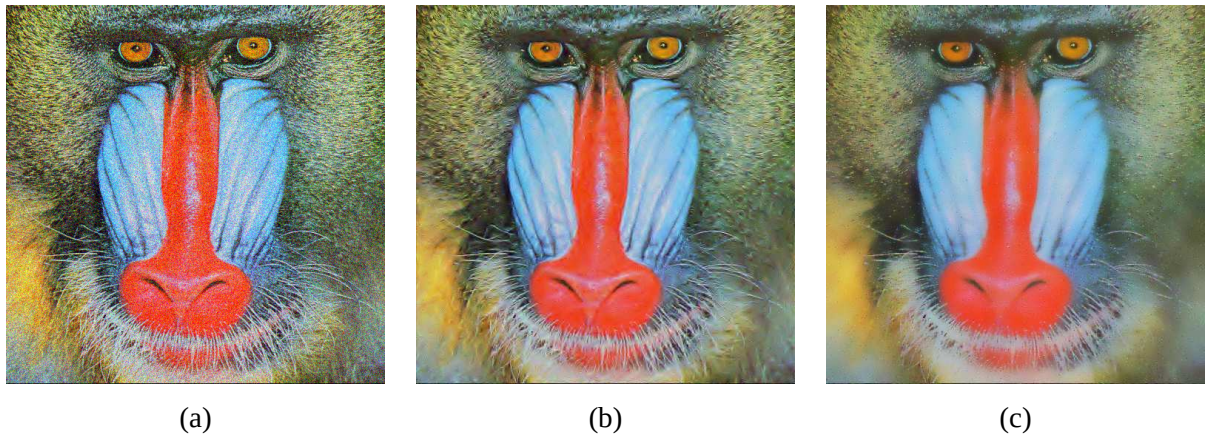


FIGURE 2.59 – Débruitage sur l'image Mandrill. (a) L'image bruitée. (b) Transformée en ondelettes séparables non décimée PSNR = 22.5679 dB. (c) Transformée en ondelettes géodésiques PSNR= 21.3662 dB.

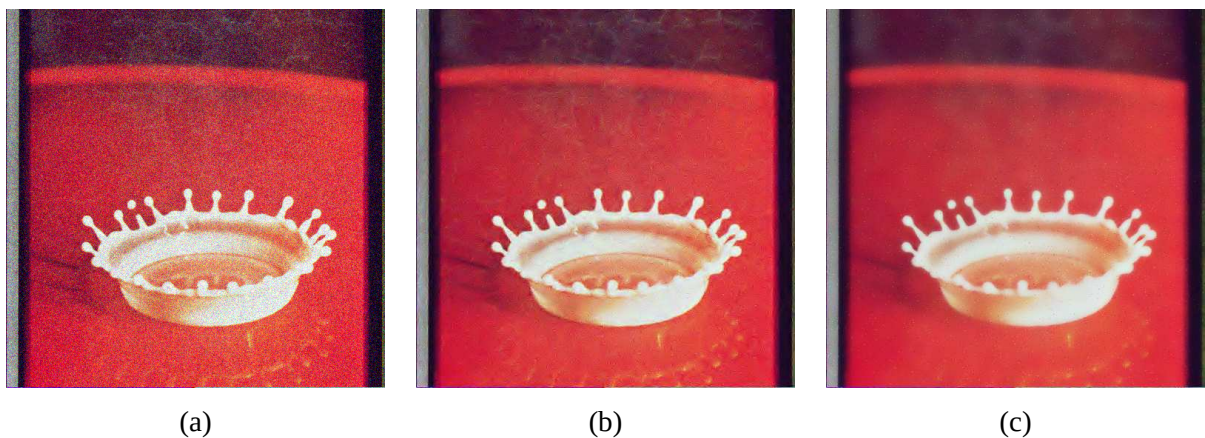


FIGURE 2.60 – Débruitage avec augmentation de seuil sur l'image Splash. (a) L'image bruitée. (b) Transformée en ondelettes séparables non décimée PSNR = 26.6779 dB. (c) Transformée en ondelettes géodésiques PSNR = 28.3384 dB.

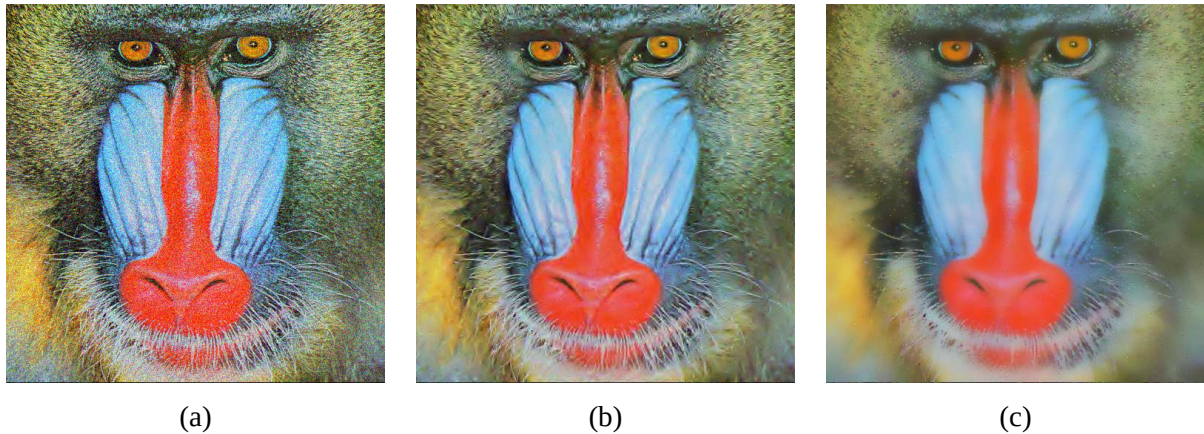


FIGURE 2.61 – Débruitage avec augmentation de seuil sur l'image Mandrill. (a) L'image bruitée. (b) Transformée en ondelettes séparables non décimée PSNR = 22.5679 dB. (c) Transformée en ondelettes géodésiques PSNR = 20.773 dB.

ces caractéristiques similaires : la restauration permet bien de supprimer le bruit et conserve les contours. Ceci étant, nous retrouvons les limites soulevées lors de la reconstruction partielle : il y a, lors du seuillage, une confusion dans notre domaine de représentation entre le bruit et la texture, et le résultat obtenu sans augmentation du seuil est meilleur si l'on privilégie la conservation de la texture.

Ces résultats valident ce que nous attendions de cette nouvelle représentation. A savoir que les informations géométriques comme les ruptures sont parfaitement représentées dans le domaine transformée et sont donc conservées lors du seuillage. En revanche, il est tout à fait logique qu'une approche géométrique soit désavantagée dans le codage de la texture.

De plus ces résultats posent la question de la stratégie de seuillage dans le cadre de ce type de représentation. Dans la solution décrite ci-dessus, nous effectuons un seuillage seulement sur la valeur des coefficients. Mais puisque nous prôtons une intégration de la géométrie dans la représentation, il faut que cette géométrie intervienne aussi durant la phase de seuillage. Concrètement cela doit se traduire sur un traitement direct sur le graphe utilisé lors de la construction des voisinages. Un travail prospectif est mené actuellement sur ce point, notamment sur l'intégration de structures de représentations plus complexes.

Nous proposons enfin de décrire l'une des particularités de cette transformée qui selon moi constitue son point fort : l'intégration naturelle de la couleur.

2.3.8 Intégration de la couleur

L'intégration dans la transformée de la couleur, et donc son influence, se fait tout naturellement lors de la construction de la topologie et se répercute, à travers la distance géodésique produite, dans l'identification de la géométrie. En effet, durant la construction du graphe des k plus proches voisins, la dimension (3 ou 5 en l'occurrence) de l'ensemble traité ne change pas le principe de la méthode. Nous cherchons alors à extraire une notion de chemin minimal, non plus à partir de distances calculées sur une différence de luminosité mais sur une différence entre des données couleur. Il faut noter d'ailleurs que tout logiquement les techniques de modélisation des variétés sont utilisées par quelques auteurs pour la manipulation des images couleur ou d'images hyperspectrales [56, 14]. Il est bien connu dans la communauté "Image couleur" que l'on peut calculer une analyse en composantes principales sur une image couleur afin d'en extraire les plans

principaux pour, par exemple, une procédure de restauration. Suivant ces nouvelles techniques d'analyse de données, l'algorithme ISOMAP est utilisé pour la conversion d'une image couleur en une image niveaux de gris dans [56]. Cependant, on peut remarquer que cette application directe n'est apparue dans la littérature qu'en 2009, donc le lien que nous exploitons dans le cadre de la transformée géodésique reste récent.

Nous pouvons donc intégrer sans difficulté la dimension couleur dans la transformée. Cette intégration de la dimension couleur dans les approches classiques (notamment basée Fourier) est plus délicate. Cette nouvelle liberté apportée par la transformée va aussi se traduire dans l'espace colorimétrique choisi. En effet les changements d'espace ne sont en général pas linéaires et donc les répercussions sur la construction du graphe des k plus proches voisins ne peuvent se résumer à un simple "facteur" d'influence.

Afin d'illustrer ce point sur les différents espaces et donc l'influence du calcul des distances couleur, nous proposons de représenter la distance géodésique au pixel central pour une même image, pour un même graphe à 8 plus proches voisins, mais dans une version en niveaux de gris, dans l'espace RGB et dans l'espace Lab (pour simuler un passage dans Lab nous avons fait le choix d'un illuminant D65 et des primaires correspondant à $sRGB$).

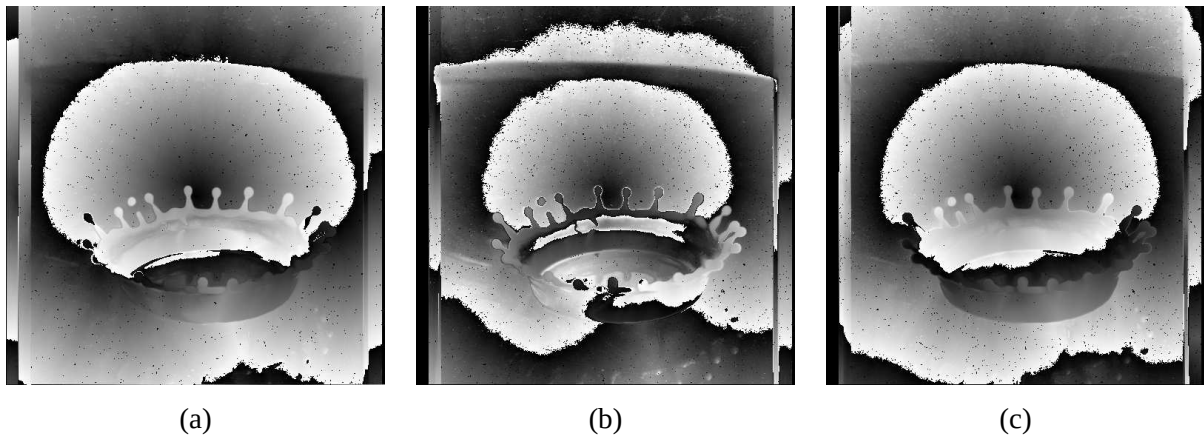


FIGURE 2.62 – Trois cas de calcul de distances géodésiques à l'échantillon central pour l'image Splash (8-ppv) : (a) distance estimée pour l'image niveaux de gris, (b) distance estimée pour l'image dans l'espace RGB , (c) distance estimée pour l'image dans l'espace Lab .

Nous constatons à partir de la figure 2.62 que le changement d'espace de RGB à Lab a une grande influence sur le comportement de la distance géodésique par rapport aux singularités de l'image. En fait, on peut remarquer une similitude des résultats entre l'image traitée dans l'espace Lab et l'image en niveaux de gris. Se pose alors la question de l'évaluation des arcs dans le graphe, puisque ce résultat semble suggérer que dans l'espace Lab , c'est principalement la composante L qui intervient.

Si nous introduisons ce changement d'espace dans la chaîne de restauration d'image, nous voyons avec la figure 2.63 que nous obtenons par ce changement d'espace une augmentation du PSNR de l'image restaurée sans pour autant détériorer le contour de l'image.

Ces dernières expérimentations montrent le potentiel et l'adaptabilité de cette nouvelle transformée, mais elles posent aussi de nombreuses questions. Cet axe me semble très prometteur et se poursuit actuellement notamment à travers une réflexion sur l'intégration de distance couleur perceptuelle dans la transformée.

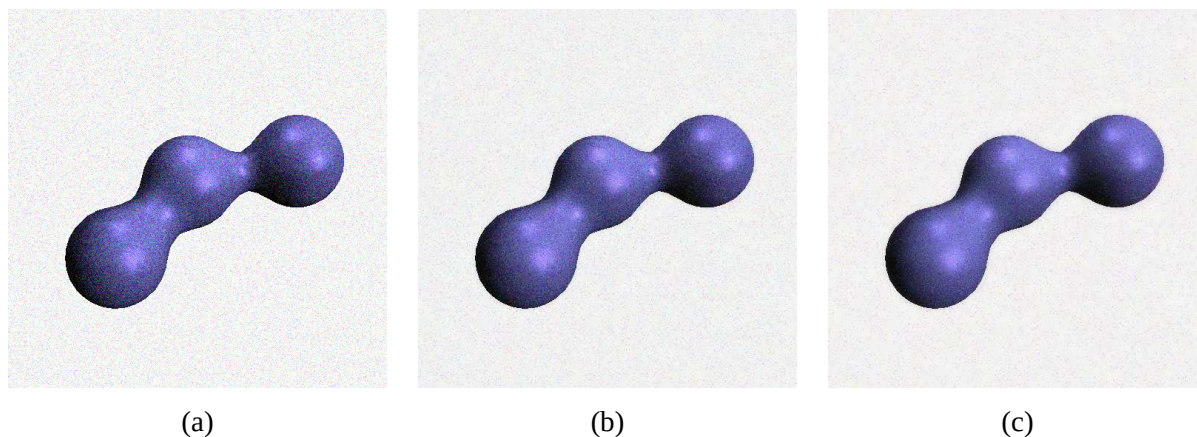


FIGURE 2.63 – Comparaison entre deux applications de débruitage pour des distances géodésiques calculées à partir d'espaces couleur différents sur l'image Metaballs. (a) Image bruitée, (b) cas de l'espace RGB : PSNR = 23.2342 dB, (c) cas de l'espace Lab : PSNR = 23.8248 dB.

Conclusion

Dans ce chapitre, nous avons posé les bases et les problématiques inhérentes à mon travail de recherche, à savoir que nous nous sommes intéressés à la définition d'une représentation multiéchelle qui prend en compte l'aspect géométrique et la dimension vectorielle des données. Pour cela, nous avons dû utiliser des connaissances et des compétences d'autres domaines, notamment de l'informatique graphique.

Dans un premier temps, à travers une collaboration avec une autre équipe du département XLIM-SIC (l'équipe Informatique Graphique et notamment E. Andrès), nous avons construit une nouvelle stratégie pour le calcul de la transformée Ridgelet/Curvelet discrète. L'un des apports a été l'utilisation du formalisme de la géométrie discrète permettant ainsi le déploiement d'une technique rapide, parfaitement reconstituible et qui a tout logiquement obtenu des résultats très intéressants pour le débruitage de données, notamment à travers ses extensions comme la transformée locale ou la version 3-D réalisée dans le cadre du travail de D. Helbet.

Il faut noter que ce premier axe a permis et a structuré un rapprochement thématique avec nos collègues informaticiens et il a donné lieu à différents encadrements et autres travaux annexes (définition d'une transformée en ondelettes invariante par rotation, étude de la rotation discrète dans le plan de Fourier ...) : il illustre bien l'intérêt de la structuration du département à partir de plusieurs domaines de compétences autour de l'image au sens large.

Dans le cadre du second axe, nous avons introduit une transformée en ondelettes géométriques adaptatives s'appuyant sur différents ingrédients, à savoir le schéma de Cohen-Daubechies-Feauveau, le filtrage bilatéral, la distance géodésique d'un signal vue comme une variété et des treillis quinquiconces entrelacés. Ceci nous a permis, à travers la distance géodésique, de prendre en compte la géométrie de l'image. En effet, la distance géodésique permet de répondre à la question qui nous intéresse plus particulièrement : "Existe-t-il une singularité entre le pixel a et le pixel b ?". Ce travail m'a permis d'élargir le champ de compétences des travaux menés et encadrés.

Les résultats notamment dans le cadre de la restauration de données, montrent un comportement de la transformée très intéressant et prometteur qui illustre la manière fine dont est extrait l'information 2-D. Par

rapport à l'existant, l'originalité de cette transformée est qu'elle ne s'adapte pas à la géométrie en utilisant une estimation d'un champ de direction des gradients de l'image mais bien en définissant une forme de segmentation irrégulière de l'image.

Cet axe, complémentaire au premier s'est développé dans le cadre d'une collaboration industrielle avec France Telecom R&D et notamment à travers la thèse de G. Lebrun. Il a, de plus, le grand intérêt d'être à l'intersection de nombreux points scientifiques, comme l'analyse de données non linéaire, la notion de variétés et les travaux sur le flot géométrique et enfin la modélisation sous forme de graphe. Cet axe est d'actualité car il nous reste à approfondir l'étude de l'intégration de la dimension couleur dans la transformée. Comme nous l'avons vu, ce qui est novateur selon nous, c'est que nous n'avons pas une transformée en niveaux de gris que nous cherchons à étendre à des données vectorielles, mais nous définissons directement une transformée adaptée à des données vectorielles. Cependant, il nous faut réfléchir à la modélisation de la couleur (et des disparités colorées). De plus, une comparaison approfondie doit être menée avec les algorithmes de type Moyennes Non-Locales [34] qui se proposent de traiter l'image par Patch et qui intègrent une exploration systématique de toute l'image pour exploiter toute information : il y a une analyse de l'organisation globale de l'image.

Enfin comme nous l'avons vu lors des applications de débruitage, ce type de transformée peut permettre aussi de remettre en question la stratégie de régularisation, à savoir que plutôt qu'un seuillage sur une notion d'amplitude de coefficients, il nous semble maintenant aussi judicieux de travailler sur la modélisation de la topologie traduite par notre graphe. Ce nouveau travail qui s'amorce requiert lui aussi des compétences scientifiques complémentaires

Comme nous l'avons indiqué l'animation de cet axe portant sur les ondelettes géométriques s'est concrétisé à travers le co-encadrement de trois thèses [132, 141, 105], s'est construit à travers l'échange avec différents chercheurs E. Andrès, B. Augereau et D. Helbert de XLIM-SIC, S. Pateux de OrangesLabs. Enfin, étant l'axe le plus ancien, il est celui qui a été le plus valorisé dans littérature à travers principalement [147, 142, 106, 41, 145, 149, 148, 146, 144, 143, 40, 43, 133].

Après avoir traité l'aspect géométrique des données, nous allons dans le chapitre suivant nous pencher sur la dimension couleur.

Chapitre 3

Prise en compte de la couleur dans les outils issus du traitement du signal

Nous proposons maintenant de décrire dans ce chapitre les travaux que nous menons depuis les années 2000, concernant les possibilités qu'offrent certains formalismes mathématiques pour encoder l'information couleur. Le point de départ a été la recherche d'un "outillage" permettant de généraliser différentes opérations du traitement du signal aux données vectorielles. Pour être plus précis, la question peut être : comment définit-on une transformée de Fourier pour une image couleur ? Nous constatons que la stratégie développée dans cet axe est différente et complémentaire du premier chapitre en ce sens que ce n'est pas l'information géométrique qui est recherchée mais bien la description de la fonction contenant des informations vectorielles définies sur une grille discrète, donc une application de $\mathbb{Z} \times \mathbb{Z}$ (ou d'un sous-domaine) à valeurs dans $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ (ou d'un sous-domaine), description à travers une base de fonctions. Les connexions disciplinaires vont être aussi avec les collègues mathématiciens.

Ces travaux ont débuté notamment à travers l'excellent travail de P. Gosselin (mémoire de DEA [95]) puis se sont poursuivis dans le cadre de la thèse de P. Denis, thèse que j'ai encadrée sous la direction de C. Fernandez. L'objectif de ce travail de thèse était notamment de voir les différentes potentialités des Quaternions mais aussi et surtout des algèbres de Clifford pour l'image couleur [62].

Etant donné les collaborations mentionnées ci-dessus avec les collègues mathématiciens, ces travaux ont trouvé tout naturellement un cadre structuré à travers l'action "Géométries et images", dans le cadre de la fédération PRIDES (Programme Régional de recherche en Images, Données et Systèmes), regroupant les laboratoires d'informatiques et de traitement des images des universités de Poitiers, la Rochelle et l'Ensm. L'objectif de cette fédération était la mise en commun des savoirs et l'impulsion de projets collaboratifs en traitement d'images au niveau de la région Poitou-Charentes. Cette action, dont j'étais co-animateur, a donc notamment permis de développer de solides collaborations sur la problématique de l'utilisation des algèbres de Clifford pour l'analyse d'images. Elle va se poursuivre dans le cadre de la nouvelle fédération MIRES (Images, Réseau et Sécurité) qui s'élargit aux équipes de XLIM à Limoges. Je continuerai à assurer un rôle de co-animation.

Les premiers travaux portant sur les transformations numériques des images couleur ont d'abord consisté à appliquer des traitements marginaux sur chacun des trois canaux composant la couleur. Toutefois, puisque ces traitements n'utilisent pas la couleur comme une information globale, ils provoquent l'apparition éventuelle de fausses couleurs. En effet, rien n'indique que le traitement individuel de chaque composante d'un

vecteur couleur a pour résultat un vecteur couleur proche de celui d'origine.

L'une des solutions a alors été de considérer un pixel couleur comme un vecteur de $\mathbb{R}^3$ correspondant à un point au sein d'un espace couleur. Différents espaces ont été proposés afin de représenter la couleur de manière plus ou moins proche de ce que la vision humaine pouvait percevoir des couleurs. La difficulté consiste alors à redéfinir pour la couleur les opérations classiques de traitement des images en niveaux de gris, comme par exemple le produit de convolution, la notion de transformée de Fourier couleur ou plus généralement toutes les approches permettant l'analyse d'une image. Nous avons abordé ce point dans le chapitre précédent.

Dans le cadre de ce nouvel axe, nous avons étudié le codage de l'information couleur des images en utilisant des représentations algébriques. Ce mode de codage, qui n'est étudié que depuis très récemment, permet de manipuler les couleurs en utilisant des opérations algébriques. Dans ce cadre nous nous sommes intéressés tout d'abord à l'utilisation du formalisme des quaternions associé aux images couleur introduit par Sangwine il y a une dizaine d'années [201]. Selon cette approche, les vecteurs associés aux couleurs sont représentés par des quaternions purs, c'est à dire des quaternions dont la partie réelle est nulle. Ensuite, nous détaillerons comment poursuivre cette approche dans le cadre plus général des algèbres de Clifford (dites aussi algèbres géométriques). Mais tout d'abord nous introduisons les définitions associées à cet outillage mathématique.

3.1 Définitions, éléments de base et premières manipulations couleur.

Ce premier paragraphe est consacré à l'introduction des outils de base concernant les quaternions et plus généralement les algèbres de Clifford. Durant cette description des outils mathématiques, nous introduirons les manipulations couleur élémentaires associées à l'outil que nous avons étudiées ou proposées.

3.1.1 L'algèbre des quaternions

Il s'agit de l'algèbre, $\mathbb{H}$, de dimension 4 sur $\mathbb{R}$, constituée des éléments de la forme

$$q = (q_0, q_1, q_2, q_3) = (S(q), V(q)). \quad (3.1)$$

Le réel $S(p)$ et le vecteur $V(p)$ sont communément appelés la partie scalaire et la partie vectorielle du quaternion q . L'algèbre est munie du produit

$$pq = (S(p)S(q) - V(p).V(q), S(p)V(q) + S(q)V(p) + V(p) \times V(q)) \quad (3.2)$$

où $V(p).V(q)$ (resp. $V(p) \times V(q)$) désigne le produit scalaire (resp. mixte ou vectoriel) des deux vecteurs $V(p)$ et $V(q)$ de $\mathbb{R}^3$. L'algèbre $\mathbb{H}$ est associative mais n'est pas commutative. Les éléments de la forme $q = (q_0, q_1, 0, 0)$ forment une sous-algèbre de $\mathbb{H}$ isomorphe à l'algèbre $\mathbb{C}$ des nombres complexes. On retrouve la forme dite de *Hamilton*, $q = q_0 + iq_1 + jq_2 + kq_3$, du quaternion q en posant : $i = (0, 1, 0, 0)$, $j = (0, 0, 1, 0)$ et $k = (0, 0, 0, 1)$.

Les règles de calcul deviennent alors

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j \quad (3.3)$$

Notons que souvent le parallèle est fait avec les nombres complexes en introduisant la notion d'une partie réelle et de trois parties imaginaires. D'ailleurs nous retrouvons un vocabulaire similaire. Le conjugué $\bar{q}$

d'un quaternion q est défini par : $\bar{q} = q_0 - iq_1 - jq_2 - kq_3$. Tout quaternion q , non nul, admet un inverse, nommément $q^{-1} = q/|q|^2$ où $|q|^2 = q\bar{q} = \bar{q}q = q_0^2 + q_1^2 + q_2^2 + q_3^2$. Le nombre réel positif $\sqrt{q\bar{q}} = \sqrt{|q|^2}$ est le *module* de q , noté $|q|$. Compte tenu de la non-commutativité, $\overline{pq}$ est le produit de $\bar{q}$ par $\bar{p}$.

Définition 2 (Ensemble des quaternions purs) *L'ensemble*

$$\mathbb{H}_0 = \{q \in \mathbb{H} \mid q = -\bar{q}\}$$

est appelé ensemble des quaternions purs (i.e. si $q = (S(q), V(q))$ alors $S(q) = 0$).

Définition 3 (Ensemble des quaternions unitaires) *L'ensemble*

$$\mathbb{H}^1 = \{q \in \mathbb{H} \mid |q| = 1\}$$

est appelé ensemble des quaternions unitaires (i.e. si $q = (S(q), V(q))$ alors $S(q)^2 + \|V(q)\|^2 = 1$)

Voyons maintenant les quatre façons de représenter les quaternions.

La représentation vectorielle. C'est la représentation $q = (S(q), V(q))$ introduite au début du paragraphe. Elle est surtout utilisée en physique et en infographie, car elle permet des calculs simplifiés sur les vecteurs de $\mathbb{R}^3$. On peut également considérer un quaternion q comme étant la somme de sa partie scalaire et vectorielle, appelées aussi partie scalaire et partie imaginaire :

$$q = S(q) + V(q). \quad (3.4)$$

On retrouve alors toutes les opérations géométriques usuelles de $\mathbb{R}^3$ lorsqu'on restreint les opérations des quaternions à l'ensemble $\mathbb{H}_0$ des quaternions purs. Soit q_1 et q_2 deux quaternions purs représentant respectivement les vecteurs V_1 et V_2 de $\mathbb{R}^3$, alors :

1. $q_1 + q_2 = V_1 + V_2$
2. $q_1 \cdot q_2 = V_1 \cdot V_2$
3. $q_1 q_2 = -V_1 \cdot V_2 + V_1 \wedge V_2$.

On remarque que le produit quaternionique de deux quaternions purs contient dans sa partie réelle le produit scalaire des deux vecteurs représentés et dans sa partie imaginaire le produit vectoriel (qui coïncide dans $\mathbb{R}^3$ avec le produit mixte). Ce résultat est présent dans certaines utilisations des quaternions pour les images couleur, comme nous le verrons dans la suite de ce chapitre.

La représentation de Hamilton. Elle s'écrit

$$q = q_0 + iq_1 + jq_2 + kq_3 \quad (3.5)$$

avec les règles de multiplication décrite précédemment. Cette écriture fait clairement apparaître que les quaternions sont une généralisation des nombres complexes et que le produit n'est pas commutatif.

La représentation exponentielle. On dispose par ailleurs d'une mise sous forme exponentielle : si q est un quaternion de module 1 ($q \in \mathbb{H}^1$), alors

$$q = \cos \theta + \mu \sin \theta \quad (3.6)$$

avec μ un quaternion pur de norme 1. On peut ainsi faire l'extension de l'exponentielle complexe de façon naturelle :

$$\mu^\theta = \cos \theta + \mu \sin \theta \quad (3.7)$$

toujours avec μ un quaternion pur de norme 1.

Notons que la propriété usuelle $e^\mu e^\nu = e^{\mu+\nu}$ n'est pas vérifiée, et $e^\mu e^\nu \neq e^\nu e^\mu$ en général. En revanche, soient $(\phi, \psi) \in \mathbb{R}^2$ et $(\mu, \nu) \in (\mathbb{H}^1 \cap \mathbb{H}_0)^2$ alors $e^{\mu\phi} e^{\nu\psi} = e^{\mu(\phi+\psi)}$. Par ailleurs si $q \in \mathbb{H}^1$ est un quaternion unitaire, on vérifie facilement que

$$q = e^{\mu\phi} \text{ avec } \mu = \frac{V(q)}{|V(q)|} \text{ et } \phi = \arctan \frac{|V(q)|}{S(q)}. \quad (3.8)$$

La forme polaire. Enfin, au même titre qu'il existe une représentation polaire des complexes du type (amplitude, phase), il existe des représentations similaires des quaternions du type (amplitude, phase, phase, phase). L'écriture suivante a été proposée par T. Bülow [35] dans le cadre de la manipulation des images en niveaux de gris ; les phases correspondent aux angles d'Euler dans un espace 4-D. Tout quaternion peut être représenté de la façon suivante :

$$q = |q| e^{i\phi} e^{k\psi} e^{j\theta} \text{ avec } (\phi, \theta, \psi) \in [-\pi, \pi[\times [-\pi/2, \pi/2[\times [-\pi/4, \pi/4] \quad (3.9)$$

L'obtention de ces différents paramètres est rappelée dans la thèse de Bülow [35]. Cette représentation permet de voir un quaternion q comme un vecteur de longueur $|q|$ et de direction (ϕ, θ, ψ) .

Nous allons décrire maintenant comment les quaternions permettent de réaliser des transformations couleur.

3.1.2 Manipuler des couleurs à l'aide des Quaternions

A travers cette section nous, allons présenter les différentes manipulations quaternioniques que l'on peut effectuer dans le cadre des images couleurs. Ces différentes manipulations ont été notamment étudiées par Gosselin en 2001 dans le cadre de son stage de DEA au département XLIM-SIC sous ma direction [95].

3.1.2.1 Transformations de $\mathbb{R}^3$

Comme nous l'avons déjà présenté, en nous appuyant sur l'analogie entre les vecteurs de $\mathbb{R}^3$ et les quaternions purs, nous pouvons écrire toutes les transformations classiques de l'espace avec uniquement des multiplications et des additions de quaternions. Nous proposons ici quelques illustrations qui nous serviront de support dans les futures transformations couleur.

La réflexion Axiale. Soient $q \in \mathbb{H}_0$ et $\mu \in \mathbb{H}^1 \cap \mathbb{H}_0$, alors

$$\hat{q} = -\mu q \mu \quad (3.10)$$

est le vecteur obtenu par la réflexion d'axe μ du vecteur q (cf. figure 3.1).

La projection orthogonale sur un vecteur. Soient $q \in \mathbb{H}_0$ et $\mu \in \mathbb{H}^1 \cap \mathbb{H}_0$ et q_1 le vecteur résultant de la projection orthogonale sur μ de q alors

$$q_1 = \frac{1}{2}(q - \mu q \mu). \quad (3.11)$$

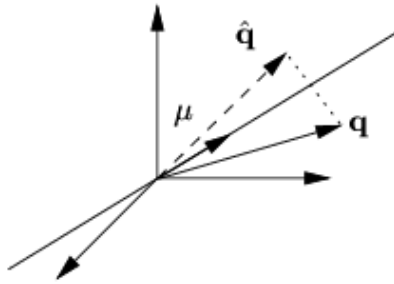


FIGURE 3.1 – Réflexion Axiale

Nous pouvons vérifier ceci par un simple calcul :

$$q_1 = \frac{1}{2}(q - \mu q \mu) = \frac{1}{2}(q - q + 2(q \cdot \mu)\mu) = (q \cdot \mu)\mu.$$

La projection orthogonale dans un plan. Soient $q \in \mathbb{H}_0$ et $\mu \in \mathbb{H}^1 \cap \mathbb{H}_0$ et q_2 est le vecteur résultant de la projection orthogonale dans le plan orthogonal à μ de q :

$$q_2 = \frac{1}{2}(q + \mu q \mu). \quad (3.12)$$

En effet,

$$q_2 = \frac{1}{2}(q + \mu q \mu) = \frac{1}{2}(q + q - 2(q \cdot \mu)\mu) = q - (q \cdot \mu)\mu.$$

Enfin, nous rappelons le lien entre les quaternions et une rotation.

La rotation axiale. Soient $q \in \mathbb{H}_0$, $\phi \in \mathbb{R}$ et $\mu \in \mathbb{H}^1 \cap \mathbb{H}_0$. Alors

$$q_3 = e^{\mu \frac{\phi}{2}} q e^{-\mu \frac{\phi}{2}} \quad (3.13)$$

est le vecteur résultant de la rotation d'axe μ et d'angle ϕ de q .

Etudions maintenant l'implication des quaternions dans le cadre des images couleur.

3.1.2.2 Quaternions et images couleur

Les premiers travaux proposant de coder les images couleur en utilisant des quaternions ont été publiés par S.J. Sangwine et al. [201]. Une couleur ne contenant que trois composantes dans l'espace RVB, il a été proposé de décrire l'information couleur sur la partie imaginaire des quaternions.

$$f : [m, n] \mapsto (f_r, f_g, f_b) \quad (3.14)$$

$$\mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{H}_0 \quad (3.15)$$

La couleur d'un pixel d'une image f aux coordonnées spatiales (m, n) sera donc codée de la manière suivante :

$$f[m, n] = f_r[m, n]i + f_g[m, n]j + f_b[m, n]k \quad (3.16)$$

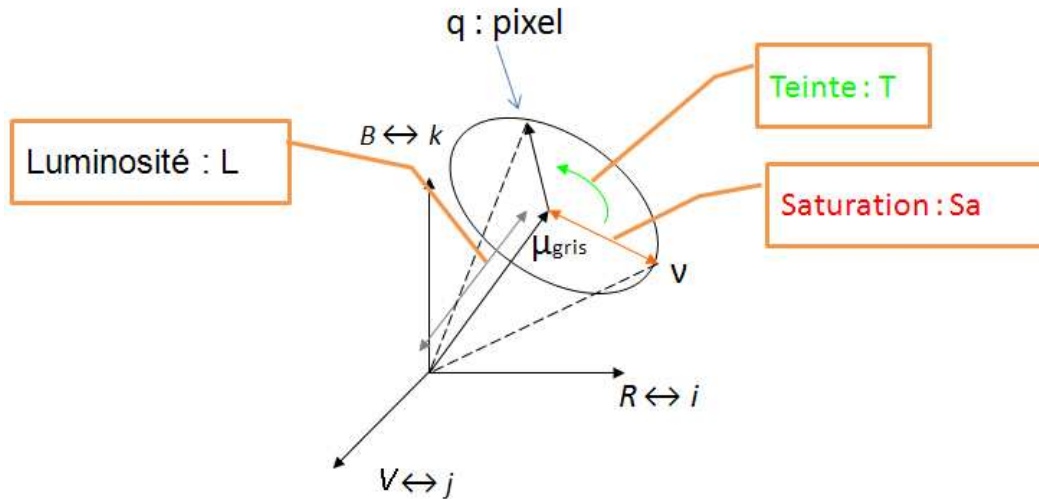


FIGURE 3.2 – Représentation Teinte, Saturation et Intensité

avec $f_r[m, n]$, $f_v[m, n]$ et $f_b[m, n]$ respectivement les composantes rouge, verte et bleue du pixel de coordonnées (m, n) .

Toutefois, ce modèle de représentation RVB n'est pas le plus adapté aux manipulations géométriques des caractéristiques colorimétriques. Par exemple, pour modifier la couleur d'un pixel sans modifier l'aspect coloré ou sa luminosité, il faut agir sur les trois composantes à travers des opérations similaires à une rotation dans l'espace des couleurs 3-D. C'est pourquoi classiquement lors de traitements couleur, on va plutôt coder l'information couleur dans un espace type *Teinte*, *Saturation* et *Intensité*. Il existe différentes représentations tentant de caractériser ces valeurs, mais dans tous les cas on retrouve toujours cette même idée : l'*Intensité* est la composante de clarté qui contient la quantité de lumière ou d'énergie ; la *Saturation* représente la pureté de la couleur et la *Teinte* la couleur pure proprement dite (rouge, vert, bleu, ...).

Cette interprétation de la couleur peut s'exprimer géométriquement à l'aide des quaternions. C'est ce qui a constitué notre point de départ dans la construction des outils. L'*Intensité* est placée sur l'axe μ_{gris} , tandis que la *Saturation* et la *Teinte* sont dans le plan perpendiculaire à cet axe qui correspond à un plan chromatique (voir figure 3.2). La *Teinte* est l'angle dans ce plan entre le pixel couleur et une couleur de référence et la *Saturation* est la distance entre le point considéré et l'axe des *Intensités*.

Considérons la couleur $f[m, n] = (f_r[m, n], f_v[m, n], f_b[m, n])$ d'un pixel codée par un quaternion pur q dans un espace couleur quelconque dont nous voulons extraire les coordonnées colorimétriques (*Intensité*, *Teinte* et *Saturation*) :

$$q = (0, f_r[m, n], f_v[m, n], f_b[m, n]). \quad (3.17)$$

Nous avons alors proposé la modélisation suivante. Il nous faut tout d'abord définir un vecteur μ (quaternion pur unitaire) représentant l'axe des *Intensités*. L'*Intensité* I de q est alors la projection de celui-ci sur μ :

$$I = q \cdot \mu. \quad (3.18)$$

Dans l'exemple illustré dans la figure 3.2, cet axe est défini comme l'axe de niveaux de gris, donc $\mu = (0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

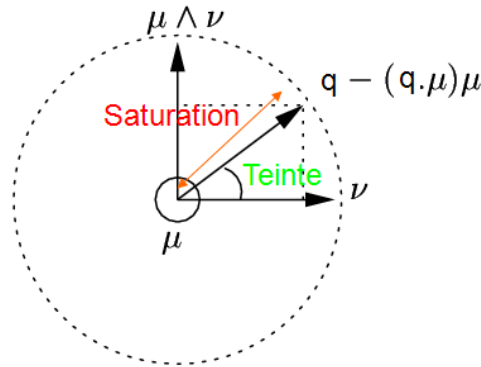


FIGURE 3.3 – Teinte et Saturation

La *Saturation* Sa étant la distance qui sépare le vecteur q de l'axe des intensités, elle s'exprime par :

$$Sa = |q - (q \cdot \mu)\mu| = |q \wedge \mu|. \quad (3.19)$$

Pour déterminer la *Teinte* T nous avons besoin d'un autre vecteur ν (quaternion pur unitaire) de référence perpendiculaire à μ afin de fixer l'origine de l'angle. Cette référence peut être par exemple fixée pour correspondre à un rouge "pur". Si nous reprenons l'exemple de la figure 3.2 avec l'axe μ défini comme l'axe de niveaux de gris alors $\nu = \sqrt{\frac{2}{3}}(i - \frac{1}{2}j - \frac{1}{2}k)$.

La *Teinte* est l'angle entre cet axe de référence et le vecteur q dans le plan perpendiculaire à μ passant par l'extrémité de q par exemple :

$$\cos T = \frac{q \cdot \nu}{|q \wedge \mu|} \text{ et } \sin T = \frac{q \cdot (\mu \wedge \nu)}{|q \wedge \mu|}. \quad (3.20)$$

Nous illustrons cette interprétation sur la figure 3.3 qui correspond à la vue par dessus de la figure précédente.

Ces manipulations ont notamment été décrites dans le travail de Gosselin [95]. On constate que l'introduction des quaternions permet une écriture simple et intuitive de la manipulation d'un espace couleur. Nous pouvons alors agir sur chacune des informations colorimétriques avec les différentes opérations rappelées en préambule. Notons que dans cette partie nous négligeons le post-traitement qui est souvent nécessaire pour maintenir les valeurs colorimétriques dans leur domaine de définition numérique. En effet, comme pour tout traitement en images, les opérations sont définies sur $\mathbb{R}$ alors que les pixels vont être codés entre 0 et 255. Les stratégies mises en oeuvre vont consister soit à effectuer un seuillage (par exemple toute valeur supérieure à 255 est ramenée à cette valeur limite) soit à effectuer une contraction et de translation de l'échelle de valeur pour ramener l'intervalle $[min, max]$ sur $[0, 255]$.

Nous présentons maintenant une manipulation particulière qui a été utilisée de nombreuses fois dans la littérature pour interpréter certains résultats dans le cadre des quaternions pour la couleur, et que nous retrouverons souvent dans la suite de ce chapitre.

3.1.2.3 La forme de Cayley-Dickson

La décomposition de Cayley-Dickson a été utilisée pour les pixels dont les couleurs sont codées avec des quaternions [81, 82] afin de séparer l'information en une partie intensité et une partie chromatique (comme nous l'avons déjà abordé dans le paragraphe précédent). Cette réécriture a surtout permis, comme nous le verrons, d'analyser les traitements et transformations réalisés avec les quaternions sous un angle "colorimétrique". Nous rappelons tout d'abord l'écriture d'un quaternion sous la forme de Cayley-Dickson. Nous avons vu qu'un quaternion peut être considéré comme un vecteur de la base (e, i, j, k) avec $e = 1$, le vecteur unitaire réel :

$$q = ae + bi + cj + dk$$

avec a, b, c, d quatre réels quelconques et a la partie réelle et (b, c, d) la partie imaginaire.

Rien ne nous empêche de choisir une autre base, par exemple $(e, \mu, \nu, \mu\nu)$ avec μ et ν deux quaternions purs unitaires orthogonaux. Leur produit $\mu\nu$ est aussi quaternion pur et orthogonal aux autres vecteurs de la base. Si nous écrivons q dans cette base :

$$q = q_e + q_\mu\mu + q_\nu\nu + q_{\mu\nu}\mu\nu \quad (3.21)$$

avec par exemple le réel $q_e = q \cdot e$, défini par le produit scalaire, et qui correspond à la projection de q sur e . De même on peut calculer $q_\mu = q \cdot \mu$, $q_\nu = q \cdot \nu$ et $q_{\mu\nu} = q \cdot \mu\nu$ des réels.

Nous constatons que l'on peut regrouper le quaternion q dans deux sous-espaces isomorphes à $\mathbb{C}$. En effet, le premier bloc (q_e, q_μ) correspond à la partie réelle et le terme de projection sur μ , le reste s'exprimant dans le plan $(\nu, \mu\nu)$ qui est le plan perpendiculaire à μ . Le quaternion peut donc s'écrire :

$$q = q_{\parallel} + q_{\perp}\nu \quad (3.22)$$

avec $q_{\parallel} = q_e e + q_\mu \mu$ et $q_{\perp} = q_\nu e + q_{\mu\nu} \mu$, $q_{\parallel}$ étant la partie parallèle à μ et $q_{\perp}$ la partie perpendiculaire.

Dans le cas de la manipulation des pixels couleur, deux quaternions purs sont utilisés pour cette décomposition : le premier vecteur μ est l'axe des gris dans le cube R,G,B.

$$\mu = \mu_{gris} = \frac{i + j + k}{\sqrt{3}}$$

Le second est l'axe perpendiculaire à μ dans la direction de la couleur rouge (pointant l'axe i)

$$\nu = \sqrt{\frac{2}{3}}(i, -\frac{j}{2}, -\frac{k}{2})$$

car le rouge est utilisé comme référence pour une teinte de valeur nulle dans le plan de chromaticité. Dans ce cas, lorsque l'on décompose une image couleur en parties parallèle et perpendiculaire à μ , il apparaît que cette décomposition permet de séparer l'information d'intensité sur la partie parallèle et l'information chromatique sur la partie perpendiculaire à μ comme nous le montrent les résultats de la figure 3.4. Nous retrouvons l'image en niveaux de gris et la représentation artificielle correspondant au plan chromatique. Notons que pour représenter la partie correspondant à la projection chromatique, nous avons fixé une valeur médiane suivant l'axe des gris. Nous retrouvons la représentation classique associée à des espaces "Luminance-Chrominance". Cette décomposition symplectique est utilisée, comme nous le verrons plus tard, dans l'analyse des images couleur par Ell et Sangwine dans [82], pour séparer l'information spectrale obtenue après transformée de Fourier en parties chromatique et achromatique, aidant ainsi à l'interprétation. Nous l'utiliserons aussi, sous différentes formes, pour décrypter la transformée de Fourier couleur.



FIGURE 3.4 – Décomposition symplectique sur des images couleur : (a) image originale ; (b) partie parallèle (c) partie perpendiculaire

3.1.2.4 Modification couleur à partir des quaternions

Notons T , Sa et I respectivement la *Teinte*, la *Saturation* et l'*Intensité* d'un pixel couleur associé au quaternion pur q défini comme précédemment (équation 3.17). Nous avons proposé différentes écritures permettant d'agir sur ces différents paramètres. Rappelons à nouveau que μ désigne l'axe *Intensité* et ν un axe perpendiculaire à μ servant de référence pour l'angle T .

Pour modifier la *Teinte*, et donc la couleur du pixel traité, il faut lui faire subir une opération de rotation autour de l'axe μ , puisque cela ne modifiera donc pas son intensité (projection sur l'axe μ) ni sa saturation (distance à l'axe μ). Pour cela nous pouvons utiliser l'opération de rotation décrite avec les quaternions :

$$\hat{q} = e^{\mu \frac{\phi}{2}} q e^{-\mu \frac{\phi}{2}} \quad (3.23)$$

Ici $\hat{q}$ a la *Teinte* de q incrémentée d'un angle ϕ . La figure 3.5(b) illustre cette transformation avec un paramètre $\phi = \frac{2\pi}{5}$. Nous constatons que par exemple les zones rouges orangées ont conservé la même intensité dans les couleurs mais qu'elles sont devenues vertes.

Pour modifier l'*Intensité* d'un facteur α et la *Saturation* d'un facteur β , nous allons décomposer le pixel couleur suivant ses informations chromatique et achromatique. Pour cela, nous rappelons que l'opération $\frac{1}{2}(q - \mu q \mu)$ correspond à la projection du vecteur q sur μ et est un vecteur dans le plan (e, μ) dont la norme est I . Pour dilater ou réduire I il suffit donc de multiplier ce vecteur par un réel α . Le raisonnement est le même pour Sa , à savoir que l'opération $\frac{1}{2}(q + \mu q \mu)$ est le vecteur résultant de la projection orthogonale dans le plan orthogonal à μ dont la norme est Sa . D'où la formule, où nous pouvons faire apparaître la forme de Cayley-Dickson que nous venons d'introduire :

$$\hat{q} = \alpha \underbrace{\left(\frac{1}{2}(q - \mu q \mu) \right)}_{(e, \mu)} + \beta \underbrace{\left(\frac{1}{2}(q + \mu q \mu) \right)}_{(\nu, \mu\nu)}. \quad (3.24)$$

La figure 3.5(c) illustre cette transformation. Dans cet exemple nous avons réduit la saturation diminuant ainsi l'éclat des différentes couleurs. Pour résumer, si l'on veut modifier tous les paramètres il faut appliquer l'opération suivante :

$$\hat{q} = \left(\frac{\alpha}{2} + \frac{\beta}{2} \right) e^{\mu \phi} q e^{-\mu \phi} - \left(\frac{\alpha}{2} - \frac{\beta}{2} \right) \mu e^{\mu \phi} q e^{-\mu \phi} \mu. \quad (3.25)$$



FIGURE 3.5 – Transformations vectorielles dans l'espace RGB. (a) Image originale. (b) Incrémentation de la Teinte de $\phi = \frac{2\pi}{5}$ avec $q \rightarrow e^{\mu\frac{\phi}{2}} q e^{-\mu\frac{\phi}{2}}$. (c) Dilatation de l'Intensité de $\alpha = 1.3$ et de la Saturation de $\beta = 0.6$ avec $q \rightarrow \left(\frac{\alpha}{2} + \frac{\beta}{2}\right) q - \left(\frac{\alpha}{2} - \frac{\beta}{2}\right) \mu q \mu$.

Avec pour résultat :

$$\begin{cases} \hat{T} = T + \phi; \\ \hat{S} = \beta S; \\ \hat{I} = \alpha I \end{cases} \quad (3.26)$$

Après transformation, l'extraction des composantes R, V et B de l'image ainsi modifiée se fait simplement en collectant les valeurs de la partie imaginaire, la partie réelle restant nulle.

Ces différentes manipulations développées dans le cadre du travail de Gosselin [95] nous permettent d'illustrer les premiers liens entre quaternions et couleur. Bien sûr ces différentes opérations peuvent aussi se calculer en utilisant un espace de type HLS et en modifiant les différentes composantes, mais la modélisation par quaternion remplit pleinement son rôle : une ré-écriture plus compacte de différentes opérations couleur permettant dans certains cas d'apporter un éclairage nouveau sur par exemples des opérations de filtrages.

Nous allons maintenant introduire le second outil mathématique étudié dans ce chapitre, à savoir les algèbres de Clifford.

3.1.3 Les algèbres de Clifford

3.1.3.1 Définition et construction

Notons que ce paragraphe reprend quelques éléments rédigés en collaboration avec M. Berthier du laboratoire LMA de La Rochelle, pour un chapitre d'un ouvrage collectif [37].

On se donne un espace vectoriel E (sur $\mathbb{R}$ pour nos données), de dimension finie, et une forme quadratique Q sur E . On construit une algèbre sur $\mathbb{R}$ qui contient E telle que si u est un vecteur de E , le produit $uu = u^2$ est égal à la valeur $Q(u)$ de la forme quadratique évaluée sur u . Dans ce cas, si u et v sont deux vecteurs de E plongés dans une telle algèbre, alors nécessairement : $uv + vu = 2B(u, v)$, en notant B la forme bilinéaire symétrique associée à Q ($Q(u) = B(u, u)$ et de plus pour tout $u, v \in ExE$ alors $B(u, v) = \frac{1}{2}[Q(u + v) - Q(u) - Q(v)]$). Le produit géométrique des deux vecteurs u et v est alors défini par la formule :

$$uv = B(u, v) + u \wedge v \quad (3.27)$$

où

$$u \wedge v = \frac{1}{2}(uv - vu)$$

est la partie antisymétrique de ce dernier.

On remarque que l'algèbre cherchée, contient les scalaires de $\mathbb{R}$ et les vecteurs de E , et elle contient également d'autres types d'objets. Par exemple, si E est l'espace vectoriel $\mathbb{R}^2$, Q la forme quadratique euclidienne et (e_1, e_2) une base orthonormée, alors l'élément $e_1 \wedge e_2$ est de carré -1; il ne peut donc pas s'agir d'un scalaire ou d'un vecteur. Nous voyons apparaître la notion de bivecteur.

D'une manière générale, chacun des éléments de la base d'une algèbre de Clifford est appelé un multivecteur. Cette notion de multivecteurs correspond à l'extension de l'interprétation géométrique que nous nous faisons des vecteurs. Un bivecteur correspondra à une portion de plan orientée du fait de ses deux dimensions. De même un trivecteur sera représenté par une portion 3D orientée de l'espace. Comme nous le verrons, il est possible d'effectuer des opérations géométriques en utilisant les multivecteurs (sous-espaces vectoriels de dimension $k < n$) et en leur appliquant des opérations algébriques. C'est pourquoi dans certains cas nous allons parler d'algèbre géométrique, par exemple notée $\mathcal{G}^3$ quand elle est associée à l'espace euclidien $\mathbb{R}^3$. Mais avant de développer cet aspect calculatoire, nous listons quelques propriétés des algèbres de Clifford et décrivons quelques exemples.

On se place maintenant dans le contexte qui nous intéresse, à savoir que l'espace vectoriel considéré est $\mathbb{R}^n$. On note $\mathbb{R}_{p,q}$ (avec $n = p + q$) l'algèbre de Clifford de l'espace vectoriel $\mathbb{R}^n$ muni de la forme quadratique Q :

$$Q(x_1, x_2, \dots, x_n) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_n^2. \quad (3.28)$$

De ce fait puisque $uu = u^2 = Q(u)$ pour tout u de $\mathbb{R}^n$ cela implique que $uv + vu = 2B(u, v) = 2u.v$ avec $u.v$ le produit scalaire dans $\mathbb{R}^n$. Dans ce cas le produit géométrique des deux vecteurs u et v s'écrit :

$$uv = u.v + u \wedge v. \quad (3.29)$$

Quelques exemples d'algèbres de Clifford De nouveau je reprends quelques éléments de [37].

Si nous prenons l'exemple le plus simple, à savoir celui de l'algèbre $\mathbb{R}_{0,1}$, à savoir l'algèbre de Clifford de $\mathbb{R}$ muni de la forme quadratique $Q(x) = -x^2$. Si e est un vecteur de $\mathbb{R}$ de norme euclidienne 1, alors $e^2 = -1$. Le couple $(1, e)$ est une base de $\mathbb{R}_{0,1}$ et on réalise un isomorphisme entre cette algèbre et $\mathbb{C}$ avec $1 \rightarrow 1$ et $e \rightarrow i$.

Soit (e_1, e_2) une base orthonormée pour le produit scalaire euclidien de $\mathbb{R}^2$. Dans $\mathbb{R}_{0,2}$, on a les relations suivantes : $e_1^2 = e_2^2 = -1$. Dans ce cas, on réalise un isomorphisme entre $\mathbb{R}_{0,2}$ et $\mathbb{H}$ avec $1 \rightarrow 1$, $e_1 \rightarrow i$, $e_2 \rightarrow j$ et $e_1 e_2 \rightarrow k$.

Soit (e_1, e_2, e_3) une base orthonormée pour le produit scalaire euclidien de $\mathbb{R}^3$. L'algèbre $\mathbb{R}_{3,0}$ est de dimension 8 et admet pour base la famille $(e_0 = 1, e_1, e_2, e_3, e_1 e_2, e_2 e_3, e_1 e_3, e_1 e_2 e_3)$.

En prenant comme exemple $\mathbb{R}_{3,0}$, il est intéressant d'introduire une terminologie particulière que nous utiliserons dans le cadre de Fourier. Dans la littérature associée aux mathématiques pour la physique, on trouve des éléments appelés Spineurs, qui servent notamment, comme nous allons le voir, à décrire une opération de rotation. Ils sont quelquefois introduits comme des éléments qui préservent le grade à travers un produit en "sandwich". Concrètement, considérons l'ensemble des éléments S qui transforme un vecteur v en un vecteur par l'opération SvS^{-1} . Si S est normalisé $S^2 = 1$ et de grade pair, alors il est appelé un spineur.

Illustrons ce concept dans le cadre de $\mathbb{R}_{3,0}$. Nous allons extraire une sous algèbre de $\mathbb{R}_{3,0}$ qui joue un rôle bien particulier notamment dans l'interprétation que l'on peut faire de la transformée de Fourier. Les éléments pairs, de la forme $q = q_0 1 + q_1 e_1 e_2 + q_2 e_2 e_3 + q_3 e_1 e_3$, forment une sous algèbre de $\mathbb{R}_{3,0}$ isomorphe à $\mathbb{H}$; on constate en effet que $(e_1 e_2)^2 = (e_2 e_3)^2 = (e_1 e_3)^2 = -1$. Ainsi le groupe $\mathbb{H}_1$ des quaternions de module 1 s'identifie aux éléments pairs de norme 1 de l'algèbre $\mathbb{R}_{3,0}$. Ces éléments particuliers vont donc être ce que l'on appelle les *spineurs* de cette algèbre. En effet, ils sont de norme 1, ils sont de grade pair et, étant donné le lien fait avec $\mathbb{H}_1$, leur action sur les vecteurs est équivalente à l'action des quaternions de norme 1 sur les quaternions imaginaires purs, à savoir une rotation.

Au delà de l'introduction du concept des spineurs, cet exemple permet d'explicitier l'intérêt de l'utilisation d'un formalisme plus complexe. Pour cela reprenons l'exemple utilisé par M. Berthier dans [37]. Comme nous venons de le voir, pour définir une rotation dans l'espace d'un vecteur, on peut tout simplement utiliser les quaternions, mais dans ce cas, dans l'opération de rotation, ce qui tourne et ce qui fait tourner sont des quaternions purs. Dans $\mathbb{R}_{3,0}$, avec le formalisme exposé ci-dessus, nous avons deux objets de nature différente les spineurs et les vecteurs, ce qui permet de lever une certaine ambiguïté.

Maintenant que nous avons introduit les concepts de base, nous allons analyser le contexte calculatoire. En effet, comme nous l'avons déjà énoncé, l'un des intérêts des algèbres de Clifford est que le produit géométrique permet de coder de façon algébrique des transformations géométriques. C'est ce que nous proposons de décrire dans le paragraphe suivant.

3.1.3.2 Un contexte calculatoire

Dans ce paragraphe, nous allons considérer l'algèbre généralement utilisée dans le cadre de la manipulation des images couleur, à savoir l'algèbre géométrique notée $\mathcal{G}_n$ qui correspond précisément l'algèbre $\mathbb{R}_{n,0}$, avec $Q(u) = u^2 \geq 0$, pour tout vecteur u de $\mathbb{R}^n$. Nous remarquons que, étant donné l'algèbre $\mathbb{R}_{n,0}$ et non $\mathbb{R}_{0,n}$, le produit de deux éléments de base sera égal à l'unité. Dans la suite, on note $e_{i_1 i_2 \dots i_k}$ l'élément $e_{i_1} e_{i_2} \dots e_{i_k}$.

Vocabulaire. Un *multi-vecteur* est un élément de l'algèbre, c'est à dire une combinaison linéaire des éléments de la base de l'algèbre. Un multi-vecteur de $\mathcal{G}_n$ s'écrit donc sous la forme :

$$A = \sum_{k=0}^n \langle A \rangle_k \quad (3.30)$$

avec $\langle A \rangle_k$ la partie *k-vectorielle* de A ou l'opérateur de *grade* k .

Par exemple, on peut exprimer un multi-vecteur A quelconque de $\mathcal{G}_3$ de la manière suivante :

$$A = A_0 e_0 + A_1 e_1 + A_2 e_2 + A_3 e_3 + A_{23} e_{23} + A_{31} e_{31} + A_{12} e_{12} + A_{123} e_{123} \quad (3.31)$$

On peut donc aussi décomposer le multi-vecteur A suivant ces parties *k-vectorielles*, par exemple le grade 0 de A est $\langle A \rangle_0 = A_0$ et le grade 2 de A est $\langle A \rangle_2 = A_{23} e_{23} + A_{31} e_{31} + A_{12} e_{12}$.

Les produits. Comme nous l'avons vu en prenant $(e_1, e_2, e_3, \dots, e_n)$ les 1-vecteurs formant la base de $\mathbb{R}^n$ on obtient par produit géométrique, les bivecteurs, les trivecteurs et ainsi de suite jusqu'au pseudo-scalaire $e_{123 \dots n}$. La généralisation des produits aux multivecteurs va alors se calculer à travers ces différents éléments de base de $\mathcal{G}_n$ en appliquant la loi d'associativité et de distributivité par rapport à l'addition.

Prenons par exemple $\mathcal{G}_2$, On indique dans le tableau 3.1 les différents produits géométriques entre les éléments de la base de $\mathcal{G}_2$.

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_{12}$
1	1	e_1	e_2	e_{12}
$\mathbf{e}_1$	e_1	1	e_{12}	e_2
$\mathbf{e}_2$	e_2	$-e_{12}$	1	$-e_1$
$\mathbf{e}_{12}$	e_{12}	$-e_2$	e_1	-1

TABLE 3.1 – Produits géométriques des éléments de la base de $\mathcal{G}_2$

On peut, dès lors, obtenir par combinaison linéaire le produit géométrique de deux multivecteurs quelconques de $\mathcal{G}_2$.

Le produit extérieur, noté $\wedge$, engendre l'algèbre extérieure de Grassmann des multivecteurs. Dans $\mathcal{G}_n$ il peut s'écrire par rapport au produit géométrique de la manière suivante, avec A_r et B_s deux multivecteurs homogènes de grade r et s :

$$A_r \wedge B_s = \langle A_r B_s \rangle_{r+s} \quad (3.32)$$

Le produit extérieur entre deux multivecteurs quelconques se généralise ensuite par distributivité.

Le produit interne permet de généraliser la notion d'orthogonalité définie par le produit scalaire de deux vecteurs de $\mathbb{R}^n$ à deux k -vecteurs. Si les deux k -vecteurs sont orthogonaux, leur produit interne est nul. Il se définit à partir du produit géométrique avec A_r et B_s deux multivecteurs homogènes de grade r et s comme :

$$A_r \cdot B_s = \langle A_r B_s \rangle_{|r-s|}. \quad (3.33)$$

De même, le produit intérieur entre deux multivecteurs quelconques se généralise ensuite par distributivité.

Notions Complémentaires. La *Réversion* est une notion qui sert notamment pour définir l'inverse et la norme. L'opération de réversion renverse l'ordre des vecteurs dans tous les multivecteurs. Donc la réversion d'un produit de vecteurs est définie par :

$$(\widetilde{ab \dots c}) = c \dots ba.$$

Par suite, pour un multi-vecteur quelconque $A = \sum_{k=0}^n \langle A \rangle_k$ de $\mathcal{G}_n$, la *réversion* $\tilde{A}$ est définie par :

$$\begin{aligned} \tilde{A} &= \sum_{k=0}^n \widetilde{\langle A \rangle_k} \\ &= \sum_{k=0}^n (-1)^{\frac{k(k-1)}{2}} \langle A \rangle_k \end{aligned} \quad (3.34)$$

Le *Produit scalaire* est défini simplement comme la partie scalaire du produit de deux multivecteurs :

$$A * B \equiv \langle AB \rangle_0.$$

Ce produit scalaire est linéaire et symétrique. Il permet d'associer à chaque multivecteur A une notion de norme scalaire et positive :

$$|A|^2 = \widetilde{A} * A \geq 0. \quad (3.35)$$

À partir de cette notion de norme, on peut alors introduire la notion d'inverse d'un multivecteur homogène de grade r non nul :

$$X_r^{-1} = \frac{\widetilde{X_r}}{|X_r|^2}. \quad (3.36)$$

La *Dualité* est définie à partir d'une notion que nous avons déjà abordée, à savoir le pseudo-scalaire. Le produit extérieur de n vecteurs conduit à un multivecteur de grade n . Pour un espace vectoriel donné, l'élément de grade maximum est unique, à multiplication par un scalaire près. Le produit extérieur de n vecteurs est alors un multiple du pseudo-scalaire de $\mathcal{G}_n$ noté $I_n = e_{123\dots n}$.

La dualité permet d'introduire une notion d'orthogonalité : pour un sous-espace F d'un espace vectoriel E , il existe un sous-espace $F^\perp$ de E qui est orthogonal à F et vérifiant : $F \oplus F^\perp = E$.

À partir du produit du pseudo-scalaire $I_n = e_{123\dots n}$ avec un multivecteur A_r de grade r , on définit le *dual* A_r^* qui est de grade $n - r$:

$$A_r^* = \widetilde{I_n} A_r. \quad (3.37)$$

Par exemple le dual d'un bivecteur formé par deux 1-vecteurs a et b dans $\mathcal{G}^3$ est le 1-vecteur B^* résultant du produit vectoriel de a et b comme l'illustre la figure 3.6.

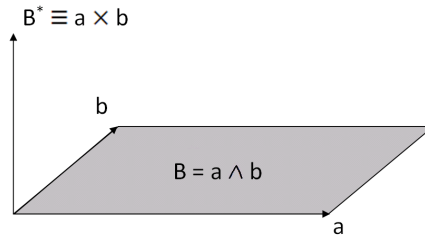


FIGURE 3.6 – le 1-vecteur B^* est le dual du bivecteur $B = a \wedge b$ dans $\mathcal{G}^3$.

Transformations géométriques de multivecteurs. Comme leur nom le suggère, les algèbres géométriques permettent de définir des transformations géométriques en utilisant le formalisme des notations algébriques. Nous pouvons tout d'abord définir des opérations de projection et réjection entre des 1-vecteurs.

Soient u, ϑ des 1-vecteurs de E , la projection $u_{\parallel}$ de u sur ϑ est donnée par :

$$u_{\parallel} = (u \cdot \vartheta) \vartheta^{-1} \quad (3.38)$$

De même, la réjection $u_{\perp}$ de u par rapport à ϑ est donnée par :

$$u_{\perp} = (u \wedge \vartheta) \vartheta^{-1} \quad (3.39)$$

Dans ce cas, on constate que si u est un vecteur de E , on dispose d'une *décomposition en partie parallèle et orthogonale relativement à ϑ* : $u = u_{\parallel} + u_{\perp}$, avec $u_{\parallel} \wedge \vartheta = 0$ et $u_{\perp} \cdot \vartheta = 0$, $u_{\perp} = (u \wedge \vartheta) \vartheta^{-1}$ et

$u_{||} = (u.\vartheta)\vartheta^{-1}$. On retrouve ainsi la décomposition symplectique vue dans le cadre des Quaternions. Si le vecteur ϑ correspond à l'axe des gris, cette décomposition permet alors de définir la partie achromatique et chromatique du vecteur associé au pixel couleur u . Remarquons que nous introduisons de nouvelles notations, à savoir ϑ l'axe des gris (μ pour les quaternions) et u en ce qui concerne le vecteur couleur (q pour les quaternions), ceci afin de lever toute ambiguïté dans la suite du propos entre les manipulations quaternioniques et celles faites dans le domaine des algèbres géométriques.

De la même façon, on en déduit que la réflexion par rapport à l'hyperplan orthogonal à un vecteur inversible ϑ de E est donnée par la formule : $s_{\vartheta}(u) = -\vartheta u \vartheta^{-1}$. De même, le symétrique d'un vecteur u par rapport à ϑ est donné par $\sigma_{\vartheta}(u) = \vartheta u \vartheta^{-1}$.

Cette notion peut se généraliser avec l'utilisation du pseudo-scalaire. A savoir que si I est le pseudo scalaire associé à un sous-espace vectoriel F d'un espace vectoriel E alors pour tout vecteur u de E on peut définir une décomposition en une partie parallèle et perpendiculaire à F :

$$u = u_{||} + u_{\perp} \quad (3.40)$$

avec

$$u_{\perp} = (u \wedge I)I^{-1} \text{ et } u_{||} = (u.I)I^{-1}.$$

Enfin, nous introduisons la notion de rotation et le lien avec le groupe des Spineurs. Nous nous plaçons dans $\mathbb{R}_{3,0}$ et nous proposons d'étudier la composée de deux réflexions. Pour cela, considérons deux plans de vecteurs orthogonaux unitaires a et b . La composée des deux réflexions par rapport à ces plans est (en utilisant les formules vues précédemment) :

$$x \rightarrow b a x a^{-1} b^{-1}. \quad (3.41)$$

Comme il est bien connu, la composée de deux réflexions est une rotation. L'axe de cette rotation est orthogonal au plan des vecteurs a et b . Ce plan est donc donné par $a \wedge b$. L'angle θ de cette rotation est le double de l'angle entre les deux vecteurs. Le produit ab peut s'écrire $ab = a.b + a \wedge b$. Puisque a et b sont unitaires et $\frac{\theta}{2}$ est l'angle entre les deux vecteurs, on obtient

$$ab = \cos \frac{\theta}{2} + \frac{\sin \frac{\theta}{2}}{|a \wedge b|} a \wedge b = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} J \quad (3.42)$$

avec J le bivecteur unitaire correspondant au plan engendré par a et b . Une rotation va donc s'écrire :

$$x \rightarrow e^{-J \frac{\theta}{2}} x e^{J \frac{\theta}{2}}. \quad (3.43)$$

Donc les rotations dans $\mathbb{R}_{3,0}$ s'écrivent :

$$x \rightarrow \tau^{-1} x \tau \quad (3.44)$$

avec τ correspondant à la somme d'un scalaire et d'un bivecteur donc des éléments de grade pair. D'après ce que nous avons introduit durant les définitions, on constate que τ est un spineur de $\mathbb{R}_{3,0}$

En généralisant, considérons maintenant l'algèbre $\mathbb{R}_{n,0}$, de sorte que $Q(u) = u^2 \leq 0$, pour tout vecteur u de $\mathbb{R}^n$. Le groupe $Spin(n)$ est défini de la façon suivante :

$$Spin(n) = (x \in \mathbb{R}_{n,0}^{(0)}, Q(x) = -1, x v x^{-1} \in \mathbb{R}^n \quad \forall v \in \mathbb{R}^n) \quad (3.45)$$

avec $\mathbb{R}_{n,0}^{(0)}$ qui désigne l'ensemble des éléments pairs de l'algèbre $\mathbb{R}_{n,0}$. Comme nous venons de l'illustrer dans $\mathbb{R}_{3,0}$, les groupes $Spin(n)$ permettent d'encoder les rotations.

3.1.4 Algèbres géométriques pour les images couleur

Comme pour les quaternions, il est en général assez naturel d'encoder l'information couleur associée à l'espace RVB sur les vecteurs, c'est à dire les éléments de grade 1, de l'algèbre $\mathcal{G}_n$. Nous pouvons, grâce à l'utilisation des transformations géométriques de $\mathcal{G}_3$, exprimer le vecteur couleur en termes de composantes perceptuelles, c'est-à-dire de teinte, de saturation et de clarté, pour ainsi définir des opérations couleur. Nous avons proposé une modélisation en ce sens notamment dans le cadre de la thèse de P. Denis [62].

3.1.4.1 Teinte, saturation et intensité à partir d'un vecteur couleur

A l'instar de ce que nous avons vu pour un quaternion donné en RVB (Rouge Vert Bleu), les opérations géométriques décrites précédemment permettent d'exprimer un pixel couleur u donné de l'espace RVB codé à l'aide de $\mathcal{G}_3$ sous la forme teinte, saturation et intensité. De nouveau, la lecture géométrique de l'espace couleur permet d'interpréter les différentes transformations algébriques : on définit ϑ le 1-vecteur associé à l'axe de gris ($\vartheta = \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3)$), v celui portant le vecteur couleur rouge pur (rappel $v = \sqrt{\frac{2}{3}}(e_1, -\frac{e_2}{2}, -\frac{e_3}{2})$), et nous notons u le 1-vecteur couleur quelconque ($u = f_r[m, n]e_1 + f_g[m, n]e_2 + f_b[m, n]e_3$). Si nous reprenons la lecture géométrique exposée dans le cadre des quaternions, et en reprenant l'écriture des différentes opérations géométriques dans le cadre de $\mathcal{G}_3$ vu précédemment, alors nous pouvons proposer la modélisation suivante :

- l'intensité est la norme de la projection du 1-vecteur couleur RVB u sur l'axe des niveaux de gris ϑ , elle s'exprime par :

$$I = |(u \cdot \vartheta) \vartheta^{-1}| \quad (3.46)$$

- la saturation est la distance du vecteur couleur RVB u par rapport à l'axe des niveaux de gris, c'est donc la norme de la réjection du vecteur couleur par rapport à cet axe :

$$Sa = |(u \wedge \vartheta) \vartheta^{-1}| \quad (3.47)$$

- la teinte est l'angle entre un 1-vecteur de référence que l'on notera v et la partie perpendiculaire à l'axe des niveaux de gris du 1-vecteur couleur u (donc par rapport à ϑ).

Comme $u_{\perp} = (u \wedge \vartheta) \vartheta^{-1}$, la teinte est donc obtenue par la formule suivante :

$$T = \cos^{-1} \left(\frac{u_{\perp}}{|u_{\perp}|} \cdot v \right) \quad (3.48)$$

Bien sûr, il est tout à fait possible d'effectuer la conversion dans l'autre sens c'est à dire d'exprimer le vecteur u en fonction de sa teinte, sa saturation et son intensité. Ceci se fait, là aussi, à partir de simples opérations géométriques.

$$u = I\vartheta + Sa e^{\vartheta T} v \quad (3.49)$$

Dans ce cas, u contient la description de la couleur du pixel étudié sous la forme R,G et B dans ses composantes 1-vectorielles. Ceci permet, comme avec les quaternions, d'effectuer des modifications sur la description couleur du pixel dans le domaine HSI puis de revenir dans le domaine initial RGB.

Nous avons proposé, durant le travail de P. Denis [63], d'utiliser ces différentes expressions pour effectuer des opérations de modification de ces trois caractéristiques couleur. Nous les rappelons dans la section suivante.

3.1.4.2 Opérations sur les caractéristiques couleur

Changement de teinte. La teinte des couleurs peut être modifiée en utilisant, comme nous l'avons vu, une rotation autour de l'axe des niveaux de gris ϑ (1-vecteur unitaire) :

$$u' = e^{-\vartheta \frac{\theta}{2}} u e^{\vartheta \frac{\theta}{2}} = I e^{-\vartheta \frac{\theta}{2}} \vartheta e^{\vartheta \frac{\theta}{2}} + S e^{-\vartheta \frac{\theta}{2}} e^{\vartheta T} v e^{\vartheta \frac{\theta}{2}}$$

$$u' = I \vartheta + S e^{\vartheta(T+\theta)} v \longrightarrow T' = T + \theta$$

La figure 3.7 illustre un tel changement de teinte. L'image d'origine (a) a été modifiée par une rotation de l'ensemble de ses pixels couleur autour de l'axe des niveaux de gris et d'un angle de $\pi/3$. Ainsi le toit rouge se retrouve vert, le seuil vert se retrouve bleu, etc.

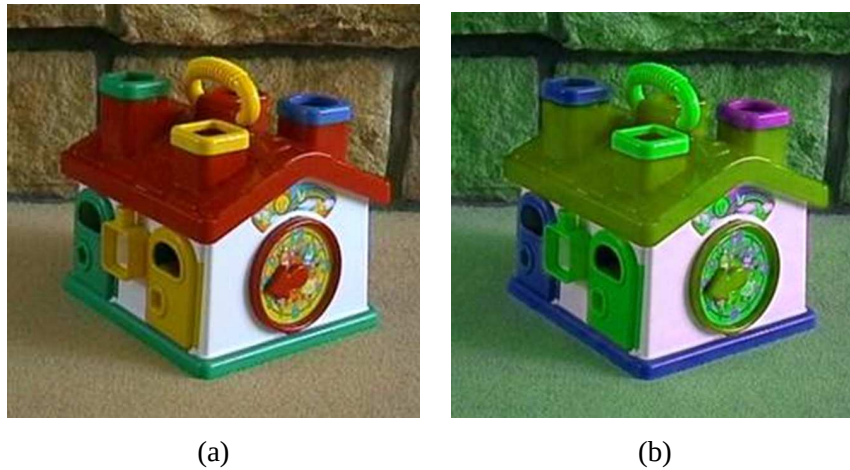


FIGURE 3.7 – Modification de la teinte : (a) image originale (b) image modifiée

Modification de la saturation. Pour modifier la saturation il suffit d'effectuer une translation de u d'un 1-vecteur pondéré par le coefficient $\beta \in \mathbb{R}$ suivant l'axe des saturations v , c'est à dire sa réjection par rapport à ϑ .

On obtient alors :

$$u' = I \vartheta + S' e^{\vartheta^* T} v = u + \beta e^{\vartheta^* T} v$$

$$u' = I \vartheta + (S + \beta) e^{\vartheta^* T} v \longrightarrow S' = S + \beta$$

La figure 3.8 illustre un tel changement de saturation. L'image d'origine (a) a été modifiée de la façon décrite en (3.1.4.2) pour obtenir l'image (b). L'effet produit par ce type d'opération est le suivant :

- Lorsqu'on retire de la saturation à l'image, comme dans le cas de la figure 3.8, l'image obtenue présente des couleurs pâles ou délavées par rapport à l'image d'origine.
- Lorsque l'image obtenue est plus saturée que l'image d'origine, ses couleurs paraissent alors plus vives.

Modification de l'intensité. Pour modifier l'intensité d'une couleur, on augmente ou on diminue sa clarté. Pour cela on effectue une translation de u d'un 1-vecteur pondéré par le coefficient $\alpha \in \mathbb{R}$ suivant l'axe des intensités. On obtient ainsi :

$$u' = u + \alpha \vartheta = I \vartheta + S e^{\vartheta^* T} v + \alpha \vartheta$$

$$I' = I + \alpha$$



FIGURE 3.8 – Modification de la saturation : (a) image originale (b) image modifiée

On remarquera qu'il est possible d'effectuer toutes ces transformations couleur grâce au fait que l'information couleur est contenue dans la partie vectorielle d'un multivecteur de $\mathcal{G}_3$.

De nouveau, nous constatons que les outils ainsi introduits permettent une écriture élégante et compacte de différentes transformations classiques calculées dans un espace couleur de type HSI. Dans ce cadre, l'algèbre géométrique remplit un peu le même rôle que dans le domaine de l'infographie, à savoir de permettre une écriture compacte se traduisant facilement dans un langage informatique. C'est pourquoi en parallèle de ces travaux, j'ai développé une Toolbox (sous la forme d'une librairie DLL) pour l'environnement Matlab ou Scilab.

Après avoir analysé et développé les définitions de bases ainsi que les manipulations élémentaires couleur, nous nous sommes penchés sur ce qui constituait notre objectif de départ : la généralisation des outils de traitement du signal aux données vectorielles.

Etant donné les modélisations exposées ci-dessus, nos premiers développements ont concerné des manipulations spatiales et donc des opérations de filtrage. C'est ce que nous proposons de décrire dans la section suivante.

3.2 Filtrage quaternionique et cliffordien : application à la détection de contours couleur

Durant ces dernières années, différentes études ont proposé d'introduire des opérations de type filtrage dans le cadre des quaternions ou de l'algèbre de Clifford. La grande majorité de ces travaux ont pour illustration la détection des contours couleur. Ils se font suivant deux approches : soit à travers une généralisation des outils de traitement du signal définis à l'origine dans le cadre des complexes [200, 64] soit avec une utilisation directe des propriétés des algèbres [28, 20]. Notons que ces deux stratégies aboutissent en général au même outil, mais leur interprétation peut être différente. Nous retrouverons par la suite cette complémentarité à travers la définition d'une transformée de Fourier couleur. Tout d'abord afin d'aborder la notion de filtrage, nous reprenons les quelques travaux de la littérature concernant la généralisation du produit de convolution.

3.2.1 Convolution

Dans le cadre des quaternions, nous pouvons définir un produit de convolution, opération qui aura bien sûr son importance dans le cadre d'une approche spatiale du traitement des images couleur. La convolution quaternionique a été étudiée par différents auteurs notamment par Pei et al. [177]. Pour simplifier les notations, nous illustrons ci-dessous les approches monodimensionnelles.

Soient f, g et h trois fonctions intégrables de $\mathbb{Z}$ à valeurs dans $\mathbb{H}$. La fonction définie par :

$$\forall x \in \mathbb{R}, f \star g[n] = \sum_{\tau \in \mathbb{Z}} f[n - \tau]g[\tau]$$

est appelée *convolution quaternionique à un côté de f et g* .

Il est important de souligner que la non commutativité du produit des quaternions permet d'envisager d'autres possibilités. Par exemple, la fonction définie par :

$$\forall x \in \mathbb{R}, h \star f \star g[n] = \sum_{\tau \in \mathbb{Z}} h[\tau]f[x - \tau]g[\tau]$$

est appelée *convolution quaternionique à deux côtés de f, g et h* .

L'une des propriétés fondamentales pour la notion de filtrage est la traduction du produit temporel de convolution en un produit terme à terme dans le domaine de Fourier (et donc la définition pour un filtre de la notion de gain fréquentiel). Du fait de la non-commutativité du produit quaternionique, cette relation n'existe plus directement dans le cadre des quaternions. Etudions par exemple le cas où l'on choisit comme transformée de Fourier, celle obtenue en remplaçant directement l'imaginaire pur i par un quaternion pur unitaire (nous détaillerons les différentes transformées Fourier quaternionique par la suite) :

Définition 4 Soient $N \in \mathbb{N}$, $f \in \mathcal{L}^2([0..N - 1], \mathbb{H})$ et $\mu \in \mathbb{H}_0 \cap \mathbb{H}^1$, alors la fonction définie par

$$\forall k \in [0..N - 1], F_{\mu}^r(f)[k] = \sum_{x=0}^{N-1} f(x)e^{-2\pi\mu\frac{kx}{N}}$$

est appelée *transformée de Fourier quaternionique discrète de f de direction μ (à droite car l'exponentielle est à droite du signal analysé)*.

Si l'on utilise cette transformée de Fourier, alors la relation entre les deux domaines va être la suivante. Soient $f, g \in L^2(\mathbb{R}, \mathbb{H})$ et h définie par $h = f \star g$ et telle que $h \in L^2(\mathbb{R}, \mathbb{H})$ alors $\forall k \in \mathbb{R}$,

$$\mathcal{F}_{\mu}^r(h)[k] = \mathcal{F}_{\mu}^r(f)[k]\mathcal{F}_{\mu}^r(g_a)[k] + \mathcal{F}_{\mu}^r(f)[-k]\nu\mathcal{F}_{\mu}^r(g_b)[k] \quad (3.50)$$

avec $g_a = \frac{1}{2}(g - \mu g \mu)$ et $g_b = \frac{1}{2}(g + \mu g \mu)$. Nous retrouvons ici la décomposition en un élément parallèle et perpendiculaire, cette décomposition se faisant en fonction de la direction μ choisie dans l'élément de base de la transformée de Fourier. Remarquons que si l'une des deux fonctions convoluées est paire, alors on retrouve la formulation classique. Le passage à la dimension deux n'apporte pas de changement particulier, les définitions et propriétés sont identiques. Nous discuterons de nouveau de cette expression à la fin de ce chapitre dans le cadre de la définition d'un banc de filtres quaternionique.

Dans le cadre de l'algèbre de Clifford, il existe très peu de travaux généralisant les outils de traitement d'image. Les principaux concernent plutôt la transformée de Fourier et nous les étudierons par la suite. Nous pouvons mentionner les travaux de Ebling et al. menés à l'origine dans le cadre des champs de vecteurs [76]. Soient F et G deux champs de multivecteurs (donc deux applications de $\mathbb{Z}^d$ dans $\mathcal{G}_n$) alors

$$F \star G[m] = \sum_{\tau \in \mathbb{Z}} F[m - \tau]G[\tau] \quad (3.51)$$

est appelée *la convolution de Clifford de F et G* . Dans le produit de convolution est utilisé le produit géométrique, donc si les deux éléments sont des champs de scalaires, on retrouve la convolution classique. Si les deux éléments sont des multivecteurs, alors, par distributivité, le résultat est une combinaison de différentes convolutions.

Nous constatons que la forme du produit de convolution reste finalement classique. Dans le cadre de notre travail, nous n'avons pas non plus remis en cause le principe du filtrage (notions de somme et de produit définis selon la formalisation utilisée). Le travail a eu plutôt comme objectif la définition des filtres. C'est pourquoi nous étudions maintenant ce qui a constitué le principal développement dans le cadre du traitement de l'image couleur par quaternions à savoir la mesure de ruptures couleur.

3.2.2 Les filtres quaternioniques et leurs extensions pour la détection de contours couleur

Se pose maintenant la définition du masque du filtre couleur. On peut reprendre tous les filtres linéaires classiques en convoluant le signal q avec une fonction réelle h :

$$q \star h[x] = \sum_{\tau=0}^{N-1} q[x - \tau \bmod N]h[\tau \bmod N]$$

mais on obtient alors le même résultat que si on avait convolué chaque couleur séparément. On peut également essayer d'utiliser certaines propriétés liées aux quaternions vues dans les paragraphes de présentation des outils (section 2.1). Nous allons décrire la façon dont Sangwine et al. [203] s'en servent pour construire une *mesure de variations locales couleur* à partir d'une généralisation quaternionique du filtre de Prewitt (en reprenant le produit de convolution à deux cotés décrit précédemment).

Nous reprenons notre modélisation des pixels sous la forme d'un quaternion pur, soit $[m, n] \mapsto f[m, n] = f_r[m, n]i + f_g[m, n]j + f_b[m, n]k$ une image couleur (codée en RVB) considérée comme une application d'un domaine de $\mathbb{Z}^2$ dans $\mathbb{H}_0$. La formule de convolution quaternionique proposée par Sangwine s'écrit

$$g[m, n] = \sum_{x=-1}^1 \sum_{y=-1}^1 h_L[x, y]f[m - x, n - y]h_R[x, y] \quad (3.52)$$

avec

$$h_L = \begin{pmatrix} \alpha & 0 & 1 \\ \alpha & 0 & 1 \\ \alpha & 0 & 1 \end{pmatrix} \quad h_R = \begin{pmatrix} \bar{\alpha} & 0 & 1 \\ \bar{\alpha} & 0 & 1 \\ \bar{\alpha} & 0 & 1 \end{pmatrix}$$

où α est un quaternion de module 1, choisi ici comme étant

$$\alpha = \exp\left(\frac{\pi}{2}\mu\right)$$

avec

$$\mu = \frac{1}{\sqrt{3}}(i + j + k)$$

Le système de coordonnées est cartésien, centré sur le pixel considéré. Le pixel de l'image filtrée se calcule aisément :

$$g(m, n) = \{[\alpha f(m+1, n+1)\bar{\alpha} + f(m+1, n-1)] + [\alpha f(m, n+1)\bar{\alpha} + f(m, n-1)] \\ + [\alpha f(m-1, n+1)\bar{\alpha} + f(m-1, n-1)]\}$$

Nous constatons que α est un spineur et que nous retrouvons l'écriture d'une opération de rotation, rotation de π autour de l'axe des gris. Lorsque le masque parcourt une zone de l'image où la couleur varie peu selon la verticale, le résultat de la convolution est proche de l'axe de gris (de vecteur directeur μ). Dans le cas contraire, le résultat met en évidence une couleur avec un niveau de saturation significatif. Ceci vient du fait que ce filtre ajoute à chaque vecteur son voisin proche tourné de π autour de l'axe des intensités μ .

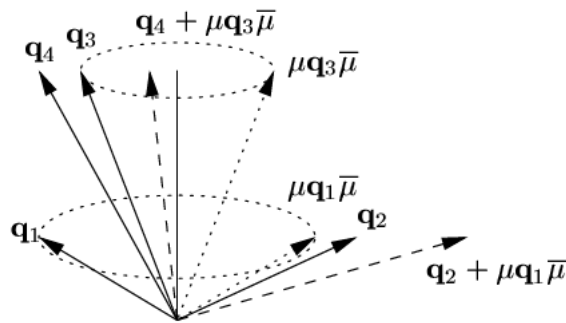


FIGURE 3.9 – Détecteur de ruptures couleur

La figure 3.9 illustre le fonctionnement de ce filtrage. Les quaternions purs q_1 , q_2 , q_3 et q_4 représentés par des vecteurs correspondent aux couleurs des pixels filtrés. On voit sur la figure 3.9 que si l'on a deux vecteurs opposés q_1 et q_2 par rapport à l'axe μ (dans l'exemple traité l'axe des gris), la somme du premier tourné de π ($\mu q_1 \bar{\mu}$) et du deuxième (q_2) aura une forte *Saturation*. Dans le cas contraire (q_3 et q_4), le résultat ($q_4 + \mu q_3 \bar{\mu}$) aura une faible *Saturation*. Nous illustrons un exemple d'utilisation sur la figure 3.10. Nous voyons apparaître sur cette image des pixels colorés au niveau des frontières de l'image, alors que les zones homogènes sont associées à des pixels gris.

Toutefois, malgré ses performances intéressantes, l'opérateur de Sangwine souffre entre autre d'une variabilité suivant le sens de parcours lors du calcul du filtrage. Nous pouvons voir sur la figure 3.11 que les ruptures liées au chapeau de Lenna ne se traduisent pas avec la même couleur dans les images résultats (a) ou (b). Pourtant, dans ces deux cas, le filtre est composé de filtres horizontaux. La différence est liée au sens d'application des filtres h_L et h_R dans la convolution de l'équation 3.52. En effet, en inversant l'ordre de leurs applications dans le schéma de convolution, les vecteurs couleur de comparaison produits sont différents. De plus, il nous a semblé que les informations couleur ne sont pas totalement exploitées dans cette première formulation. C'est pourquoi nous avons mené différents travaux proposant de généraliser l'opérateur de Sangwine comme notamment dans la thèse de P. Denis [62].

3.2.3 Extension de l'opérateur de Sangwine par des opérations géométriques

Si nous reprenons la formulation quaternionique d'origine, nous avons pour chaque pixel q_1 un vecteur de comparaison avec son voisin q_2 , et nous avons constaté que plus la différence de couleur est grande entre

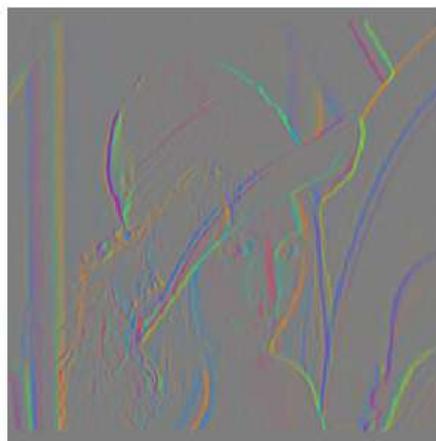
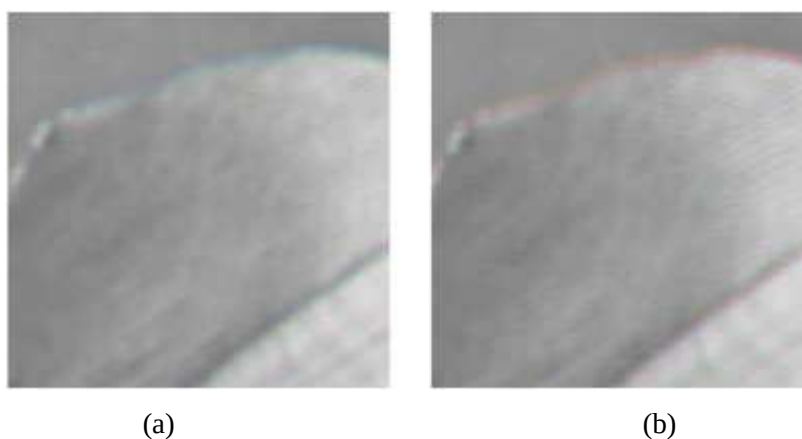


FIGURE 3.10 – Filtrage de Sangwine sur l'image Lenna

FIGURE 3.11 – Mesure des ruptures de Sangwine sur l'image Lenna : (a) zoom sur les contours du chapeau ; (b) zoom identique mais avec l'image traitée par un filtrage dans lequel h_R et h_L ont été permutés.

pixels voisins et plus ce vecteur de comparaison est éloigné de l'axe des niveaux de gris.

Dans le cadre de la thèse de F. Petit [179], des travaux ont été menés pour tenter de donner un sens à la couleur de ce vecteur de comparaison. Cette thèse s'est déroulée dans l'équipe Icones en collaboration avec le laboratoire L3I de La Rochelle et avait pour objectif de restaurer puis d'analyser des images sous-marines (avec notamment une application autour de la segmentation des poissons). Ce travail illustre bien l'intérêt de l'utilisation des quaternions car, en résumant rapidement, on peut dire que les images sous-marines se caractérisent par un axe des gris modifié tendant vers le bleu. Avec le formalisme introduit précédemment, après identification de ce nouvel axe, l'écriture des différentes opérations comme les filtrages restent identiques, puisque seul le paramétrage de μ_{gris} est modifié. C'est pourquoi j'ai collaboré à l'encadrement dans le cadre de ce travail de thèse pour la définition de différentes opérations reposant sur les quaternions. Une stratégie utilisant la transition de couleur après filtrage quaternionique a été aussi proposée pour sélectionner les contours liés à certaines informations colorimétriques. Nous reprenons sur la figure 3.12 les expérimentations menées par F. Petit, illustrant la teinte des coefficients après calcul du détecteur de Sangwine. La figure 3.13 présente les résultats de la détection de contours suivant un critère de teinte. Ces techniques sont détaillées dans [181, 180].

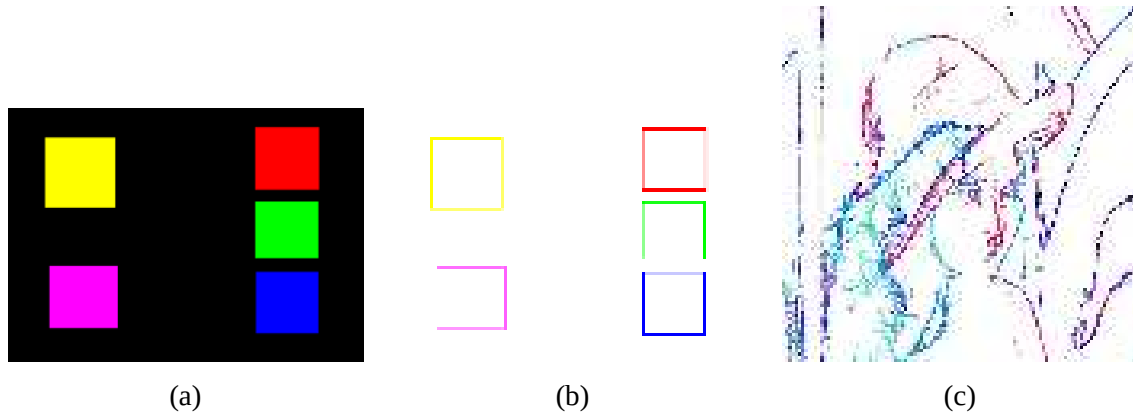


FIGURE 3.12 – Analyse de la couleur des contours après le détecteur de Sangwine (a) Image de test originale avec des patches couleur et un fond noir, (b) couleur des contours, (c) Couleur des contours pour l'image Lena



FIGURE 3.13 – Détecteur de contours de couleurs spécifiques sur l'image Pills avec une détection (a) jaune, (b) orange.

Dans le cadre de la thèse de P. Denis, nous avons ensuite proposé de donner un sens à la distance entre le vecteur de comparaison et l'axe des niveaux de gris (correspondant donc à la saturation de ce vecteur de comparaison), que nous noterons q_{dist} . En effet cette distance nous donne bien un renseignement sur la présence ou non d'une rupture dans l'information couleur de l'image. Plus la saturation est importante, plus la rupture couleur dans l'image est significative. Cette distance peut être calculée au moyen d'opérations quaternioniques élémentaires car elle est la norme de la réjection du vecteur de comparaison par rapport à μ [63] :

$$q_{dist} = \frac{1}{2}(q_{sang} + \mu q_{sang}\mu) \quad (3.53)$$

avec q_{sang} le résultat du filtrage suivant les masques proposés par Sangwine.

La méthode de filtrage de Sangwine donne deux vecteurs différents ($q_1 + \mu q_2 \bar{\mu}$ et $q_2 + \mu q_1 \bar{\mu}$) pour comparer les couleurs portées par les vecteurs q_1 et q_2 . Les deux distances S_1 et S_2 calculées avec notre approche sont bien sûr identiques. Ensuite cette opération de filtrage peut être appliquée en utilisant des filtres horizontaux, verticaux et dans les deux directions diagonales. On peut alors combiner ces différentes distances, calculées sur la saturation obtenue pour chacune de ces directions, pour obtenir une notion de module de gradient couleur (comme lors d'une approche classique). La procédure de fusion sera définie suivant l'application (norme L_2 par exemple).

La figure 3.14 montre les résultats de cette approche avec, pour chaque ligne, l'image originale suivie du calcul de la norme du gradient couleur, et enfin une carte de contours obtenue par seuillage empirique de

cette norme L_2 du gradient.

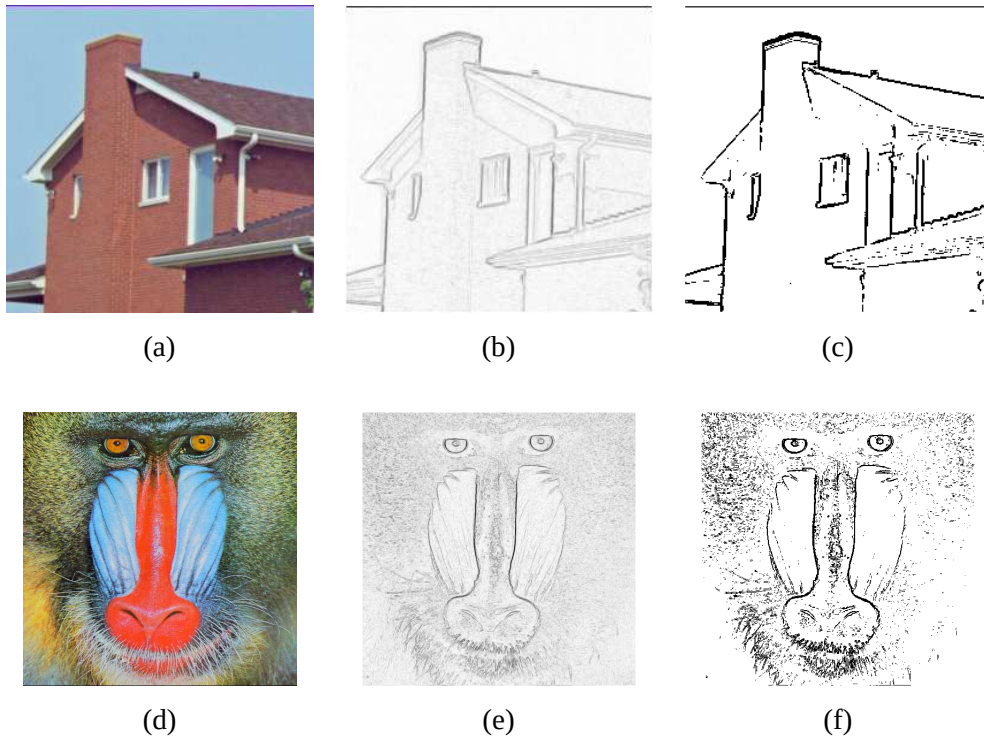


FIGURE 3.14 – De gauche à droite : images originales, norme du gradient couleur et carte des contours

Cependant, l'inconvénient de cette méthode réside dans le fait que ce gradient n'est calculé que sur une mesure de saturation. Par exemple lorsque l'information d'un pixel est uniquement de type achromatique, la saturation est nulle et donc la norme du gradient associé en ce point sera elle aussi nulle.

Nous avons alors proposé d'utiliser les algèbres géométriques afin d'apporter une description des couleurs par rapport à l'axe achromatique des niveaux de gris. Cette description géométrique va se faire à travers le produit géométrique suivant $f[m, n]\vartheta$ avec ϑ l'axe des gris. Ce produit, nous le rappelons, se décompose en une somme de deux parties :

- $f[m, n] \wedge \vartheta$ le bivecteur représenté par le plan porté par les vecteurs couleur $f[m, n]$ et ϑ . Lorsque la norme de cette partie est nulle, le vecteur couleur analysé est colinéaire à ϑ c'est à dire qu'il est porté par l'axe des niveaux de gris.
- $f[m, n] \cdot \vartheta$, est un scalaire correspondant à la projection du vecteur couleur $f[m, n]$ sur l'axe ϑ , autrement dit, son intensité.

Le processus de calcul de ce nouvel opérateur couleur va alors être réalisé à travers le filtrage introduit ci-dessous qui comprend à la fois le filtrage selon l'approche de Sangwine sur sa partie vectorielle, l'intensité sur la partie scalaire et la dimension chromatique sur la partie bivectorielle :

$$g[m, n] = \sum_{x=-1}^1 \sum_{y=-1}^1 v[x, y] f[m-x, n-y] u[x, y] \quad (3.54)$$

avec

$$v = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ \vartheta & \vartheta & \vartheta \end{pmatrix} \quad \text{et} \quad u = \begin{pmatrix} 1 & 1 & 1 \\ 0 & \vartheta & 0 \\ \vartheta^{-1} & \vartheta^{-1} & \vartheta^{-1} \end{pmatrix}$$

où ϑ est un vecteur non nul de $\mathbb{R}^3$. Pour mémoire, nous considérons l'image couleur comme une application f comme à valeurs dans la partie vectorielle de $\mathcal{G}_3$:

$$f[m, n] = f_r[m, n]e_1 + f_g[m, n]e_2 + f_b[m, n]e_3$$

où (e_1, e_2, e_3) est une base orthonormée de $\mathbb{R}^3$.

Notons que nous introduisons la valeur unité et ϑ dans la partie centrale du masque, cela permet ainsi d'introduire dans le filtrage le produit $f(m, n)\vartheta$. Le résultat de la convolution s'écrit donc (les produits sont effectués dans $\mathcal{G}_3$) :

$$\begin{aligned} g[m, n] &= \{[\vartheta f[m+1, n+1]\vartheta^{-1} + f[m+1, n-1]] + [\vartheta f[m, n+1]\vartheta^{-1} + f[m, n-1]] \\ &+ [\vartheta f[m-1, n+1]\vartheta^{-1} + f[m-1, n-1]]\} + f[m, n]\vartheta \end{aligned} \quad (3.55)$$

avec $f[m, n]\vartheta = f[m, n] \cdot \vartheta + f[m, n] \wedge \vartheta$. Le résultat est la somme d'un vecteur, d'un scalaire et d'un bivecteur. Chacune de ces parties contient une information spécifique. Nous avons vu que $\vartheta\vartheta\vartheta^{-1}$ est le symétrique du vecteur v par rapport au vecteur ϑ . On en déduit que si le masque parcourt une zone de l'image où la couleur varie peu selon la direction verticale, la partie vectorielle du résultat de la convolution est colinéaire au vecteur ϑ et l'on obtient ainsi un critère d'homogénéité. Cette partie vectorielle correspond au filtre de Sangwine. La partie bivectorielle permet de savoir si la couleur du pixel central est colinéaire à la couleur codée par ϑ et la partie scalaire mesure la différence de norme entre ces deux couleurs.

Nous proposons d'illustrer pas à pas notre filtrage sur une image synthétique contenant différents éléments, à savoir des parties achromatiques, des ruptures franches de teinte et des ruptures de saturation (cf. figure 3.15a). Nous prenons le cas le plus classique, à savoir que ϑ correspond à l'axe des gris.

Dans ce cas, l'approche complète est, pour chaque coordonnée :

- Calcul de la norme de la partie bivectorielle. Cette norme permet de décider si l'information en chaque pixel est de type achromatique. Nous voyons très bien dans la figure 3.15b qu'il y a deux zones où la norme est quasi nulle. Ces deux zones correspondent aux deux rectangles au centre de l'image qui ne présentent pas d'information couleur.
- Construction d'une image binaire (appelée masque) par seuillage de cette norme. Ce seuillage permet de séparer les zones de norme très faible des autres zones de l'image où les informations de chromatisme prédominent sur l'information d'intensité.
- Filtrage de la partie scalaire qui représente l'information d'intensité de l'image, par un ensemble de filtres de Prewitt dans chacune des directions horizontale, verticale et diagonales afin d'obtenir un gradient d'intensité. Ensuite, on utilise le masque défini à l'étape précédente pour ne conserver que l'information achromatique. Le résultat de cette étape est illustré par la figure 3.15c.
- Fusion du gradient d'intensité sur les zones achromatiques de l'image avec le gradient de saturation. Pour cette fusion, la valeur maximale entre les deux gradients peut être une solution.

La figure 3.15d illustre les ruptures mises en évidence par cette méthode. Comparées au gradient précédent de la figure 3.14, les différentes ruptures sont mises en valeur aussi efficacement. Cependant là où la

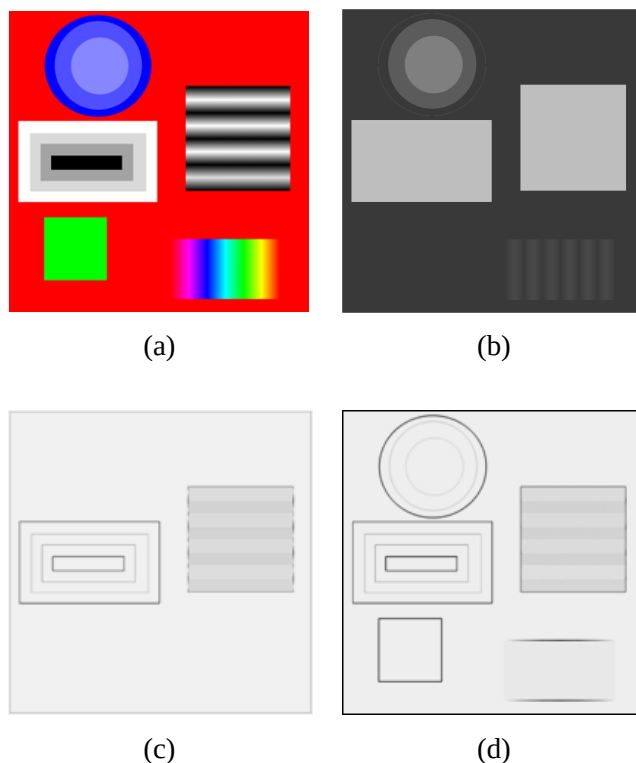


FIGURE 3.15 – (a) Image synthétique présentant différentes ruptures : achromatique, de saturation et de teinte ; (b) Norme de la partie bivectorielle (valeur ramenée entre 0 et 255 et inversée) ; (c) Norme du gradient d'intensité, calculé suivant les masques de Prewitt et masqué grâce à la norme de la partie bivectorielle ; (d) Gradient final.

méthode classique échoue dans l'extraction des contours achromatiques, cette technique permet de résoudre le problème comme nous le voyons par la mise en évidence de contours à l'intérieur du rectangle du centre gauche.

La figure 3.16 illustre un autre exemple de détection des ruptures. De nouveau, nous observons l'efficacité de la méthode, puisque quelque soit leur nature (chromatique ou achromatique) les différentes ruptures apparaissent dans l'image résultat.



FIGURE 3.16 – (a) Image originale ; (b) Gradient final.

Enfin, la figure 3.17 présente une comparaison avec des méthodes classiques. Nous constatons que, étant donné sa construction, le détecteur basé sur les algèbres géométriques permet de faire une synthèse des

différents résultats.

Notre objectif, dans cette présentation, n'est pas bien sûr de faire une comparaison exhaustive, puisque l'outil n'a pas été défini pour être optimal et peut être modifié à l'envie selon les informations que l'on souhaite mettre en évidence. Nous avons simplement introduit pas à pas une méthode de traitement d'images couleur qui utilise l'algèbre géométrique à travers la manipulation géométrique des couleurs au moyen d'opérations algébriques [63]. Si dans la construction actuelle la comparaison a été faite par rapport à l'axe des niveaux de gris, il est bien sûr tout à fait possible d'envisager une comparaison des pixels avec d'autres vecteurs couleurs pour définir de nouvelles opérations réellement colorimétriques. C'est ce qui a été fait dans le cadre des images sous-marines.

Nous développons actuellement, à partir de ce filtrage et en ajoutant la dimension teinte dans le calcul du gradient une procédure de paramétrage d'un snake couleur. Nous présentons sur la figure 3.18 un exemple de segmentation par un snake couleur paramétré suivant les valeurs d'un gradient basé algèbre géométrique.

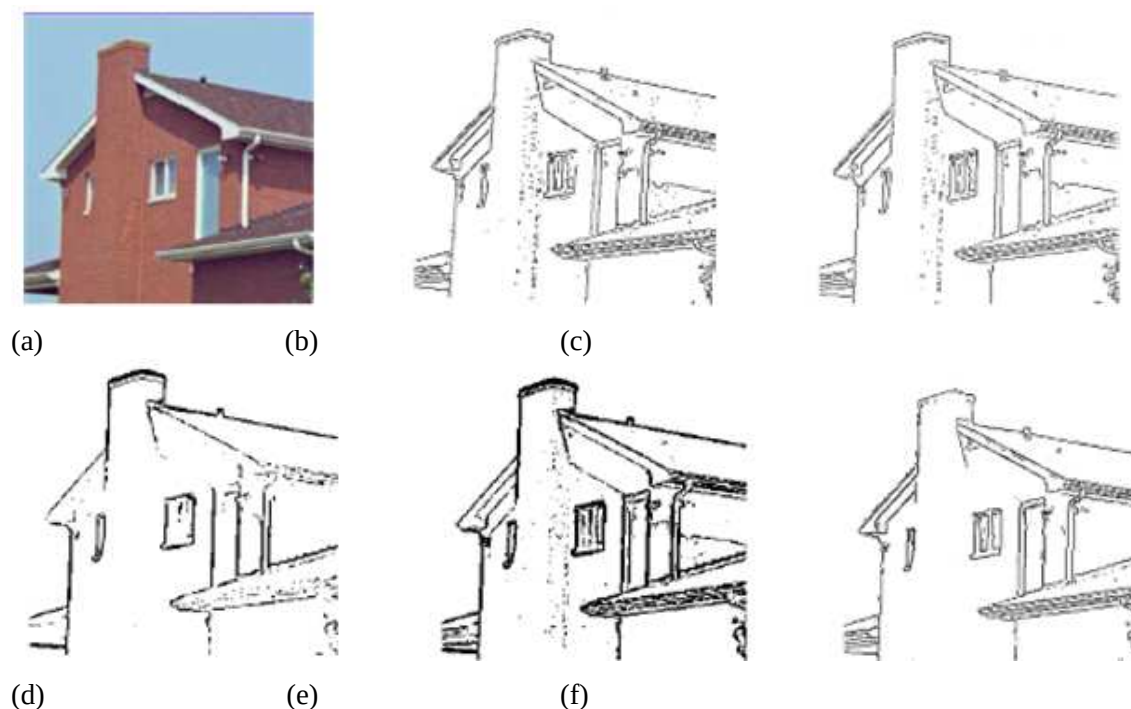


FIGURE 3.17 – (a) image originale (b) méthode marginale (c) Di Zenzo (d) gradient de saturation (e) gradient AG final (f) Carron

Nous proposons maintenant d'analyser un autre axe de généralisation à la couleur des outils classiques de traitement d'image, à savoir la transformée de Fourier. En effet, ces dernières années de nombreux travaux ont proposé de généraliser l'analyse fréquentielle aux images couleur avec des stratégies à base de quaternion ou d'algèbre de Clifford [20, 199, 64]. Cependant, comme nous allons le voir l'interprétation de ces représentations et leur utilisation sont des problèmes toujours ouverts.

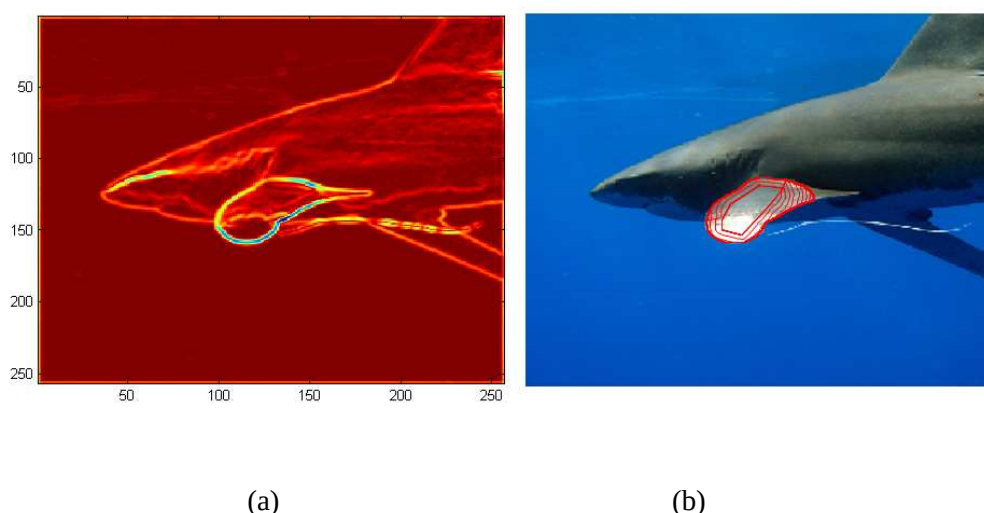


FIGURE 3.18 – *Segmentation par Snake couleur basé sur les algèbres géométriques : (a) valeur du gradient (b) différentes itérations du Snake*

3.3 Les transformées de Fourier couleur

La définition et la compréhension d'une Transformée de Fourier pour les images couleur a été, à l'origine, l'un des objectifs de l'utilisation de ces nouveaux outils mathématiques et a constitué un axe prépondérant de recherche que j'ai mené ces dernières années. Cependant, comme nous le constaterons, les résultats ne sont pas forcément à la hauteur des attentes.

Avant d'étudier l'extension de Fourier avec le formalisme des algèbres, il est intéressant d'analyser dans la littérature comment les chercheurs ont tenté de traduire l'information fréquentielle dans le domaine couleur.

3.3.1 Analyse spatio-chromatique d'images couleur

Nous allons tout d'abord revoir comment des variations spatiales de couleur peuvent être détectées dans le domaine fréquentiel couleur. Pour cela, nous nous appuyons sur les travaux de McCabe et al. [155] pour décrire les contrastes de couleur. Dans ces premières approches, la notion d'espace fréquentiel couleur est définie à partir du formalisme des nombres complexes.

3.3.1.1 La transformée de Fourier spatio-chromatique

Contrairement aux approches dites marginales qui consistent à traiter les différents canaux séparément, la transformée doit prendre en compte les données vectorielles d'une image couleur dans leur ensemble. Les premiers travaux dans cette voie ont proposé de se placer dans un espace couleur qui, idéalement, est indépendant de l'intensité, ceci afin de pouvoir effectuer des opérations portant uniquement sur les composantes chromatiques de l'image. On peut, pour cela, utiliser par exemple une représentation par des nombres complexes : les parties réelle et imaginaire codent les deux composantes chromatiques d'un pixel.

McCabe et Caelli [155] choisissent ainsi de représenter les composantes chromatiques dans l'espace YUV d'un pixel de coordonnées (m, n) par un nombre complexe $U[m, n] + iV[m, n]$. On peut dès lors calculer

les composantes fréquentielles chromatiques avec un calcul classique de transformée de Fourier complexe :

$$F_{cr}[o, p] = \frac{1}{\sqrt{N}\sqrt{M}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (U[m, n] + iV[m, n]) e^{\frac{-2i\pi(mo+np)}{N}} \quad (3.56)$$

Le couple (o, p) correspond aux coordonnées fréquentielles de la transformée de Fourier de l'image. Il est à noter que la propriété de symétrie hermitienne n'est plus valable dans ce contexte puisque les données de départ sont complexes. D'autres travaux ont proposé d'utiliser d'autres espaces couleur, l'idée étant toujours d'avoir une partie achromatique et deux composantes chromatiques que l'on code dans un complexe.

Il existe d'autres approches utilisant par exemple la représentation polaire en lien avec l'espace couleur LHS [79]. La transformée de Fourier est alors calculée à partir d'une fonction dont les valeurs en chaque position sont exprimées sous la forme polaire :

$$z[m, n] = f_S[m, n] e^{(if_H[m, n])} \quad (3.57)$$

où f_S est la saturation et f_H est la teinte du pixel à la coordonnée $[m, n]$.

Ces approches ont été utilisées avec succès notamment en ce qui concerne le calcul d'intercorrélation [79]. Plus récemment au sein de l'équipe Icones, un travail a été mené autour de la généralisation des modèles paramétriques, et notamment de la densité spectrale associée, aux images couleur [191, 190]. Cette généralisation a elle aussi utilisé le modèle complexe.

Cependant, il n'en demeure pas moins qu'il est difficile à partir de ces représentations de donner un sens au spectre résultant de la transformée de Fourier complexe. En effet, il n'est pas immédiat de traduire le coefficient fréquentiel F en termes de comportement spatial couleur, il n'existe d'ailleurs que très peu de travaux sur ce sujet. L'une des pistes, reprise par la suite, dans le cadre des quaternions est la notion de chemin couleur.

3.3.1.2 Notion de chemin couleur

Une explication de l'espace fréquentiel obtenu à partir de la transformée de Fourier complexe est proposée par McCabe et all. [155] en utilisant la notion de chemin couleur (figure 3.19). Cette approche se propose d'interpréter les coefficients de Fourier à travers un diagramme de phase. Nous reprenons le principe des courbes de Lissajous avec une représentation paramétrique de définition générale :

$$x(t) = a \sin(t) \text{ et } y(t) = b \sin(nt + \theta). \quad (3.58)$$

Dans le cas de la transformée de Fourier complexe, le coefficient de Fourier peut s'écrire sous sa forme polaire : $k.e^{i\theta}$. Si l'on raisonne dans le domaine spatial et donc après calcul d'une transformée de Fourier inverse, les coefficients $F_{cr}[o, p]$ du spectre chromatique interagissent avec les fonctions exponentielles de la transformée de Fourier inverse. Prenons l'exemple le plus simple d'un spectre ne contenant qu'un coefficient non nul $F_{cr}[o, p] = (k.e^{i\theta})\delta_{\alpha_1, \alpha_2}[o, p]$ avec :

$$\delta_{\alpha_1, \alpha_2}[o, p] = \begin{cases} 1 & \text{si } o = \alpha_1; p = \alpha_2 \\ 0 & \text{sinon} \end{cases} \quad (3.59)$$

Dans ce cas, si l'on calcule la TF inverse, on obtient :

$$\begin{aligned}
f[m, n] &= \sum_o \sum_p e^{[i(\frac{om}{M} + \frac{pn}{N})]} k e^{[i\theta] \delta_{\alpha_1, \alpha_2}[o, p]} \\
&= k e^{[i(\frac{\alpha_1 m}{M} + \frac{\alpha_2 n}{N}) + \theta]} \\
&= k \cos((\frac{\alpha_1 m}{M} + \frac{\alpha_2 n}{N}) + \theta) + i k \sin((\frac{\alpha_1 m}{M} + \frac{\alpha_2 n}{N}) + \theta)
\end{aligned}$$

Si nous reprenons l'équation (3.58), nous sommes dans le cas de la courbe de Lissajous avec $a = b$ et $n = 1$ donc la courbe est une ellipse dans le plan chromatique. Dans ce cas, k est le rayon de balayage du chemin et va donc représenter l'amplitude de la variation de couleur dans le domaine spatial, bien sûr le couple (α_1, α_2) va représenter la vitesse d'oscillation et nous constatons que θ est l'angle initial de la variation dans les composantes chromatiques (il est appelé angle couleur dans l'article de McCabe).

L'interprétation de ce chemin couleur va bien sûr dépendre de l'espace couleur de l'image complexe analysée. Si l'on reprend l'espace YUV, après transformée de Fourier inverse, le résultat est une matrice complexe décrivant des variations spatiales suivant les composantes U et V de l'image.

Suivant cette approche, la notion de chemins couleur va alors se construire par la juxtaposition du diagramme de phase et de la variation couleur dans le domaine spatial.

Un premier exemple est donné avec la figure 3.19. Pour obtenir de telles variations de couleur, on initialise les composantes fréquentielles comme suit ($\alpha_1 > 0, \alpha_2 > 0$) :

- $F_{cr}[o, p] = k \cdot \delta_{\alpha_1, \alpha_2}[o, p]$ pour le chemin c_1 et
- $F_{cr}[o, p] = k \cdot \delta_{-\alpha_1, -\alpha_2}[o, p]$ pour le chemin c_2 .

On remarque que les deux chemins c_1 et c_2 , associés à des coordonnées fréquentielles différentes ont le même angle $\theta = 0$ et la même amplitude $m = k$. Afin de bien comprendre le lien entre le chemin et le contenu de l'image, il faut se référer à l'organisation du plan UV que nous illustrons sur la figure 3.3.1.2. Par exemple, en ce qui concerne le chemin c_1 nous constatons une oscillation débutant par le bleu ($\theta = 0$) et ensuite passant par le vert puis le rouge, suivant ainsi le chemin schématisé par c_1 .

Si l'on combine les deux chemins, on obtient donc le chemin c_3 lié au spectre $F_{cr}[o, p] = k \cdot (\delta_{\alpha_1, \alpha_2}[o, p] + \delta_{-\alpha_1, -\alpha_2}[o, p])$. La particularité du troisième chemin est qu'il correspond en terme spatial à une variation de couleur qui suit uniquement l'axe u associé aux réels, et qui va donc correspondre à des oscillations variant du bleu au vert-marron.

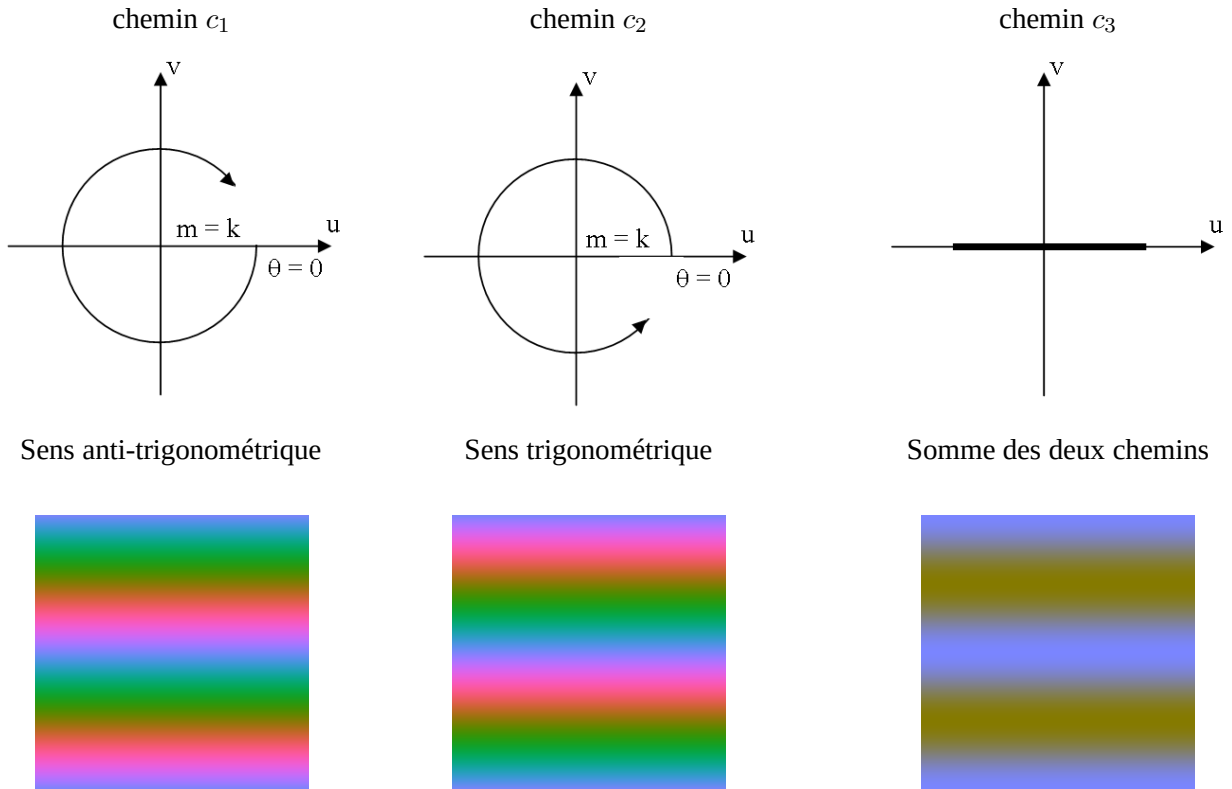
De la même façon, une variation de couleur qui suit l'axe des imaginaires (et donc l'axe V) est possible en initialisant le spectre par

$$F_{cr}[o, p] = i \cdot k \cdot (\delta_{\alpha_1, \alpha_2}[o, p] + \delta_{-\alpha_1, -\alpha_2}[o, p])$$

qui est la somme des deux chemins c_4 et c_5 d'angle $\theta = \frac{\pi}{2}$ et d'amplitude $m = k$ caractérisés par les quadruplets $i \cdot k \cdot \delta_{\alpha_1, \alpha_2}$ et $i \cdot k \cdot \delta_{-\alpha_1, -\alpha_2}$. Cette fois ci, le chemin couleur est associé aux variations spatiales sur l'axe v autrement dit les oppositions de couleur entre le rouge et le vert.

On peut bien sûr envisager de faire des combinaisons linéaires des chemins couleur de base, comme illustré dans la figure 3.21 Le résultat de cette somme est le chemin elliptique c_{12} illustré à droite. La forme de l'ellipse est définie par les amplitudes m_1 et m_2 et l'orientation de la couleur par les angles θ_1 et θ_2 .

Suivant ce concept l'idée est donc de décrire l'information spectrale obtenue par la transformée de Fourier au moyen de la notion de chemin couleur. Chaque chemin couleur correspond à un point de U et V dans le spectre. Il apparaît donc possible d'analyser des images couleur au moyen de l'information spectrale obtenue à partir de cette lecture de transformée de Fourier. Cependant les variations passent par de nombreuses



initialisation avec $k \cdot \delta_{\alpha_1, \alpha_2}[o, p]$ initialisation avec $k \cdot \delta_{-\alpha_1, -\alpha_2}[o, p]$ initialisation avec $k \cdot (\delta_{\alpha_1, \alpha_2} + \delta_{-\alpha_1, -\alpha_2})$

FIGURE 3.19 – Le chemin couleur c_3 est obtenu à partir de la somme des chemins couleur c_1 et c_2 , il est décrit uniquement sur l'axe des réels.

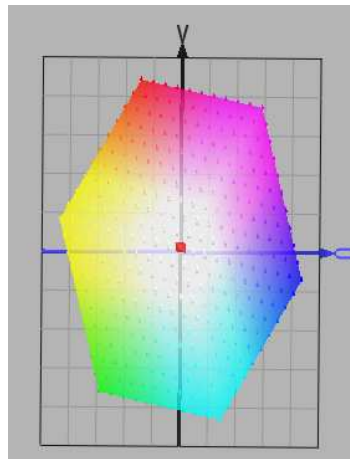
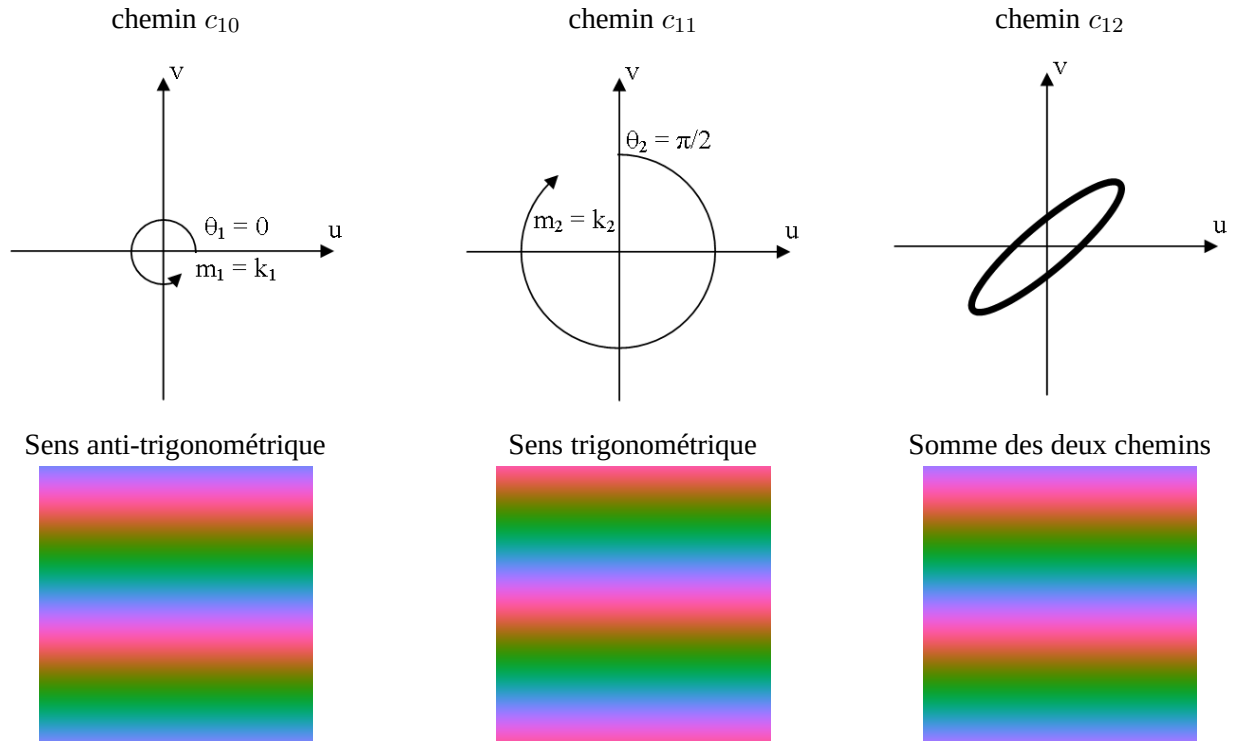


FIGURE 3.20 – Variation couleur dans le plan UV .

couleurs en parcourant chaque chemin, il peut alors paraître difficile d'isoler une couleur plutôt qu'une autre dans l'analyse obtenue. Pour réussir, il faudra décomposer le spectre afin d'obtenir tous les chemins couleurs élémentaires utilisés pour obtenir telle ou telle variation. Cette démarche nous semble difficilement applicable sur des images réelles. Toutefois, à notre connaissance, elle est la seule à tenter de donner un sens au spectre chromatique. De plus, comme nous allons le voir, cette représentation sera reprise par Sangwine

initialisation avec $m_1.\text{delta}_{-\alpha_1, -\alpha_2}$ initialisation avec $i.m_2.\text{delta}_{-\alpha_1, -\alpha_2}$

combinaison des deux

FIGURE 3.21 – Le chemin couleur c_{12} est obtenu à partir de la somme des chemins couleur c_{10} et c_{11} . Ici même si les coordonnées fréquentielles (o, p) sont identiques, les deux chemins c_{10} et c_{11} ont des amplitudes (m_1 et m_2) et des phases (θ_1 et θ_2) différentes.

dans le contexte Quaternionique.

Pour conclure sur cette transformée de Fourier, nous proposons de faire lien entre cette transformée chromatique proposée il y a donc maintenant quelques années et la modélisation moderne des images couleur dans l'algèbre géométrique que nous développons en début de ce chapitre (à savoir la modélisation de chaque pixel suivant un 1-vecteur $p = f_r[m, n]e_1 + f_g[m, n]e_2 + f_b[m, n]e_3$). La caractérisation du plan chromatique UV dans le cadre de l'algèbre géométrique est représenté sur la figure 3.3.1.2. Le plan est formé à partir des deux vecteurs U et V , avec par exemple

$$U = -0.6 * e_1 - 1.14 * e_2 + 0.75 * e_3 \quad (3.60)$$

défini par la matrice de passage entre l'espace RGB et l'espace YUV. Il correspond donc au bivecteur $U \wedge V$.

Suivant cette illustration, on constate que la transformée de Fourier chromatique peut s'interpréter comme une transformée de Fourier complexe calculée sur la composante $m||_{U \wedge V}$. La définition de ce plan (et donc du bivecteur) peut évoluer selon les applications. Cette interprétation de la transformée de Fourier chromatique montre la souplesse de la modélisation d'une image couleur suivant l'algèbre géométrique et son caractère unificateur des différentes approches. Par exemple, nous présentons sur la figure 3.23, une expérimentation que j'avais développée dans le cadre du Workshop sur les algèbres géométriques qui a eu lieu à La Rochelle en 2009 [42]. Cette expérimentation est partie du principe que, puisque nous avons construit une représentation fréquentielle chromatique adaptative, nous pouvons définir une transformée en ondelettes

couleur à partir simplement d'un banc de filtres classiques qui s'adapte au contenu colorimétrique. Suivant cette idée, la figure 3.23 présente la décomposition en ondelettes de $f||_{U \wedge V}$, avec $U \wedge V$ défini perpendiculairement à l'axe d'inertie maximum du nuage de points associé à l'image couleur. Nous pouvons voir que le choix d'une direction adaptée à l'image permet une représentation fine des détails couleur de l'image (puisque l'analyse est faite suivant la direction présentant le plus de dispersion).

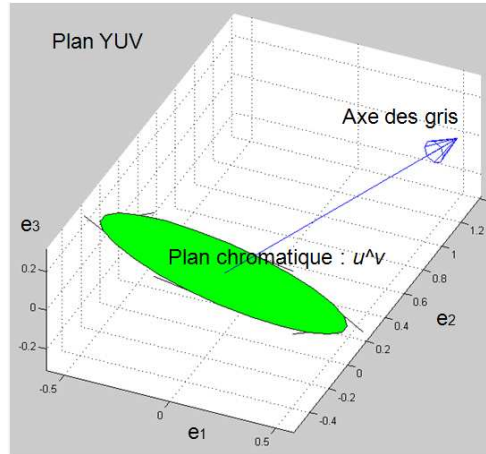


FIGURE 3.22 – Modélisation du plan UV dans le cadre de l'algèbre géométrique.

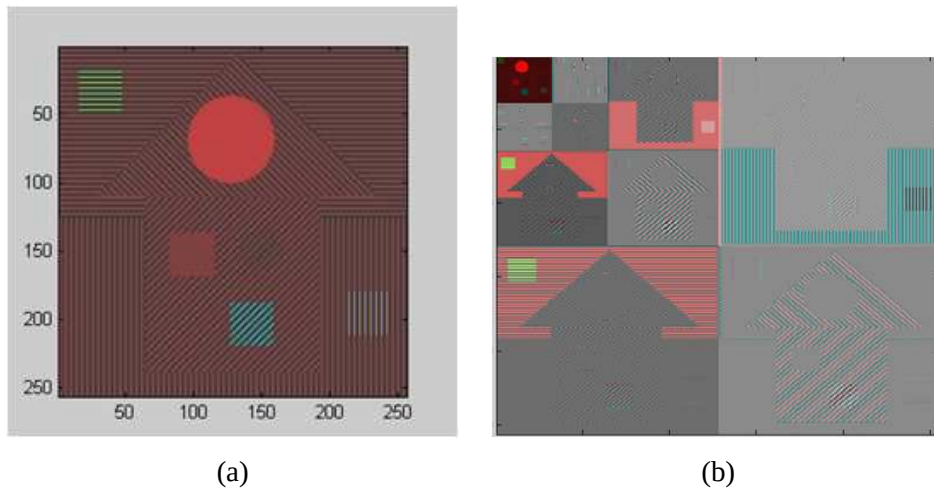


FIGURE 3.23 – Banc de filtres chromatiques. (a) Image de test (b) décomposition en ondelettes suivant un certain plan $u \wedge v$

Nous proposons maintenant d'étudier la généralisation de la transformée de Fourier dans l'algèbre des Quaternions.

3.3.2 Transformées de Fourier quaternioniques

Les transformées de Fourier quaternioniques (ou aussi TFQ) ont été introduites dans le cadre des images couleur par Ell et Sangwine en 1992 [83] afin de généraliser la transformée de Fourier complexe pour les images couleur. Dans le cadre des quaternions, il existe plusieurs versions, avec des objectifs différents, des

transformées de Fourier différentes notamment parce que le produit quaternionique n'est pas commutatif. Alors que Bülow utilise une TFQ pour analyser les structures 2D des images en niveaux de gris [35], Ell et Sangwine sont parmi les premiers à associer les images couleur aux transformées de Fourier quaternioniques [83]. Nous proposons de décrire maintenant ces différentes transformées.

3.3.2.1 Définition

La première version est celle proposée par Todd Ell [79] dans sa thèse et utilisée par Sangwine en 1996 dans [201]. Soit f une fonction discrète à deux variables m et n et à valeurs quaternioniques, la transformée de Fourier quaternionique (ou TFQ) telle qu'elle est définie est la fonction à valeurs quaternioniques définie par :

$$F_{côtés_1}[o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-2\mu_1 \pi \frac{om}{M}} f[m, n] e^{-2\mu_2 \pi \frac{pn}{N}} \quad (3.61)$$

où μ_1 et μ_2 des quaternions unitaires purs, choisis comme étant j et k dans la version de Ell.

Cette définition est une variante de celle exposée par Thomas Bülow dans sa thèse [35] pour analyser les images en niveaux de gris en mettant l'accent sur l'aspect spatial. Pour cela, dans la version de Bülow, la dimension horizontale de l'image est associée à l'imaginaire pur i tandis que la dimension verticale est liée à l'imaginaire j :

$$F_{ij}[o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-2i\pi \frac{om}{M}} f[m, n] e^{-2j\pi \frac{pn}{N}} \quad (3.62)$$

Cette TFQ peut être appliquée sur des signaux réels, complexes et quaternioniques. Dans le cadre des travaux de Bulow, elle est utilisée sur des signaux à valeurs réelles qui correspondent à l'information d'intensité des pixels en niveaux de gris. Elle permet alors notamment de faire apparaître des propriétés de symétries associées au spectre obtenu par cette transformée. Plus précisément, il est connu que l'on peut écrire n'importe quelle fonction 1-D comme la somme d'une fonction paire et d'une fonction impaire $f(x) = f_p(x) + f_i(x)$ et que ces deux parités se dissocient dans le domaine de Fourier. L'idée de Bülow est d'étendre ces notions de symétries au cas des signaux 2D avec pour but de définir une notion de signal analytique 2-D. Un signal 2D f se décompose en effet selon les parties paires et impaires le long des axes horizontal m et vertical n : $f(x, y) = f_{pp}(x, y) + f_{ip}(x, y) + f_{pi}(x, y) + f_{ii}(x, y)$ avec f_{pp} correspondant à la partie de f qui est paire le long de m et n , f_{ip} correspondant à la partie qui est impaire le long de m et paire le long de n , etc. Dans la TF complexe 2-D, ces différentes parités vont se mélanger dans les parties réelle et imaginaire.

Nous illustrons ci-dessous les symétries présentes dans la transformée de Fourier Quaternionique de Bülow pour un signal à valeurs réelles :

$$F^q[o, p] = F_{pp}^q[o, p] + iF_{ip}^q[o, p] + jF_{pi}^q[o, p] + kF_{ii}^q[o, p] \quad (3.63)$$

On a donc une caractérisation de toutes les symétries spatiales possibles en 2D séparées dans le spectre quaternionique. De plus, même si cette information n'est pas nécessaire à l'analyse d'un signal réel, avec un signal décrit sur un quaternion quelconque, les mêmes propriétés de symétries séparent les parties paires et impaires suivant m et n dans le spectre pour chaque composante du quaternion (partie réelle, première partie imaginaire, deuxième partie imaginaire et troisième partie imaginaire) de manière indépendante. Nous

reviendrons sur cette définition dans le prochain chapitre, dans lequel notre problématique sera justement la mise en valeur des directions d'analyse.

Dans le cadre de la couleur, Sangwine a proposé une variante aux transformées précédentes en ajoutant à l'expression de l'exponentielle, une direction caractérisée par un quaternion unitaire pur μ [202]. Le plus souvent, pour ne privilégier aucune couleur pour l'analyse spectrale, on prend le quaternion neutre $\mu = \mu_{gris} = (i + j + k)/\sqrt{3}$ qui correspond à l'axe achromatique de l'espace RVB. La définition est la suivante :

$$F_{gauche}[o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-2\mu\pi(\frac{om}{M} + \frac{pn}{N})} f[m, n] \quad (3.64)$$

Il est possible de placer l'exponentielle à droite de la fonction f , on obtient alors la définition d'une autre transformée de Fourier puisque le produit quaternionique n'est pas commutatif.

Nous proposons maintenant de caractériser et de donner un sens à cette nouvelle transformée couleur.

Dans la suite de ce chapitre, nous nous intéressons uniquement à la transformée de Fourier quaternionique à gauche (équation 3.64) qui est la plus utilisée pour analyser les fréquences des images couleur. Aussi lorsqu'on mentionnera une TFQ il s'agira de cette dernière. La version de Bülow sera utilisée dans le chapitre suivant.

3.3.2.2 Définition numérique de l'espace de Fourier quaternionique

Dans le cadre de la TF complexe, il est bien connu que si l'on veut reconstruire un signal à valeurs réelles à partir de données dans le domaine fréquentiel, il faut imposer certaines conditions sur ces données pour tenir compte de la propriété de symétrie hermitienne de la transformée de Fourier complexe. Il n'existait pas de travaux caractérisant les contraintes sur le domaine de Fourier quaternionique Couleur. Cela a fait l'objet d'une partie de la thèse de P. Denis [62]. Nous proposons de reformuler d'une manière compacte cette caractérisation.

Pour obtenir après une transformée de Fourier quaternionique inverse (TFQI) une fonction quaternionique pour laquelle la partie réelle est nulle, c'est à dire une image couleur, l'espace de Fourier quaternionique doit respecter certaines conditions. L'information spectrale est décrite par des quaternions a priori quelconques, elle est donc de la forme suivante :

$$F[o, p] = F_r[o, p] + F_i[o, p]i + F_j[o, p]j + F_k[o, p]k \quad (3.65)$$

en un point du spectre de coordonnées fréquentielles (o, p) , avec $o \in]-\frac{M}{2}; \frac{M}{2}]$ et $p \in]-\frac{N}{2}; \frac{N}{2}]$ et (M, n) étant la taille de l'image.

La TFQ inverse est donnée par :

$$f[m, n] = \frac{1}{\sqrt{MN}} \sum_o \sum_p e^{2\mu\pi(\frac{om}{M} + \frac{pn}{N})} F[o, p] \quad (3.66)$$

où $\mu = \mu_i i + \mu_j j + \mu_k k$ est le quaternion unitaire pur donnant la direction de l'analyse. Comme nous allons le voir, ce quaternion pur a une influence sur l'organisation des informations dans le domaine fréquentiel, c'est pourquoi dans la très grande majorité, cette direction est associée à l'axe des gris permettant de séparer l'information en une information fréquentielle de luminance et une information fréquentielle de chrominance.

Décrivons sous quelle condition la composante réelle f_r est nulle après TFQ inverse. En développant le calcul, nous obtenons la relation suivante :

$$\begin{aligned} f_r[m, n] &= \frac{1}{\sqrt{MN}} \sum_o \sum_p \left[\cos \left(2\pi \left(\frac{om}{M} + \frac{pn}{N} \right) \right) F_r[o, p] \right. \\ &\quad \left. - \sin \left(2\pi \left(\frac{om}{M} + \frac{pn}{N} \right) \right) (\mu_i F_i[o, p] + \mu_j F_j[o, p] + \mu_k F_k[o, p]) \right] \end{aligned} \quad (3.67)$$

On constate par la suite que pour obtenir une image couleur f (avec $f_r = 0$) il est nécessaire que $F_r[0, 0] = F_r[\frac{M}{2}, \frac{N}{2}] = 0$ et pour tout $o \in [1; \frac{M}{2} - 1]$ et $p \in [1; \frac{N}{2} - 1]$:

$$F_r[-o, -p] = -F_r[o, p] \quad (3.68)$$

Il faut de plus, pour tout $o \in [0; \frac{M}{2}]$ et $p \in [0; \frac{N}{2}]$:

$$\begin{aligned} F_i[-o, -p] &= F_i[o, p] \\ F_j[-o, -p] &= F_j[o, p] \\ F_k[-o, -p] &= F_k[o, p] \end{aligned} \quad (3.69)$$

Tout logiquement, on retrouve ainsi une propriété équivalente à celle déjà connue avec l'analyse de Fourier complexe, à savoir que le spectre quaternionique d'une image couleur contient l'analogue d'une symétrie anti-hermitienne. Etudions maintenant l'information portée par le coefficient spectral F .

3.3.2.3 Interprétation du spectre quaternionique

Nous avons proposé différentes pistes pour donner une interprétation de ce que représentait l'information décrite par les coefficients du spectre :

- Décrire le spectre quaternionique avec la représentation exponentielle (nous retrouvons une approche similaire dans [9]).
- Procéder à une analyse des atomes élémentaires [64].

Sangwine et al. proposent de décomposer chaque coefficient spectral en deux parties, l'une simplexe et l'autre perplexe, afin de séparer l'information de luminosité de celle de chrominance [205, 82]. Nous décrivons aussi cette approche.

Interprétation du spectre quaternionique avec la notation exponentielle. On peut essayer d'exprimer le spectre de Fourier en utilisant les notions de phase, de module et d'axe qui caractérisent la notation exponentielle d'un quaternion. Rappelons que pour une coordonnée fréquentielle donnée, le coefficient de Fourier appartient à $\mathbb{H}$, $F[o, p] = q_0 + iq_1 + jq_2 + kq_3$ alors il existe un réel φ et ν appartenant à $\mathbb{H}_0 \cap \mathbb{H}^1$ tels que :

$$F[o, p] = q = |q|e^{\nu\varphi} \quad (3.70)$$

Le quaternion ν est l'axe associé à $F[o, p]$, et le réel φ son angle.

Classiquement, dans le cadre de l'image, se pose le problème de la représentation graphique du spectre. Par la suite, on représente le logarithme du module de la transformée de Fourier quaternionique comme cela est utilisé dans les visualisations de spectres. Pour la phase et pour l'axe de la transformée de Fourier,

il faut pondérer ces informations par l'amplitude du module (pour des problèmes d'instabilité des valeurs numériques de phase et d'axe). La phase étant exprimée entre $-\pi/2$ et $\pi/2$, la manière la plus simple pour la représenter est de l'intégrer dans la composante de Teinte d'un espace couleur HSV. Enfin l'axe est quant à lui codé sur trois composantes couleur d'un espace RVB car c'est un quaternion pur. Décrivons maintenant ces trois informations.

Le module. Son interprétation ne présente pas de difficulté car elle est identique à celle que nous connaissons lorsque nous effectuons une transformée de Fourier complexe. Elle correspond bien à une description de l'énergie suivant les fréquences spatiales.

La phase et l'axe. Il existe quelques propositions dans la littérature [82, 62].

Notons tout d'abord que lorsque la direction de la TFQ est associée à un vecteur couleur perpendiculaire aux couleurs contenues dans l'image analysée, l'angle obtenu par la TFQ reste le même et correspond à la couleur verte. Ce vert correspond à un angle de $\pi/2$ qui est la valeur limite de la fonction $\arctan$, donc quand l'information de phase n'est pas pertinente. Par exemple si la direction de la TFQ est $\mu = i$ (donc associée au rouge) alors pour les images bleue, verte et cyan, images dans lesquelles aucun pixel n'est fonction du « rouge », l'angle apparaîtra avec une valeur de $\pi/2$.

Si nous choisissons la direction d'analyse de la transformée de Fourier suivant l'axe des niveaux de gris, il est possible d'avoir un peu plus d'information sur ces deux notions. Nous proposons de reprendre l'exemple construit par Ell et Sangwine[82] qui effectuent une transformée de Fourier quaternionique de direction μ_{gris} sur une image comportant un fond uni de couleur jaune saturé à la moitié de la valeur maximale, et un objet carré dont la couleur est opposée dans l'espace couleur HSV à celle du fond, soit le bleu, lui aussi à moitié saturé. Nous ajoutons une deuxième image qui est construite comme la première mais avec les couleurs inversées. Les résultats sont présentés sur la figure 3.24.

La vue du module nous montre une forme de lobe caractéristique d'une impulsion, donc elle est cohérente. Puisque la direction d'analyse est l'axe des gris, on observe qu'il n'y a pas de couleur verte associée à la phase. On remarque également les oscillations des valeurs de phase et d'angle suivant la direction attendue, à savoir une direction « Sud-Est » correspondant à l'axe sur lequel se situe le petit carré en haut à gauche. La deuxième figure montre qu'en échangeant les couleurs dans l'image d'origine on ne change pas le module, ni même l'orientation des oscillations de phase et d'axe comme attendu. De plus, on constate que la phase obtenue avec la deuxième image présente des oscillations complémentaires à celles obtenues avec la première image, ce qui est cohérent car les couleurs analysées par la TFQ sont complémentaires. On observe le même comportement entre les deux vues de l'axe. Nous avons donc une traduction cohérente du contenu colorimétrique dans ces trois informations. Cependant, lorsque l'on veut aller plus loin et extraire une information à partir des deux mesures de phase et d'axe (c'est un travail qui a été notamment mené durant la thèse de F. Petit [179]) les résultats sont décevants. Toutes ces constations restent fragiles et empiriques et, à notre connaissance, il n'existe pas une proposition satisfaisante dans la littérature.

Analyse par atomes élémentaires. Notre deuxième stratégie pour proposer une interprétation du spectre quaternionique consiste à raisonner en termes d'atomes élémentaires. Ceci permet d'illustrer simplement quelle fonction "couleur" se traduit dans le domaine fréquentiel par un unique coefficient non nul (en respectant les symétries exposées précédemment). En initialisant un point dans le domaine fréquentiel par une valeur, on veut interpréter le résultat obtenu dans le domaine spatial après transformée de Fourier inverse.

Nous nous intéressons tout d'abord à la composante réelle du spectre (initialisation de F_r). Afin de

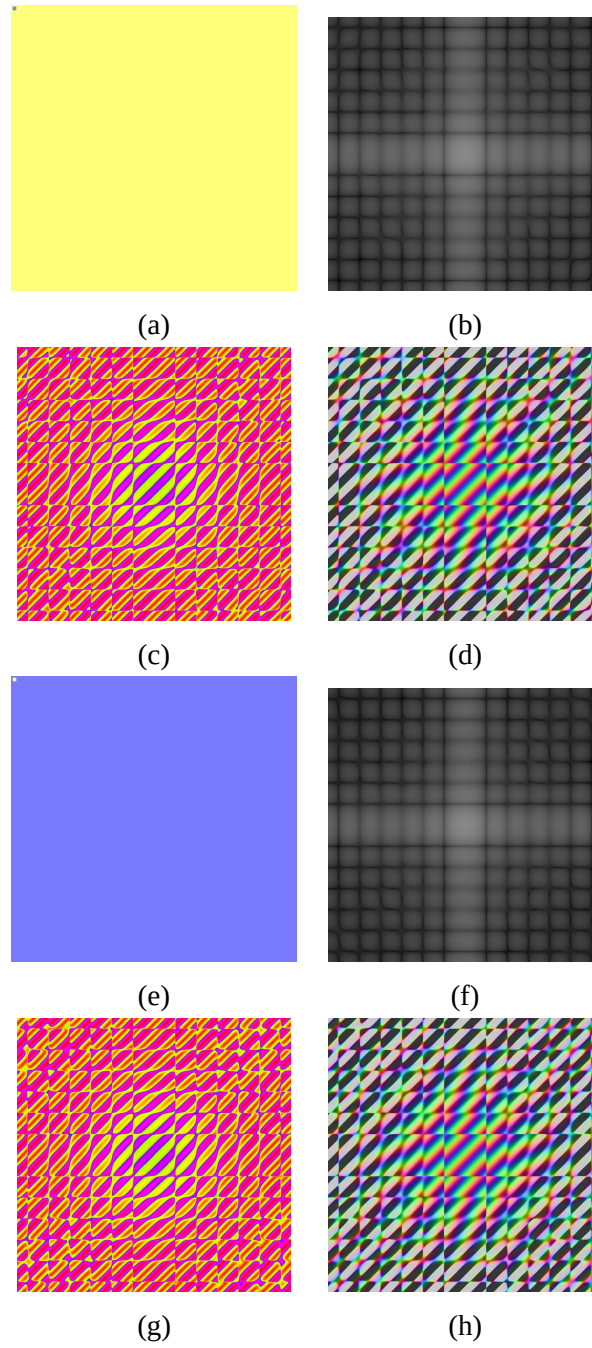


FIGURE 3.24 – Images originales (a) et (e) ; leur module (b) et (f) ; leur phase (c) et (g) et leur axe (d) et (h) par la TFQ

respecter les symétries du spectre couleur, un atome élémentaire est défini dans le domaine fréquentiel par : $F[o, p] = K_r \cdot \delta_{o_0, p_0}[o, p] - K_r \delta_{-o_0, -p_0}[o, p]$. Après un calcul simple de la TFQ inverse, on constate que cet atome correspond dans le domaine spatial à :

$$f[m, n] = 2\mu(K_r) \sin\left(2\pi\left(\frac{o_0 m}{M} + \frac{p_0 n}{N}\right)\right) \quad (3.71)$$

La partie réelle dans le domaine spatial est donc bien nulle puisque f est égale à une sinusoïde multipliée par un quaternion pur μ . On constate ensuite que l'information codée dans la partie réelle du spectre quaternionique est associée aux évolutions impaires (vitesse d'oscillation dépendant bien évidemment des coordonnées fréquentielles de l'initialisation) pondérées par la direction d'analyse μ , donc suivant les trois composantes couleur.

Nous pouvons faire le même calcul sur l'une des parties imaginaires e avec $e = i, j$ ou k : $F[o, p] = e \cdot (K_e \cdot \delta_{o_0, p_0}[o, p] + K_e \delta_{-o_0, -p_0}[o, p])$. On obtient :

$$f[m, n] = e \left[2(K_e) \cos\left(2\pi\left(\frac{o_0 m}{M} + \frac{p_0 n}{N}\right)\right) \right] \quad (3.72)$$

On constate de nouveau que la partie réelle dans le domaine spatial est nulle et que la partie associée à e capte les évolutions paires suivant la composante couleur associée à la partie complexe e et ceci quelque soit la direction d'analyse de la TFQ.

Pour résumer :

- L'atome associé à une composante imaginaire dans le spectre est une variation paire dans le domaine spatiale sur la couleur associée. Donc un Dirac dans le domaine fréquentiel sur une composante i donnera une variation de couleur suivant la composante i dans le domaine spatial. Il apparaît également que la direction μ n'a pas d'influence lors de l'insertion d'un dirac dans les composantes imaginaires de l'espace d'insertion fréquentiel.
- Une initialisation avec un Dirac positionné sur la composante réelle de l'espace fréquentiel, il en résulte une variation spatiale impaire de couleur suivant la direction donnée par la direction μ .

Bien sûr étant donné la linéarité de la TFQ, ces principes se généralisent en présence de plusieurs atomes, auquel cas les couleurs obtenues respectent la loi de la synthèse additive.

Nous illustrons sur la figure 3.25 l'initialisation du spectre avec des images en couleur dans l'espace RVB .

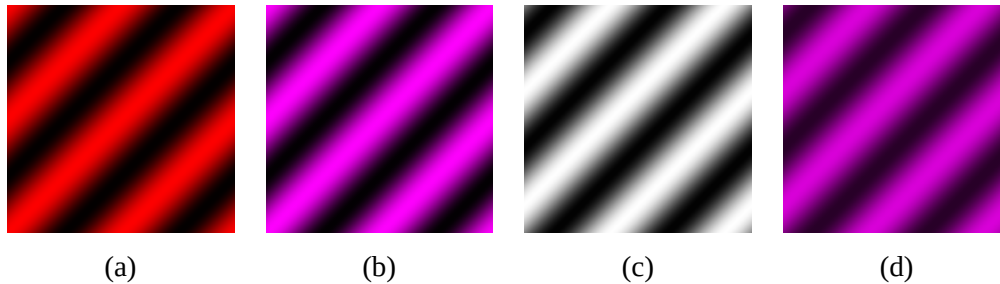


FIGURE 3.25 – Exemples d'initialisation : (a) init avec μ_{Lum} $F_i[o, p] = K_i$ et $F_i[-o, -p] = K_i$ (b) init avec μ_{Lum} $F_i[o, p] = K_i$, $F_k[o, p] = K_k$, $F_i[-o, -p] = K_i$ et $F_k[-o, -p] = K_k$ (c) init avec $\mu_{Lum} = \frac{i+j+k}{\sqrt{3}}$, $F_r[o, p] = K_r$ et $F_r[-o, -p] = -K_r$ (d) init avec $\mu_{Magenta} = \frac{i+k}{\sqrt{2}}$, $F_r[o, p] = K_r$ et $F_r[-o, -p] = -K_r$

Pour la partie réelle, on remarque donc que la variation de couleur obtenue est fonction de la direction donnée à la transformée de Fourier. Par exemple, si la direction μ vaut $(i + k)/\sqrt{2}$, le résultat sera une variation impaire sur les deux composantes i et k (figure 3.25). Ceci équivaut à une variation avec autant d'amplitude dans le rouge que le bleu soit, d'après la synthèse additive, le magenta.

Interprétation du spectre quaternionique en y associant luminosité et chromaticité. Une dernière façon d'analyser le contenu fréquentiel obtenu par la transformée de Fourier quaternionique a été proposée par Sangwine et al. [82]. Elle consiste à séparer les coefficients du spectre en utilisant les notations parallèle/perpendiculaire.

Tout coefficient du spectre F à la coordonnée spectrale $[o, p]$ peut être séparé en une partie parallèle et une partie perpendiculaire à l'axe μ de la transformée en utilisant la décomposition symplectique :

$$F[o, p] = F_{\parallel}[o, p] + F_{\perp}[o, p] \quad (3.73)$$

décrite précédemment. Dans la suite nous notons ν un quaternion imaginaire pur orthogonal à μ le quaternion pur unitaire associé à cette décomposition symplectique.

Cette dernière équation peut s'écrire sous forme polaire :

$$F[o, p] = A_{\parallel}[o, p] \exp^{\mu\theta_1[o, p]} + A_{\perp}[o, p] \exp^{\mu\theta_2[o, p]} \nu. \quad (3.74)$$

A partir de cette expression, Sangwine et al. développent dans [82] une interprétation du spectre. Pour alléger les notations, nous omettons les coordonnées fréquentielles $[o, p]$

Si l'on raisonne par rapport au domaine spatial et donc après application d'une TFQ inverse, les parties parallèle et perpendiculaire du coefficient $F[o, p]$ multiplient les fonctions exponentielles de la transformée de Fourier quaternionique.

La partie parallèle interagit avec les exponentielles de la transformée sans changer son axe :

$$\exp[\mu(\alpha m + \beta n)] A_{\parallel} \exp[\mu\theta_1] = A_{\parallel} \exp[\mu(\alpha m + \beta n + \theta_1)]$$

où $\alpha = (2\pi o)/M$ et $\beta = (2\pi p)/N$ sont les fréquences horizontales et verticales. La phase θ_1 est donc simplement celle de la partie parallèle et $A_{\parallel}$ est son amplitude.

Ensuite lorsque que l'on multiplie les exponentielles de la transformée par la partie perpendiculaire $A_{\perp} e^{\mu\theta_2} \nu$, on obtient (ceci seulement parce que ν est perpendiculaire à μ) :

$$\exp[\mu(\alpha m + \beta n)] A_{\perp} \exp[\mu\theta_2] \nu = A_{\perp} \exp[\mu(\alpha m + \beta n + \theta_2)] \nu$$

Ainsi comme la direction de la transformée de Fourier utilisée est ici $\mu = \mu_{gris}$, l'angle θ_2 représente l'angle initial du vecteur ν qui est perpendiculaire au vecteur μ . Comme ν tourne autour de μ , il définit un plan. Pour rappel, nous nous plaçons après calcul de la transformée Fourier inverse, nous raisonnons dans le domaine spatial. Dans ce cas, un plan perpendiculaire à l'axe μ , qui est l'axe des gris, va correspondre à un plan de chrominance. Nous retrouvons alors la notion de chemin couleur introduite dans le cadre des complexes : l'angle θ_2 correspond à la valeur initiale du chemin couleur dans ce plan de chrominance et $A_{\perp}$ matérialise l'amplitude de la rotation.

Avec cette approche on analyse le contenu de l'information fréquentielle avec des vecteurs qui effectuent des rotations dans l'espace couleur. Ces rotations peuvent être caractérisées par des chemins couleur comme nous l'avons vu précédemment puisque l'on effectue les rotations sur le plan de chrominance. La théorie de Mc-Cabe[155] est alors utile pour appréhender la construction de chemins couleur qui sont des sommes de chemins de bases. C'est ce qu'ont proposé de développer Ell et al [82].

Cependant, comme dans le cadre de la TF complexe, cette modélisation reste peu aisée à manipuler dans un souci d'analyse.

3.3.2.4 Applications de Fourier quaternionique

A notre connaissance, on ne peut trouver dans la littérature que quelques applications dans lesquelles est utilisée la transformée de Fourier quaternionique. L'information fréquentielle contenue dans le spectre quaternionique est par exemple utilisée en tatouage d'image numérique [18, 19]. La TFQ a aussi été utilisée pour calculer des corrélations d'images couleur [204, 163, 176].

On peut construire de simples applications de filtrage en travaillant directement sur le spectre, à travers un fenêtrage. Ceci a permis ainsi de développer des approches de type décomposition de Gabor par fenêtrage Gaussien de la TFQ couleur [250]. Dans cet esprit, nous illustrons sur la figure 3.3.2.4 quelques exemples de fenêtrages fréquentiels, et notamment les résultats obtenus pour un filtrage passe-haut suivant cette méthode.

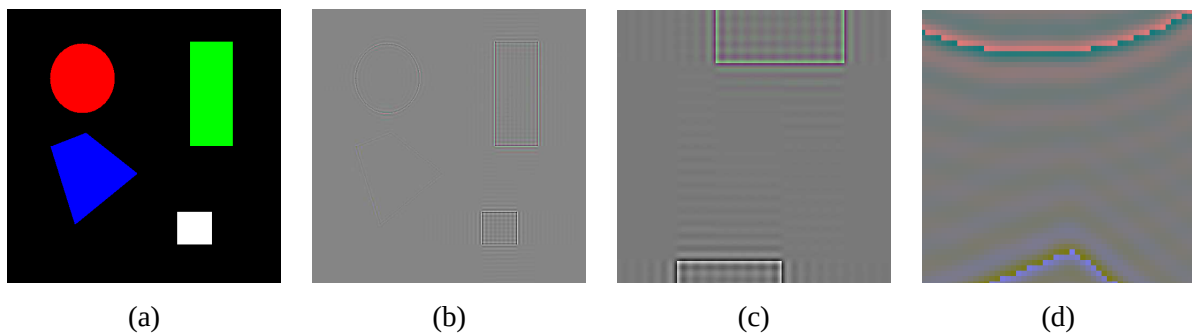


FIGURE 3.26 – Résultats du filtrage quaternionique passe-haut : (a) image couleur simple ; (b) résultat du passe-haut (image représentée avec une inversion des couleurs) ; (c) détails des contours des rectangles vert et blanc ; (d) détails des contours du cercle rouge et du quadrilatère bleu. On remarque que le filtrage fréquentiel préserve les couleurs des images originales.

Nous constatons que ce filtrage conserve les couleurs des images originales et que le contenu haute-fréquence est bien isolé du reste des images. En effet les contours des rectangles, du cercle et du quadrilatère sont de la même couleur que dans l'image originale. Puisque nous avons utilisé un simple fenêtrage dans le domaine fréquentiel et, même si le produit fréquentiel n'est pas équivalent au produit de convolution spatial avec les quaternions, nous remarquons qu'il apparaît un type d'artefacts équivalent à ceux introduits par un fenêtrage fréquentiel par une fonction porte sur des images en niveaux de gris. De plus, on constate que les oscillations induites par notre fenêtrage fréquentiel varient entre les couleurs originales des images et leurs complémentaires (en synthèse additive). On illustre ainsi l'interprétation du spectre dont les coefficients traversent des chemins elliptiques sur le plan chromatique tandis qu'ils gardent à peu près la même clarté. En effet, la direction utilisée ici est l'axe des niveaux de gris μ_{gris} .

De nouveau, comme nous l'avons présenté pour la transformée de Fourier chromatique, nous pouvons définir une décomposition en ondelettes quaternionique couleur en utilisant directement un fenêtrage dans le domaine fréquentiel. Nous rejoignons en cela les travaux sur Gabor. Cependant, la difficulté d'interprétation des coefficients que l'on vient d'exposer reste présente, le fenêtrage n'y change rien. Les auteurs qui décomposent l'image couleur la transformée de Fourier quaternionique et Gabor utilisent tous, à ma connaissance le module des coefficients. Nous proposons en annexe de ce mémoire une construction originale d'un banc de filtres quaternionique reconstituable qui est une autre manière de définir une analyse par ondelettes pour les images couleur dans le cadre des quaternions.

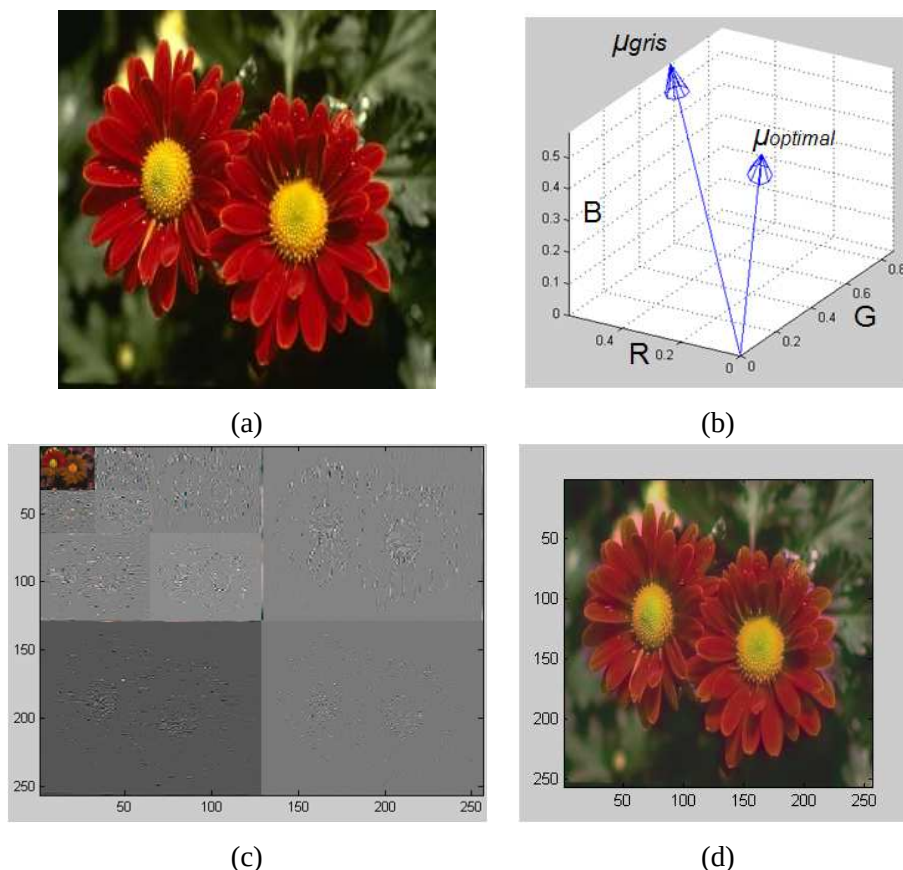


FIGURE 3.27 – Compression de l'image Flowers à partir d'une décomposition en ondelettes quaternionique couleur adaptée (a) Image originale (b) Axe utilisé pour la transformée (c) Décomposition en ondelettes (d) Image reconstruite

Nous allons maintenant proposer et étudier différentes approches utilisant les algèbres de Clifford pour analyser les images couleur, qui sont actuellement les définitions les plus avancées en ce qui concerne Fourier pour les images couleur.

3.3.3 Approche fréquentielle par algèbres de Clifford pour les images couleur

A notre connaissance, la première transformée de Fourier définie dans le cadre des algèbres de Clifford a été proposée par Brackx et al. [28, 29] avec une étude particulière pour les fonctions bidimensionnelles. Revoyons brièvement cette approche.

3.3.3.1 $\mathcal{G}_2$ et images couleur

On définit la Transformée de Fourier Cliffordienne Continue 2D (TFCC2D) sur des vecteurs de $\mathbb{R}_n$ en se basant sur [28, 29] par :

$$F(\mathbf{u}) = \int_{\mathbb{R}_2} \exp(2\pi(\mathbf{u} \wedge \mathbf{x})) f(\mathbf{x}) dV(\mathbf{x}) \quad (3.75)$$

avec $\mathbf{u}$ la composante fréquentielle et $\mathbf{x}$ la composante spatiale.

Lorsque l'on travaille en 2D on choisira les vecteurs $\mathbf{u} = ue_1 + ve_2$ et $\mathbf{x} = xe_1 + ye_2$. L'exponentielle du produit externe s'exprime donc par $\exp^{\mathbf{u} \wedge \mathbf{x}} = \cos(uy - vx) + \sin(uy - vx)e_{12}$ et est constituée d'une partie scalaire et d'une partie bivectorielle.

D'une vision plus numérique, l'équation peut s'écrire :

$$F(u, v) = \mathcal{F}_{\mathcal{H}^+}[f](u, v) = \int_{\mathbb{R}} \int_{\mathbb{R}} \exp(2\pi(uy - vx)e_{12}) f(x, y) dx dy \quad (3.76)$$

On peut noter que les expressions des transformées présentées auparavant étaient basées sur un produit scalaire, à savoir que l'équivalent dans $\mathbb{R}_2$ est un noyau de la transformée qui s'écrit

$$\exp^{2\pi(ux + vy)e_{12}} = \exp^{2\pi(\mathbf{u} \cdot \mathbf{x})e_{12}}$$

. Dans la transformée de Bracks et al. il est utilisé le produit extérieur. L'une des idées sous-jacentes est d'introduire l'analyse de Clifford dans le noyau de Fourier en remplaçant l'opérateur "exponentielle" à valeur scalaire par un opérateur "exponentielle" à valeur dans l'algèbre de Clifford.

Bien sûr dans le cadre des images numériques, la transformée doit être discrétisée. Nous proposons de définir tout simplement la transformée de Fourier Cliffordienne discrète (TFCD) de Bracks comme suit :

$$F[o, p] = \mathcal{F}_{\mathcal{H}^+}[f][o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \exp(2\pi(\frac{om}{M} - \frac{pn}{N})e_{12}) f[m, n] \quad (3.77)$$

La transformée inverse est donnée par :

$$f[m, n] = \mathcal{F}_{\mathcal{H}^-}[F][m, n] = \frac{1}{\sqrt{MN}} \sum_{o=0}^{M-1} \sum_{p=0}^{N-1} \exp(-2\pi(\frac{om}{M} - \frac{pn}{N})e_{12}) F[o, p] \quad (3.78)$$

Pourtant, l'algèbre géométrique $\mathcal{G}_2$ n'est pas naturelle pour représenter une image couleur. Les travaux de Bracks sont présentés dans le cadre des fonctions de $\mathbb{R}^2$ vers $\mathbb{R}_2$.

Si l'on veut quand même utiliser $\mathcal{G}_2$ pour coder l'information couleur d'une image numérique, on peut par exemple au point de coordonnées $[m, n]$ introduire la fonction f égale à :

$$f[m, n] = f_r[m, n]e_0 + f_g[m, n]e_1 + f_b[m, n]e_2$$

où les fonctions f_r, f_g et f_b correspondent à chacune des composantes rouge, verte et bleue de l'image. Cependant cet encodage de l'information couleur sur l'algèbre $\mathcal{G}_2$, interdit l'utilisation des transformations géométriques couleur que nous avons rappelées. En effet, ces opérations, comme les projections, les rotations, les réjections ainsi que les translations ne sont définies dans les algèbres géométriques que pour des multivecteurs simples, c'est à dire s'exprimant sur un grade unique de l'algèbre. Le choix effectué ici code l'information rouge sur le grade 0 et les informations verte et bleue sur le grade 1 de l'algèbre et empêche donc l'utilisation de telles transformations.

Cette non homogénéité dans le codage de l'information couleur a aussi des conséquences dans la traduction des informations dans le domaine spectral après transformée de Fourier. En effet par simple calcul on peut voir que les parties scalaires et bivectorielles sont uniquement réservées à la description des variations de la composante rouge avec les variations paires de la composante rouge sur la partie scalaire du spectre et les variations impaires seront décrites par la partie bivectorielle du spectre. Les parties vectorielles décrivent quant à elles les variations pour les composantes verte et bleue mais les variations paires de la composante verte seront mélangées avec les variations impaires de la composante bleue sur la première partie vectorielle du spectre alors que ce sera l'inverse sur la seconde partie vectorielle.

Pour lutter contre ce problème de non homogénéité dans la description de l'information couleur, on peut envisager d'utiliser un espace Luminance-Chrominance :

$$f[m, n] = f_L[m, n]e_0 + f_{c_1}[m, n]e_1 + f_{c_2}[m, n]e_2.$$

Nous n'avons pas poursuivi l'analyse de cette stratégie, mais nous nous sommes plutôt tournés vers l'étude de l'algèbre géométrique $\mathcal{G}_3$ qui présente beaucoup plus d'avantages que $\mathcal{G}_2$ dans la manipulation ainsi que l'analyse des couleurs.

3.3.3.2 $\mathcal{G}_3$ et images couleur

Tout d'abord, si nous reprenons les liens existants entre les quaternions et l'algèbre $\mathcal{G}_3$ (pour mémoire les éléments pairs forment une sous algèbre de $\mathbb{R}_{3,0}$ isomorphe à $\mathbb{H}$), une solution simple pour généraliser la transformée de Fourier au cadre des algèbres géométriques revient à coder l'information colorimétrique de l'image sur les trois parties bivectorielles d'un multivecteur de $\mathcal{G}_3$, et dans ce cas, une transformation de Fourier peut être définie comme utilisant les bivecteurs de $\mathcal{G}_3$, par une identification directe avec celle définie avec les quaternions (equation 3.64). L'analyse du domaine fréquentiel est donc identique et nous ne reviendrons pas dessus.

Dans cette partie, nous allons revoir la nouvelle transformée de Fourier appliquée aux images couleur, utilisant l'algèbre géométrique $\mathcal{G}_3$ que nous avons décrit notamment dans [62]. La manière dont sera encodée la couleur préservera la possibilité d'effectuer des transformations géométriques sur les vecteurs couleur. En effet, une couleur va être considérée comme un 1-vecteur de $\mathcal{G}_3$.

On codera donc l'image par une fonction de $\mathbb{Z}^2 \rightarrow \mathcal{G}_3$ avec $f[m, n] = f_R[m, n]e_1 + f_G[m, n]e_2 + f_B[m, n]e_3$ pour une image RVB. Avec f_R , f_G et f_B les trois fonctions représentant les composantes couleurs dans l'espace RVB .

On peut trouver différents travaux proposant l'extension de la transformée de Fourier dans $\mathcal{G}^3$ à travers l'utilisation du pseudoscalaire. Ces travaux ont été menés indépendamment dans des cadres différents : nous avons mené cette réflexion pour la problématique du traitement des images couleur (cela rejoint aussi les travaux de Mawardi [154]), Ebling et al. ont construit la transformée de Fourier pour la manipulation de champ de vecteurs [77] (comme pour le produit de convolution). Tous ces travaux obtiennent la même forme pour la transformée, à savoir pour une image de dimension $M \times N$, nous définirons la « transformation de Fourier Cliffordienne couleur » dans $\mathcal{G}_3$ discrète (TFCG3D) par :

Pour $o = 0, \dots, M - 1$ et $p = 0, \dots, N - 1$

$$F[o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \exp^{-2\pi e_{123} \left(\frac{om}{M} + \frac{pn}{N} \right)} f[m, n] \quad (3.79)$$

o et p étant les coordonnées dans le domaine fréquentiel tandis que m et n sont les coordonnées dans le domaine spatial.

On observe que le noyau de la transformée est différent de celui de la définition proposée par Brackx [28, 29], en effet ici, on revient à la définition classique correspondant au produit scalaire.

Comme pour la plupart des transformées de Fourier hypercomplexes [86, 81, 177, 77], il est possible d'effectuer la transformation de Fourier Cliffordienne couleur définie précédemment en utilisant des transformations de Fourier complexes rapides classiques, à travers le découpage en une partie parallèle et une partie perpendiculaire.

Puisque l'élément de l'algèbre géométrique du noyau est le pseudoscalaire e_{123} et que dans $\mathcal{G}_3$ c'est un élément qui commute avec tous les autres, la transformée de Fourier ainsi définie conserve de nombreuses propriétés (on note $\mathcal{F}$ la TF définie par l'équation 3.79) :

- la linéarité : si $f(\mathbf{x}) = \alpha f_1(\mathbf{x}) + \beta f_2(\mathbf{x})$ alors $\mathcal{F}\{f\}(\mathbf{u}) = \alpha \mathcal{F}\{f_1\}(\mathbf{u}) + \beta \mathcal{F}\{f_2\}(\mathbf{u})$;
- retard : si $f_d(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$ alors $\mathcal{F}\{f_d\}(\mathbf{u}) = e^{-e_{123}\mathbf{u}\cdot\mathbf{x}} \mathcal{F}\{f\}(\mathbf{u})$;
- facteur d'échelle : si $a \in \mathbb{R}^+$ et $f_a(\mathbf{x}) = f(a\mathbf{x})$ alors $\mathcal{F}\{f_a\}(\mathbf{u}) = \frac{1}{a^3} \mathcal{F}\{f\}(\frac{\mathbf{u}}{a})$;
- shift : si $\mathbf{u}_0 \in \mathbb{R}^3$ et $f_0(\mathbf{x}) = f(\mathbf{x})e^{e_{123}\mathbf{u}_0\cdot\mathbf{x}}$ alors $\mathcal{F}\{f_0\}(\mathbf{u}) = \mathcal{F}\{f\}(\mathbf{u} - \mathbf{u}_0)$;
- dérivée : $\mathcal{F}\{\nabla f\}(\mathbf{x}) = e_{123}\mathbf{u} \mathcal{F}\{f\}(\mathbf{u})$;
- théorème de convolution : $\mathcal{F}\{f_1 * f_2\}(\mathbf{x}) = \mathcal{F}\{f_1\}(\mathbf{u}) \mathcal{F}\{f_2\}(\mathbf{u})$.

Afin de bien comprendre cette nouvelle transformée, le plus simple est de décomposer cette transformée en différentes transformées complexes élémentaires. Pour cela, il suffit de faire des regroupements judicieux. Par exemple si l'on prend la partie de f associé à e_{23} et e_1 :

$$F[o, p] = \frac{1}{\sqrt{MN}} \sum_{m,n} \exp^{-2\pi e_{123}(\frac{om}{M} + \frac{pn}{N})} (f_{23}[m, n]e_{23} + f_1e_1) \quad (3.80)$$

$$= \frac{1}{\sqrt{MN}} \sum_{m,n} \exp^{-2\pi e_{123}(\frac{om}{M} + \frac{pn}{N})} (f_{23}[m, n]e_{123} + f_1)e_1 \quad (3.81)$$

$$(3.82)$$

Par identification $e_{123} \rightarrow i$ on retrouve une transformée de Fourier complexe classique. Donc en effectuant plusieurs fois ce type de regroupement, on obtient l'écriture suivant :

$$\begin{aligned} F[o, p] = & \mathfrak{F}_c[f_0 + f_{123}e_{123}][o, p]1 + \mathfrak{F}_c[f_1 + f_{23}e_{123}][o, p]e_1 \\ & + \mathfrak{F}_c[f_2 + f_{31}e_{123}][o, p]e_2 + \mathfrak{F}_c[f_3 + f_{12}e_{123}][o, p]e_3 \end{aligned} \quad (3.83)$$

avec $\mathfrak{F}_c$ la transformée de Fourier classique.

Cette formulation est bien utilisée pour accélérer les calculs avec l'utilisation de l'algorithme FFT. A partir de cette formulation, nous constatons immédiatement que si on découpe l'information couleur du pixel f de coordonnées m et n en trois composantes, rouge $r[m, n] = f_1[m, n]$, verte $v[m, n] = f_2[m, n]$ et bleue $b[m, n] = f_3[m, n]$, alors la transformée de Fourier couleur est équivalente à trois transformées de Fourier classiques appliquées sur chacun des plans couleur, que l'on recombine ensuite :

$$F[o, p] = \mathfrak{F}_c[f_1][o, p]e_1 + \mathfrak{F}_c[f_2][o, p]e_2 + \mathfrak{F}_c[f_3][o, p]e_3$$

De cette constatation, on peut caractériser rapidement le spectre couleur. Tout d'abord le spectre ne contient pas d'information sur ses parties scalaires et trivectorielles. Par contre l'information de couleur

est séparée sur la partie vectorielle pour les variations paires et la partie bivectorielle pour les variations impaires. On voit aussi que chacune des variations impaires décrites dans la partie bivectorielle est associée à la composante duale de la partie vectorielle. Autrement dit e_{23} est associé à f_1 , e_{31} est associé à f_2 et e_{12} est associé à f_3 . Chaque composante couleur est donc traitée de la même façon dans le spectre ce qui est déjà complètement différent de ce que nous avons constaté avec la transformée de Fourier Cliffordienne dans $\mathcal{G}_2$.

La conclusion est que cette transformée de Fourier Cliffordienne qui utilise la partie pseudoscalaire dans le noyau de l'exponentielle est donc équivalente à une transformée de Fourier marginale qui consiste à appliquer une transformée de Fourier complexe sur chaque composante couleur. Ceci limite bien sûr son intérêt.

On peut alors s'interroger sur l'utilisation d'un autre élément dans le noyau de l'exponentielle, à savoir par exemple un bivecteur. Ceci fait l'objet de la section suivante.

3.3.3.3 Une transformée de Fourier couleur utilisant $\mathcal{G}_4$

Nous allons étudier dans ce dernier paragraphe la version la plus récente de la transformée de Fourier qui a été proposée par nos collègues de La Rochelle dans [20]. Le papier d'origine a une vision assez formelle, nous proposons ici un bref résumé en mettant en avant surtout le principe.

L'idée de départ est la constatation que la transformée de Fourier complexe 2-D d'une fonction à deux variables $f : x \mapsto f(x, y)$ qui s'écrit donc :

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp(-i(ux + vy)) dx dy \quad (3.84)$$

peut, si l'on introduit le formalisme de l'algèbre de Clifford, s'interpréter comme une rotation dans le plan complexe.

Dans le cadre de l'algèbre de Clifford, la fonction à valeurs réelles va être plongée dans $\mathbb{R}_{2,0}$ par :

$$f : (x, y) \mapsto f_r(x, y)e_1 + 0.e_2$$

avec f_r la partie réelle, et la transformée de Fourier s'écrit :

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [\cos(\frac{1}{2}(ux + vy)) + e_{12} \sin(\frac{1}{2}(ux + vy))](f_r(x, y)e_1 + 0.e_2) \cos(-\frac{1}{2}(ux + vy)) + e_{12} \sin(-\frac{1}{2}(ux + vy)) dx dy \quad (3.85)$$

Cette transformée se lit comme une rotation dans $\mathbb{R}_{2,0}$ qui s'écrit :

$$x \rightarrow \tau^{-1} x \tau \quad (3.86)$$

avec $\tau = \cos(\frac{1}{2}(ux + vy)) + e_{12} \sin(\frac{1}{2}(ux + vy))$ correspondant à la somme d'un scalaire et d'un bivecteur donc des éléments de grade pair, τ est un spineur de $\mathbb{R}_{2,0}$.

Nous avons vu qu'un pixel couleur se code fort logiquement dans $\mathbb{R}_{3,0}$ par :

$$f : (x, y) \mapsto f_r(x, y)e_1 + f_g(x, y)e_2 + f_b(x, y)e_3$$

l'idée proposée par T. Batard [20] pour étendre la transformée de Fourier couleur est d'utiliser l'écriture adaptée aux rotations dans $\mathbb{R}^3$. De nouveau, nous avons vu que les rotations dans $\mathbb{R}_{3,0}$ s'écrivent :

$$x \rightarrow \tau^{-1} x \tau$$

avec τ correspondant à la somme d'un scalaire et d'un bivecteur donc des éléments de grade pair, τ est un spineur de $\mathbb{R}_{3,0}$. L'exponentielle de la transformée de Fourier va correspondre à la somme d'un scalaire et d'un bivecteur, et non à la somme d'un scalaire et d'un vecteur comme dans la version de Sangwine. Finalement la transformée de Fourier pour les images couleur plongées dans $\mathbb{R}_{3,0}$ va s'écrire :

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\left(\frac{B}{2}(ux + vy)\right) f(x, y) \exp\left(-\frac{B}{2}(ux + vy)\right) dx dy \quad (3.87)$$

avec B un bivecteur de norme 1.

Cependant pour des contraintes d'implantation, les collègues de La Rochelle ont plutôt considéré le groupe $Spin(4)$ que le groupe $Spin(3)$ (ayant pour conséquence la séparation en deux transformées de Fourier complexes permettant l'utilisation de la FFT).

De ce fait, dans la suite, une image couleur est considérée comme une application f de $\mathbb{R}^2$ à valeurs dans la partie vectorielle de $\mathcal{G}_4 = \mathbb{R}_{4,0}$:

$$f : (x, y) \mapsto f_r(x, y)e_1 + f_g(x, y)e_2 + f_b(x, y)e_3 + 0e_4 \quad (3.88)$$

La composante sur e_4 est nulle, de la même façon que pour une fonction réelle la composante imaginaire est nulle.

En conséquence la transformée de Fourier couleur va utiliser un élément pair de l'algèbre $\mathbb{R}_{4,0}$ pour construire le Spineur. Pour comprendre la formulation finale, nous reprenons les détails sur $Spin(4)$ donnés par exemple dans [94] (référence permettant de décortiquer les différents concepts). Tout élément pair, x , de l'algèbre $\mathbb{R}_{4,0}$ est de la forme $x = x_01 + x_1e_1e_2 + x_2e_1e_3 + x_3e_1e_4 + x_4e_2e_3 + x_5e_2e_4 + x_6e_3e_4 + x_7e_1e_2e_3e_4$. On note

$$i = e_1e_2 \quad j = e_2e_3 \quad k = e_1e_3 \quad i' = e_4e_3 \quad j' = e_4e_1 \quad k' = e_4e_2 \quad \mathbb{I} = e_1e_2e_3e_4$$

d'où les relations

$$ij = k \quad jk = i \quad ki = j \quad i^2 = -1 \quad j^2 = -1 \quad k^2 = -1$$

$$i\mathbb{I} = \mathbb{I}i = i' \quad j\mathbb{I} = \mathbb{I}j = j' \quad k\mathbb{I} = \mathbb{I}k = k'$$

On en déduit que tout élément x de $Spin(4)$ s'écrit sous la forme :

$$x = u + \mathbb{I}v$$

où u et v sont des quaternions tels que $u\bar{u} + v\bar{v} = 1$ et $u\bar{v} + v\bar{u} = 0$

Dans ce cas la transformée de Fourier couleur plongées dans $\mathbb{R}_{3,0}^4$ va s'écrire :

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\left(\frac{e_{1234}B}{2}(ux + vy)\right) \frac{B}{2}(ux + vy) f(x, y) \exp\left(-\frac{e_{1234}B}{2}(ux + vy)\right) \exp\left(-\frac{B}{2}(ux + vy)\right) dx dy \quad (3.89)$$

avec B un bivecteur de norme 1. De nouveau on constate qu'il faut définir une direction d'analyse. Simple-ment nous sommes passés d'un vecteur dans la définition quaternionique à un bivecteur B correspondant à un plan. Le produit $e_{1234}B$ correspond au plan orthogonal à B dans $\mathbb{R}^4$. Batard et al. montrent que la définition de la transformée de Fourier proposée, sous sa forme générale, permet de retrouver les deux transformées de Sangwine et Bülow [20].

Si l'on analyse précisément la transformée ainsi définie, nous retrouvons le principe de la séparation de la transformée en deux transformées complexes

$$F[o, p] = \frac{1}{N} \sum_{m,n=0}^{N-1} e^{\frac{2B\pi(mo+np)}{2N}} f\|_B[n, m] e^{\frac{-2B\pi(mo+np)}{2N}} \quad (3.90)$$

$$+ \frac{1}{N} \sum_{m,n=0}^{N-1} e^{\frac{2e_{1234}B\pi(mo+np)}{2N}} f\|_{e_{1234}B}[n, m] e^{\frac{-2e_{1234}B\pi(mo+np)}{2N}}$$

Nous indiquons sur la figure 3.28 une visualisation géométrique de la séparation ainsi faite. Nous constatons que de nouveau nous retrouvons le principe de la transformée de Fourier chromatique. Prenons l'exemple le plus immédiat : si $B = \vartheta \wedge e_4$ avec ϑ l'axe des gris alors dans ce cas $f\|$ correspond à $f\|_{\vartheta}$ et nous retrouvons exactement la transformée exposée en début de chapitre.

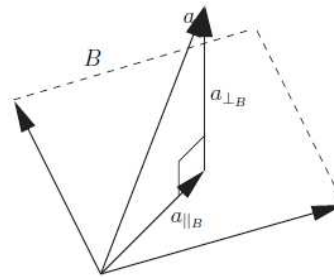


FIGURE 3.28 – Visualisation géométrique de la transformée de Fourier dans $\mathcal{G}_4$

Comme pour les autres transformées, nous pouvons appliquer un fenêtrage dans le domaine fréquentiel. Nous reprenons par exemple ici les expérimentations menées par nos collègues de La Rochelle [20]. La figure 3.3.3.3 montre le résultat d'un filtrage directionnel. L'image originale est l'image de gauche. Le bivecteur choisi pour paramétrer la transformée est $B = e_1 \wedge e_4$ correspondant à la couleur rouge. Le masque utilisé dans le domaine de Fourier parallèle vaut 0 sur l'ensemble $\{|\arg(z) - \pi/2| < \varepsilon\} \cup \{|\arg(z) + \pi/2| < \varepsilon\}$, et 1 ailleurs. On peut voir sur l'image de droite que les lignes rouges ont disparu alors que les lignes vertes (qui sont dans la partie orthogonale) ne sont pas modifiées.

Notons pour finir que les auteurs proposent que la métrique utilisée dans le calcul de la transformée de Fourier se soit pas forcément la métrique euclidienne. Dans ce cas, la définition du bivecteur $e_{1234}B$ correspondant au plan orthogonal à B va être modifié en fonction par exemple du contenu de l'image. Le principe d'une adaptation de la métrique à l'information que l'on cherche à analyser se retrouve dans de nombreuses applications, notamment pour le traitement des images hyperspectrales. Par exemple, dans le cadre de la thèse de S. Rousseau [196](thèse financée par la DGA) nous avons mené une réflexion sur une notion de métrique adaptée pour la segmentation d'images hyperspectrales [197].

Enfin, nous appliquons le même type d'expérimentation que pour les transformées précédentes, à savoir la définition d'une transformée en ondelettes en utilisant les liens exprimées précédemment entre cette nouvelle transformée et la transformée chromatique d'origine. De nouveau, je reprends une expérimentation développée dans le cadre du Workshop sur les algèbres géométriques [42]. Nous illustrons sur la figure 3.30

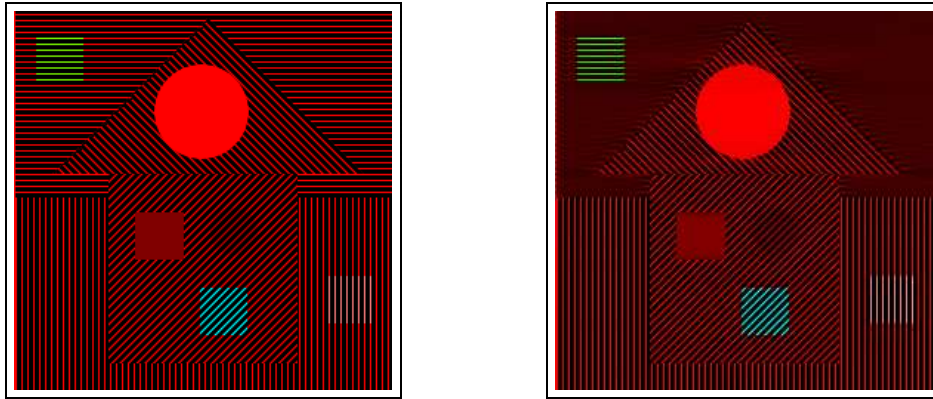
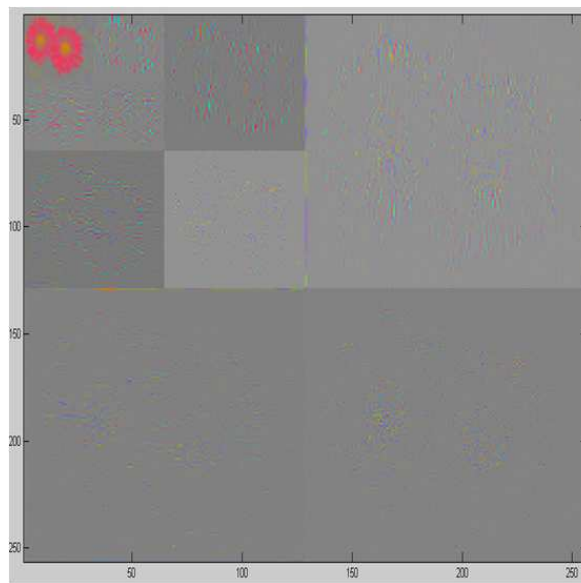


FIGURE 3.29 – Image originale - Image après filtrage directionnel tiré de [20]

la décomposition obtenue en utilisant le bivecteur $B = e_1 \wedge e_4$ correspondant à la couleur rouge. Nous constatons que la décomposition met bien en évidence les informations liées majoritairement au rouge.

ar exemple, nous présentons sur la figure 3.23, une expérimentation que j'avais développée dans le cadre du Workshop sur les algèbres géométriques qui a eu lieu à La Rochelle en 2009 [42].


 FIGURE 3.30 – Décomposition en ondelettes $\mathcal{G}^4$ couleur de l'image Flowers

Cette nouvelle transformée a permis d'introduire une approche généraliste et unificatrice. Toutefois, au travers des expérimentations exposées ci-dessus, nous constatons qu'en terme numérique, aucune information supplémentaire en termes d'analyse n'est apportée. L'utilisation de ces nouveaux formalismes demandent, tels qu'ils sont présentés dans ce chapitre, encore de nombreux approfondissements.

Nous proposons en annexe de ce mémoire une proposition que nous avons faite concernant la construction d'un banc de filtres quaternionique pour les images couleur. Nous réutilisons pour cela la définition de la transformée de Fourier quaternionique de Sangwine. Cette construction fait le lien avec le chapitre suivant, puisqu'il nous permet de faire le point sur les décompositions en ondelettes utilisant ces nouveaux formalismes et qu'il utilise une démarche purement signal, tous ces éléments composent entre autre le chapitre

suivant.

3.4 Conclusion

Ce chapitre avait pour objectif de présenter les travaux menés dans le cadre de l'utilisation de formalismes algébriques pour analyser et traiter les images numériques couleur. J'ai créé et animé cet axe de recherche dès mon recrutement au sein du département XLIM-SIC. Il a trouvé son implication scientifique dans de nombreux encadrements dont deux thèses et il a été un élément fédérateur dans la construction d'un thème scientifique transverse au sein de la fédération Prides. Ce sujet de recherche est encore très confidentiel et il m'a permis d'accéder à des approches complémentaires des démarches classiques utilisées en image.

Scientifiquement, nous avons observé que les solutions développées ces dix dernières années consistent à coder l'information couleur par de simples éléments de l'algèbre utilisée. La manipulation des vecteurs couleur utilisant par exemple des produits vectoriels ou encore des produits matriciels se résument alors à des opérations simples permettant de manier les éléments de l'algèbre.

Il existait des travaux antérieurs et c'est pourquoi j'ai tout d'abord rappelé les prémisses de ces approches, à savoir l'analyse de l'information chromatique grâce à l'utilisation de nombres complexes, notamment à travers la définition d'une transformée de Fourier « chromatique ». En effet cette transformée permettait d'analyser partiellement le comportement des couleurs dans un espace fréquentiel. Mais malgré les travaux sur la notion de « chemin couleur » proposée initialement par McCabe, nous avons vu que la question de l'interprétation n'était pas simple.

La nature complètement tridimensionnelle de la couleur, a tout naturellement mené à une généralisation de ces travaux vers un formalisme comportant plus de dimensions. C'est pourquoi dès 2001, j'ai débuté un travail sur l'opportunité d'utiliser les quaternions pour manipuler les images couleur, comme il a été proposé il y a 15 ans par Sangwine. Comme nous l'avons vu les quaternions, contrairement aux nombres complexes, peuvent décrire complètement une information couleur, par exemple sur leurs parties imaginaires. On peut alors manipuler un vecteur couleur par de simples opérations comme le produit. Nous avons alors proposé différentes ré-écritures d'opérations couleur à travers des transformations géométriques sur des vecteurs couleur au moyen de ce formalisme. Parmi les travaux menés au département XLIM-SIC, nous avons proposé une extension du détecteur de ruptures couleur de Sangwine.

Nous avons ensuite introduit une approche très récente et innovante pour la manipulation des images couleur à savoir les algèbres géométriques. Dans ce chapitre, nous avons vu que les algèbres géométriques $\mathcal{G}_3$ permettent de manipuler des entités géométriques telles que scalaires, vecteurs, bivecteurs et trivecteurs de manière indépendante. Des règles permettent d'affecter certains objets au moyen d'autres et celles-ci sont régies par les différents produits. De nouveau, nous avons proposé différentes ré-écritures d'opérations couleur en utilisant ce formalisme et construit une nouvelle approche spatiale de détection de ruptures couleur, en généralisant celle définie avec les quaternions.

Enfin, dans la dernière partie de ce chapitre, nous avons discuté de la question scientifique qui était le point de départ de mon travail sur les quaternions : la définition d'une transformée de Fourier pour les images couleur.

Nous avons tout d'abord analysé la transformée quaternionique proposée par Sangwine et Ell. Nous avons vu que malgré les travaux récents sur ce point (notamment les nôtres), il existe encore de nombreuses questions en ce qui concerne la complète caractérisation du domaine fréquentiel quaternionique. C'est pour-

quoi, nous avons développé dès 2005 un volet de recherche concernant une transformation de Fourier Cliffordienne appliquée aux images couleurs. Tout naturellement nous avons déployé une solution qui reposait l'algèbre géométrique $\mathcal{G}_3$ avec l'incorporation du pseudo-scalaire dans l'exponentielle. L'étude numérique de cette transformation a mis en évidence que les composantes fréquentielles couleur se séparent suivant les parties paires ou impaires sur les composantes 1-vectorielles et bi-vectorielles du spectre de manière indépendante. Cependant cette transformée de Fourier Cliffordienne qui utilise le pseudoscalaire dans le noyau est équivalente à une transformée complexe appliquée sur chaque composante couleur. Nous avons conclu ce chapitre en présentant les travaux les plus récents menés sur $\mathcal{G}_4$ dans le cadre de la théorie des groupes par nos collègues de La Rochelle. Ce travail de nature plutôt théorique nécessite encore de l'attention sur sa partie numérique car nous n'avons pas pu montrer un apport numérique majeur par rapport à l'analyse simple chromatique décrite au début du chapitre (associée bien sûr à différentes projections de l'image sur des sous-espaces de représentations).

Le sujet exposé dans ce chapitre est un sujet d'étude très récent et de ce fait il ouvre de nombreuses questions. Les techniques apportent une manière élégante de décrire certains problèmes. Cependant, le bilan nous semble pour l'instant mitigé et même à travers les évolutions les plus récentes, nous n'avons pas pu dépasser certaines limites dans l'analyse de l'image couleur. C'est pourquoi, nous sommes retournés aux sources du traitement de l'image, à savoir l'analyse du signal pour définir une représentation à la fois adaptée à un support bidimensionnel mais aussi à la couleur. Ceci fait l'objet du chapitre suivant.

Cet axe portant sur l'apport des algèbres géométriques et des quaternions pour le traitement d'images couleur s'est concrétisé à travers le co-encadrement ou l'intervention dans l'encadrement de trois thèses [62, 179, 196]. Ce thème scientifique est au cœur de la fédération PRIDES et se poursuit dans le cadre de la nouvelle fédération MIRES (Images, Réseau et Sécurité) permettant de nombreux échanges notamment avec nos collègues de La Rochelle. Au niveau national, j'ai assuré la co-animation avec D. Tschumperlé du Greyc une action du GDR ISIS ayant pour sujet les outils mathématiques pour l'image couleur, dans laquelle les algèbres de Clifford avait pleinement leurs places. Enfin, ce thème trouve une application à travers une collaboration contractuelle avec la DGA pour la détection de défaut dans les images hyperspectrales. Ces travaux ont été principalement valorisés dans [64, 37, 45, 63, 180, 181, 197].

Chapitre 4

Ondelettes analytiques

Dans ce troisième chapitre, je vais aborder mes travaux de recherche et d'encadrement les plus récents concernant les décompositions multiéchelles. Ils se situent à l'intersection des deux chapitres précédents et ils concluent d'une certaine façon cette partie concernant la recherche d'un outil numérique. L'objectif reste le même, à savoir définir une transformée en ondelettes en prenant pleinement en compte l'aspect bidimensionnel des images, comme dans le premier chapitre, et en intégrant la dimension couleur comme dans le second chapitre. Le point de départ de ces travaux se situe dans nos études sur les quaternions et l'algèbre de Clifford, menés à l'origine pour la manipulation de la couleur. Ces réflexions nous ont amenés à étudier les travaux de Bülow, Felsberg et Sommer. L'université de Khmel'nyts'kyi a proposé d'utiliser les outils mathématiques que nous avons décrits dans le chapitre précédent, non pas pour manipuler la couleur, mais pour généraliser finement la notion de signal analytique aux images en niveaux de gris. Etant donné l'importance du signal analytique dans la construction de différents outils de traitement mais aussi dans l'extraction d'information, ces travaux ont donné naissance à une nouvelle branche de recherche au sein de la communauté image, pour l'instant s'appuyant sur un nombre réduit d'équipes. Finalement, certaines des dernières avancées n'utilisent plus l'outillage de départ (l'algèbre de Clifford), c'est d'ailleurs notre cas, mais gardent ce souci de retour aux sources dans la définition des opérations pour l'image, à travers par exemple, des questions autour de la notion de fréquence instantanée et donc de phase.

C'est tout logiquement que nous prenons part à cette réflexion, afin, notamment, de définir de nouvelles définitions de transformée en ondelettes qui vont s'efforcer de capter au mieux l'information 2-D mais cette fois, non pas à travers une lecture géométrique mais plutôt par l'extension rigoureuse 2-D d'outils venant du signal.

Cet axe de recherche a pris largement son essor à travers le travail mené actuellement par R. Souillard dans le cadre de sa thèse que j'encadre [222]. Je présenterai à la fin de ce chapitre les tous derniers développements et les pistes suivies.

Mais tout d'abord pour décrire pleinement ce nouvel axe de recherche, il nous faut retourner 50 ans en arrière, à travers la définition du signal analytique.

4.1 Signal analytique et ses extensions

Comme nous l'avons indiqué dans l'introduction, nous allons rappeler des éléments relatifs à la notion de fréquence instantanée 1-D et de phase avec pour objectif son extension aux images. Tout en abordant la notion de fréquence instantanée, nous allons introduire le signal analytique. En effet la fréquence instantanée

se calcule à partir du signal analytique, et finalement c'est lui qui est le point central dans la modélisation de l'information. Notre propos va se concentrer suivant deux points de vue : théorique et pratique, dans le cadre tout d'abord du signal 1-D. Ensuite nous décrirons les pistes d'extension de la notion de signal analytique aux images, et plus généralement aux signaux multi-dimensionnels.

4.1.1 Rappel sur le signal analytique 1-D

Dans le cadre de l'analyse d'un signal, l'apport recherché d'une représentation autre que le domaine spatial ou temporel est lié à sa capacité à décrire le signal de façon compréhensible et adaptée à l'application concernée. La représentation fréquentielle est par exemple particulièrement bien adaptée pour décrire les signaux possédant un caractère périodique marqué. Les signaux que nous rencontrons en pratique sont très souvent non-stationnaires. En général pour ce type de problème, nous supposons une stationnarité locale ; l'analyse spectrale peut se faire sur des fenêtres de courtes étendues sur lesquelles il sera supposé que le signal est stationnaire. Par la suite se sont développées, pour faire de l'analyse spectrale, des méthodes du type temps-fréquence qui présentent souvent l'avantage de ne pas nécessiter d'hypothèse de stationnarité locale. Dans le cadre, de l'analyse Temps-Fréquence, une caractéristique est prépondérante : la fréquence instantanée.

La fréquence instantanée La fréquence instantanée est une des grandeurs adaptées pour la description spectrale des signaux non-stationnaires. Le concept de fréquence instantanée se rattache naturellement au concept plus général de fréquence. Cependant, la notion de fréquence instantanée est ambiguë puisqu'elle sous-entend une mesure instantanée sur un signal non stationnaire. La définition de la fréquence instantanée est problématique et est à l'origine de diverses interprétations. La définition moderne de fréquence instantanée est due à Ville et reprend les résultats de travaux antérieurs effectués par Gabor . Le point de départ est un signal réel exprimé sous la forme la plus classique du point de vue théorique, à savoir la sinusoïde :

$$s(t) = a(t) \cos(\phi(t)) \text{ avec } \phi(t) = wt + \theta \quad (4.1)$$

Ce signal est défini par trois paramètres : l'amplitude a , la pulsation $w = 2\pi f$ et la phase θ . Le problème est qu'en pratique on mesure pour t fixé la quantité $s(t)$, ce qui laisse inconnu les grandeurs a et ϕ .

La fréquence instantanée f_i de ce signal est définie comme la dérivée de la phase ϕ du signal analytique noté z_s associé au signal réel s :

$$z_s(t) = a(t) \exp(j\phi(t)) \text{ et } f_i(t_0) = \frac{1}{2\pi} \frac{d\phi}{dt}(t_0) \quad (4.2)$$

Nous constatons que la notion de fréquence instantanée est liée à la définition d'un signal analytique. Nous proposons de revoir ce concept.

Un pas important dans la définition de la notion de fréquence instantanée fut réalisé par Gabor qui proposa une méthode pour associer un signal complexe à un signal réel de manière unique. Cette définition permit de proposer une définition pour des moments fréquentiels des signaux d'ordre n :

$$\frac{\int \xi^n |Z(\xi)|^2 d\xi}{\int |Z(\xi)|^2 d\xi} \quad (4.3)$$

avec $Z(\xi)$ le spectre du signal complexe. On constate dans cette expression que si Z est remplacé par le spectre du signal réel alors tous les moments d'ordre pair sont nuls.

La méthode envisagée pour calculer ce signal analytique a consisté à annuler les fréquences négatives du spectre du signal et à doubler l'amplitude des fréquences positives de manière à conserver l'énergie du signal. Le signal complexe z_s ainsi obtenu est donc appelé le signal analytique associé au signal réel s :

$$z_s(t) = s(t) + j\mathbf{TH}[s](t) \quad (4.4)$$

où $\mathbf{TH}$ dénote la transformée de Hilbert. La transformée de Fourier du signal analytique est :

$$Z_s(\xi) = 2H(\xi)S(\xi) = S(\xi)(1 + j(-j\text{sign}(\xi))) \quad (4.5)$$

ce qui correspond bien à la relation indiquée initialement, où H est l'échelon unité et la fonction "*sign*" est définie par :

$$\text{sign}(\xi) = \begin{cases} -1 & \text{si } \xi < 0 \\ 1 & \text{si } \xi > 0 \end{cases}$$

Notons que dans le cadre des définitions précédentes en $\xi = 0$, on est amené à poser $\text{sign}(0) = 0$ et $H(0) = 1/2$. Nous pouvons constater que la transformée de Hilbert va prendre la forme suivante dans le domaine de Fourier :

$$\mathbf{TH}[s](\xi) = -j\text{sign}(\xi)S(\xi)$$

Le signal analytique associé au signal réel peut s'obtenir également par des opérations dans le domaine temporel, à travers une opération de filtrage :

$$z(t) = s * (\delta(t) + j\frac{1}{\pi} * \frac{v.p.}{t}) \quad (4.6)$$

et donc

$$\mathbf{TH}[s](t) = \frac{1}{\pi} v.p. \int \frac{s(t-\tau)}{\tau} d\tau \quad (4.7)$$

avec $v.p.$ qui désigne la valeur principale de Cauchy de l'intégrale.¹ Les deux signaux s et $\mathbf{TH}[s]$ sont souvent dits en quadrature puisqu'en théorie ils sont déphasés de $\pi/2$. Par exemple, $\mathbf{TH}[\cos(2\pi t)] = \sin(2\pi t)$.

Il nous faut maintenant préciser le lien entre signal analytique et enveloppe et phase instantanées dans le cadre d'un signal quelconque. Tout d'abord il nous faut nous placer dans le cadre des signaux à spectre étroit (ce qui aura une certaine importance par la suite).

On dit d'un signal qu'il est à spectre étroit si on peut trouver deux quantités, B et $\omega_0 > 0$, telles que

$$|S(\omega)| = 0 \text{ pour } |\omega - \omega_0| > B \quad (4.9)$$

Pour des pulsations éloignées de la pulsation ω_0 , le spectre s'annule. En pratique, cela revient à dire que le spectre ne s'étend pas jusque 0 ni jusqu'à $+\infty$. Pour un tel signal, qui serait réel, on va pouvoir construire une représentation suivant le modèle de la sinusoïde et donc faire apparaître cette notion d'enveloppe et de phase instantanées.

Pour cela, on construit le signal analytique puis on fait "glisser" le spectre du signal analytique de manière à amener cette pulsation ω_0 à l'origine. Le signal obtenu par ce décalage sur l'axe fréquentiel du spectre du

1. Pour mémoire, cela signifie qu'à l'intervalle d'intégration de l'intégrale singulière on enlève un intervalle symétrique $[x_0 - \epsilon, x_0 + \epsilon]$ autour de chaque singularité (ici en 0) et que l'on fait tendre ensuite vers 0 :

$$v.p. \int_{-\infty}^{+\infty} \frac{s(t-\tau)}{\tau} d\tau = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{-\epsilon} \frac{s(t-\tau)}{\tau} d\tau + \int_{\epsilon}^{+\infty} \frac{s(t-\tau)}{\tau} d\tau \quad (4.8)$$

signal analytique, s'appelle l'enveloppe complexe du signal de départ $e_s(t)$. Etant donnée cette opération de décalage, e_s vérifie la relation suivante :

$$e_s(t) = e^{-j\omega_0 t} Z_s(t) \quad (4.10)$$

Afin de faire apparaître les composantes en quadrature, on peut écrire cette enveloppe complexe sous la forme

$$e_s(t) = s_1(t) - js_2(t) \quad (4.11)$$

avec

$$s_1(t) = s(t) \cos \omega_0 t + \mathbf{TH}[s](t) \sin \omega_0 t \quad (4.12)$$

et

$$s_2(t) = s(t) \sin \omega_0 t - \mathbf{TH}[s](t) \cos \omega_0 t \quad (4.13)$$

Ces signaux s_1 et s_2 sont appelés composantes de Rice du signal de départ ou encore composantes en quadrature, car on peut définir le signal de départ à partir de ces composantes. En effet, on peut remarquer que $s(t) = \Re[Z_s(t)] = \Re[e_s(t)e^{j\omega_0 t}]$ et donc

$$s(t) = s_1(t) \cos \omega_0 t + s_2(t) \sin \omega_0 t \quad (4.14)$$

On remarque que tout signal à spectre étroit peut être vu comme un signal obtenu par modulation de 2 porteuses en quadrature (cosinus et sinus) par des signaux s_1 et s_2 dont la largeur de bande est limitée à $[B, B]$. Ces deux signaux sont les composantes de Rice. De ces résultats nous retrouvons aussi la modélisation recherchée, à savoir :

$$s(t) = a(t) \cos(\omega_0 t + \phi(t)) \text{ avec } a(t) = \sqrt{s_1(t)^2 + s_2(t)^2} \text{ et } \phi(t) = -\arctan \frac{s_2(t)}{s_1(t)} \quad (4.15)$$

De ces rappels, on peut aussi retenir que numériquement plus le spectre d'un signal est à bande étroite, plus l'approximation du signal en quadrature réalisée par transformation de Hilbert sera satisfaisante.

4.1.1.1 Lien avec l'analyse de fonction.

Terminons ce paragraphe par une remarque importante concernant le qualificatif "analytique" associé au signal z que nous venons de décrire. Traditionnellement, on qualifie d'analytique une fonction développable en série entière sur son domaine de définition. Un des résultats fondamentaux de la théorie des fonctions d'une variable complexe est qu'une telle fonction est analytique si et seulement si elle remplit les conditions dites de Cauchy-Riemann.

Plus précisément, soit F une fonction définie sur un ouvert B de $\mathbb{C}$. Notons $F(\chi) = u(x, y) + jv(x, y)$ avec $\chi = x + jy$ et u et v des fonctions à valeurs réelles. On dira que la fonction F est holomorphe, ou de façon équivalente satisfait les conditions de Cauchy-Riemann, si u et v vérifient :

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ et } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad (4.16)$$

en tout point du domaine B . Si F est holomorphe alors sa partie réelle u est harmonique, c'est-à-dire est solution de l'équation de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \Delta u = 0 \quad (4.17)$$

et il en va de même pour la partie imaginaire v .

Le signal analytique z peut être considéré comme la restriction au bord $v = 0$ d’une fonction holomorphe $F(u, v)$ définie sur le demi-plan $\{(u, v) \in \mathbb{C}, v \geq 0\}$ et dont la partie réelle est le signal s de départ. C’est ce qui justifie la terminologie employée. La thèse de De Knock décrit longuement ce parallèle [120].

Cette remarque est importante car c’est le point de départ de la généralisation du signal analytique proposée par Felsberg [87] pour les images en niveau de gris. Les conditions de Cauchy Riemann sont remplacées par une équation faisant intervenir l’opérateur de Dirac et l’on parle alors de fonctions monogènes. Nous répréciserons cette approche à la fin de ce chapitre.

4.1.1.2 Interprétation géométrique du signal analytique

En complément de la lecture historique Temps-Fréquence du signal analytique à travers la notion de module et de phase instantanée, T. Bülöw a proposé dans le cadre de sa thèse [35] une interprétation géométrique très originale du signal analytique. Toutefois, comme nous l’avons exprimé précédemment, pour avoir une parfaite lecture physique de la phase locale, il faut que le signal soit à bande étroite. Bülöw se place alors dans le cadre des filtres de Gabor complexes qui sélectionnent une certaine bande de fréquences positives à travers une fenêtre Gaussienne. Le signal ainsi obtenu est complexe, et correspond à un signal analytique. La modélisation du signal issu du filtrage de Gabor complexe est de nouveau la notion de module et de phase, mais avec une interprétation différente : le module $a(t)$ va indiquer la présence relative d’un élément, et la phase $\phi(t)$ décrit la forme de cet élément. Nous résumons dans la table 4.1 les différentes formes identifiées par Bülöw [35].

Phase ϕ	Stade de l’oscillation	Forme du signal
≈ 0	autour d’un maximum	“bosse”
$\approx \frac{\pi}{2}$	décroissance	“descente”
$\approx \pm\pi$	autour d’un minimum	“creux”
$\approx -\frac{\pi}{2}$	croissance	“montée”

TABLE 4.1 – Correspondance entre phase instantanée et structure locale du signal.

Cette vision géométrique de la phase 1D a renforcé l’intérêt de définir un signal analytique 2-D pour analyser la géométrie des images.

Nous proposons maintenant de revoir les différentes pistes formulées pour l’extension du concept de signal analytique à des signaux 2-D à valeurs scalaires.

4.1.2 Extension de la notion de signal analytique aux images en niveaux de gris

En 1D, nous avons vu précédemment qu’à un signal réel peut être associé le signal complexe dit analytique, déterminé de manière unique. La particularité du signal analytique est de ne pas posséder de fréquences négatives, ce qui a d’abord l’avantage, comme nous l’avons dit de donner un sens aux moments fréquentiels, et d’introduire le calcul de la fréquence instantanée. Ce signal analytique a été obtenu simplement en annulant les fréquences négatives et en doublant l’amplitude des fréquences positives (pour conserver l’énergie totale).

En 2D, la notion de fréquences négatives et positives n'est pas établie, ce qui justifie que l'extension de la notion de signal analytique aux images reste un problème ouvert.

Considérons l'exemple standard 2-D :

$$s(x, y) = \cos(x\omega_{x1} + y\omega_{y1}) \cos(x\omega_{x2} + y\omega_{y2}) \quad (4.18)$$

Le signal bi-dimensionnel s est réel et possède deux composantes de fréquences pures $(\omega_{x1} + \omega_{x2}, \omega_{y1} + \omega_{y2})$ et leurs symétriques par rapport à l'origine. Afin de mieux visualiser la dispersion de l'information, nous représentons le signal sous forme exponentielle (permettant d'identifier les pics dans le domaine fréquentiel)

$$s(x, y) = 1/4 \left[e^{j(x(\omega_{x1} + \omega_{x2}) + y(\omega_{y1} + \omega_{y2}))} + e^{j(x(\omega_{x2} - \omega_{x1}) + y(\omega_{y2} - \omega_{y1}))} \right. \\ \left. + e^{j(x(\omega_{x1} - \omega_{x2}) + y(\omega_{y1} - \omega_{y2}))} + e^{j(x(-\omega_{x1} - \omega_{x2}) + y(-\omega_{y1} - \omega_{y2}))} \right] \quad (4.19)$$

Pour étendre le principe de signal analytique aux images, on peut tenter d'appliquer le principe 1D suivant les deux dimensions, à savoir multiplier par quatre l'amplitude des fréquences situées dans le quadrant $(\omega_x > 0, \omega_y > 0)$ et à annuler le reste du spectre, c'est l'approche de Hahn [99]. La traduction de cet opérateur dans le domaine des fréquences serait donc :

$$S_{\mathcal{H}_T}(\xi_1, \xi_2) = -S(\xi_1, \xi_2) \text{sign}(\xi_1) \text{sign}(\xi_2) \quad (4.20)$$

Nous avons alors ce que l'on appelle la transformée de Hilbert totale :

$$s_{\mathcal{H}_T}(x, y) = s(x, y) * \left(\frac{1}{\pi^2 xy} \right) \quad (4.21)$$

Mais cette solution ne convient que pour certains signaux (ceux dont le spectre est entièrement contenu dans ce quadrant). Si nous reprenons l'exemple de l'équation (4.19), alors si $\omega_{x2} > \omega_{x1} > 0$ et $\omega_{y1} > \omega_{y2} > 0$, dans ce cas nous constatons que l'une des composantes va disparaître lors de la transformation.

Conserver tout autre quadrant en annulant les autres ne constituerait bien sûr pas une meilleure solution.

Une solution peut être de conserver deux quadrants et dans ce cas le signal analytique va être construit à partir de deux signaux (chacun correspondant à un quadrant). Cependant cette solution n'est pas satisfaisante puisque nous avons alors deux phases et deux modules ce qui entraîne une ambiguïté dans l'interprétation.

Pour contrecarrer les défauts de la transformée de Hilbert totale, une variante de la solution totale a été proposée par Zhu et al [254] . Il est suggéré que, comme en 1D, le signal analytique s'écrive comme le résultat d'un filtrage :

$$Z(\xi_x, \xi_y) = 2H(\xi_x, \xi_y)S(\xi_x, \xi_y) \quad (4.22)$$

mais avec comme filtre :

$$H(\xi_x, \xi_y) = \frac{1}{4} [\text{sign}(\xi_x) + \text{sign}(\xi_y) + 2] \quad (4.23)$$

Nous constatons que ce filtre permet de sélectionner 3 quadrants sur 4 (avec des poids différents). Si nous voulons faire apparaître une notion de transformée de Hilbert 2D, nous pouvons écrire le gain fréquentiel ainsi introduit de la manière suivante

$$H(\xi_x, \xi_y) = \frac{1}{2} (1 + j[-\frac{1}{2}j(\text{sign}(\xi_x) + \text{sign}(\xi_y))]) \quad (4.24)$$

et ainsi le noyau de la transformation de Hilbert 2D correspondante est égal à :

$$K(\xi_x, \xi_y) = -\frac{1}{2}j(\text{sign}(\xi_x) + \text{sign}(\xi_y)) \quad (4.25)$$

Dans le domaine temporel, ce noyau s'écrit alors :

$$k(x, y) = \delta(y)v.p. \left(\frac{1}{\pi x} \right) + \delta(x)v.p. \left(\frac{1}{\pi y} \right) \quad (4.26)$$

Notons que nous voyons ainsi apparaître la notion de transformées de Hilbert partielles dans les directions x et y dans le domaine spatial par :

$$s_{\mathcal{H}_1}(x, y) = s(x, y) * \left(\frac{\delta(y)}{\pi x} \right) \quad s_{\mathcal{H}_2}(x, y) = s(x, y) * \left(\frac{\delta(x)}{\pi y} \right) \quad (4.27)$$

avec $\delta(x)$ et $\delta(y)$ les distributions de Dirac dans les directions x et y . Ce qui correspond dans le domaine fréquentiel à :

$$S_{\mathcal{H}_1}(\xi_x, \xi_y) = -jS(\xi_x, \xi_y)\text{sign}(\xi_x) \quad S_{\mathcal{H}_2}(\xi_x, \xi_y) = -jS(\xi_x, \xi_y)\text{sign}(\xi_y) \quad (4.28)$$

L'inconvénient de cette autre solution est de ne pas traiter également toutes les composantes. En effet, si le signal analysé fait partie d'un des deux quadrants contenant une fréquence "négative", le signal après transformation resté inchangé, et donc il ne devient pas complexe, comme attendu.

Afin de généraliser l'approche fréquentielle liée à la définition du signal analytique, une solution peut être de couper le plan en deux, ou plus généralement pour des signaux multi-dimensionnels, il s'agit de définir un hyper-plan découpant le domaine fréquentiel, puisque la transformée de Fourier d'une fonction à valeur réelle a une symétrie hermitienne et présente donc une redondance d'information sur la moitié de l'espace [254]. Dans ce cas, la fonction de transfert du filtre qui permet de transformer le signal réel en signal analytique est définie par :

$$H(\xi_x, \xi_y) = \frac{1}{2}(1 + \text{sign}(a\xi_x + b\xi_y)) \quad (4.29)$$

avec $a\xi_x + b\xi_y = 0$ la droite coupant le plan en deux. Nous constatons qu'une composante réelle aura toujours ces deux impulsions situées de part et d'autre de cette droite, et donc l'une des deux sera annulée lors de l'application du filtre. La seule limite concerne les composantes éléments de la ligne (tout comme la fréquence nulle en 1D) qui vont rester inchangés. C'est pourquoi une dernière famille de solution va être étudiée par Grassin et al. avec un découpage suivant un demi-plan non symétrique [97].

En pratique, il est souhaitable de lier le choix de la droite au contenu spectral de l'image à analyser. Il faut effectivement éviter de couper une composante spectrale par la droite, car cette composante se retrouve alors en deux morceaux séparés.

Dans le cadre de nos travaux, nous nous sommes intéressés aux dernières évolutions concernant la définition d'un signal analytique 2-D qui utilisent la notion de quaternion.

4.1.3 Extension de la notion de signal analytique par transformée de Fourier quaternionique

4.1.3.1 Transformée de Fourier Quaternionique pour les images niveaux de gris

Ces dernières années s'est développée une nouvelle approche reposant sur les quaternions. Cette approche repose sur la transformée de Fourier quaternionique définie par Bülow [35], que nous avons déjà citée dans le chapitre précédent :

Soit s une fonction discrète à deux variables m et n à valeurs quaternioniques, une transformée de Fourier quaternionique (ou TFQ) discrète est une fonction à valeurs quaternioniques pouvant s'écrire de la manière suivante :

$$S[o, p] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-2\mu_1 \pi \frac{om}{M}} s[m, n] e^{-2\mu_2 \pi \frac{pn}{N}} \quad (4.30)$$

avec μ_1 et μ_2 des quaternions unitaires purs.

Nous rappelons que dans le cadre des travaux de Bulow, cette TFQ sera utilisée sur des signaux à valeurs réelles qui correspondent à l'information d'intensité des pixels en niveaux de gris. Bülow a défini cette transformée de Fourier en associant la direction horizontale m à la première exponentielle décrite avec le quaternion i et la direction verticale n à la deuxième exponentielle liée au quaternion j , afin de faire apparaître des propriétés de symétries qui sont associées au spectre obtenu par cette transformée. Donc dans ce cas, $\mu_1 = j$ et $\mu_2 = j$ et s est à valeurs dans $\mathbb{R}$.

Cette définition permet alors notamment de mettre en évidence des propriétés de symétries associées au spectre. Plus précisément, il est bien connu que l'on peut écrire n'importe quelle fonction comme la somme d'une fonction paire et d'une fonction impaire $s = s_p + s_i$ et que ces deux parités se dissocient dans le domaine de Fourier. L'idée de Bülow est d'étendre ces notions de symétries au cas des signaux 2-D, ce qui permettra d'introduire une notion de signal analytique 2-D.

Un signal 2-D s se décompose en effet selon les parties paires et impaires le long des axes horizontal m et vertical n : $s = s_{pp} + s_{ip} + s_{pi} + s_{ii}$ avec s_{pp} correspondant à la partie de s qui est paire le long de m et n , s_{ip} correspondant à la partie qui est impaire le long de m et paire le long de n , etc.

Nous illustrons ci-dessous comment les symétries spatiales vont se distribuer dans la transformée de Fourier Quaternionique de Bülow pour un signal à valeurs réelles :

$$S^q[o, p] = S_{pp}^q[o, p] + iS_{ip}^q[o, p] + jS_{pi}^q[o, p] + kS_{ii}^q[o, p] \quad (4.31)$$

On a une caractérisation de toutes les symétries spatiales possibles en 2D séparées dans le spectre quaternionique.

Les coefficients spectraux peuvent s'écrire sous une forme polaire, comme nous l'avons vu dans le chapitre précédent :

$$q = |q| e^{i\phi} e^{k\psi} e^{j\theta} \text{ avec } (\phi, \theta, \psi) \in [-\pi, \pi] \times [-\pi/2, \pi/2] \times [-\pi/4, \pi/4] \text{ pour une fréquence donnée.}$$

Nous voyons apparaître un module et trois phases. Nous avons constaté que dans le cadre des images couleur, l'analyse de l'information portée par la TFQ de Sangwine posait certaines difficultés.

Dans le cadre des images niveaux de gris, le résultat de la TFQ de Bülow exprimée sous une forme polaire permet de faire apparaître une propriété importante, qui était absente dans la transformée de Fourier complexe 2-D, à savoir la propriété de translation.

Il est bien connu en 1-D que lorsque l'on translate un signal, la phase du spectre est modifiée proportionnellement. Cette même opération appliquée sur une image entraîne une ambiguïté, à savoir qu'une

translation selon m ou n se traduit dans la même phase, puisque unique. La transformée TFQ lève cette ambiguïté :

Soient les signaux 2D s_1 et s_2 , et leurs QFT's $S_1 = |S_1|e^{i\varphi_1}e^{j\theta_1}e^{k\psi_1}$ et $S_2 = |S_2|e^{i\varphi_2}e^{j\theta_2}e^{k\psi_2}$.

On a la relation suivante :

$$s_2(x, y) = s_1(x - r_x, y - r_y) \Leftrightarrow \begin{pmatrix} \varphi_2(\xi_x, \xi_y) \\ \theta_2(\xi_x, \xi_y) \end{pmatrix} = \begin{pmatrix} \varphi_1(\xi_x, \xi_y) - 2\pi\xi_x r_x \\ \theta_1(\xi_x, \xi_y) - 2\pi\xi_y r_y \end{pmatrix} \quad (4.32)$$

qui correspond bien à la généralisation de la propriété de translation 1-D.

Le troisième angle ψ est plus difficile à interpréter, il semble plus en lien avec la notion de structures. Nous discuterons plus précisément de l'information portée par cette troisième phase lors de l'analyse du signal analytique 2-D.

Enfin, nous décrivons une dernière propriété de cette nouvelle transformée de Fourier, propriété qui est fondamentale pour la définition d'un signal analytique 2-D. En effet, il existe une généralisation du concept de symétrie Hermitienne au cas 2-D dans le domaine quaternionique. La notion de conjugaison quaternionique est un cas particulier de l'involution introduite par Chernov [52] :

$$\begin{aligned} \alpha(q) &= -iqi = q_0 + iq_1 - jq_2 - kq_3 \\ \beta(q) &= -jqj = q_0 - iq_1 + jq_2 - kq_3 \\ \gamma(q) &= -kqk = q_0 - iq_1 - jq_2 + kq_3 \end{aligned} \quad (4.33)$$

Notons que Ell et Sangwine ont proposé dans [80] une étude sur la généralisation de ce concept d'involution.

A partir de ces relation, Bülow introduit une généralisation du concept de symétrie hermitienne :

soit s une fonction discrète à deux variables m et n à valeurs quaternioniques, alors s présente une symétrie hermitienne au sens des quaternions si

$$\begin{aligned} s[n, -m] &= \alpha(s[n, m]) \\ s[-n, m] &= \beta(s[n, m]) \\ s[-n, -m] &= \gamma(s[n, m]) \end{aligned} \quad (4.34)$$

Dans ce cas on peut montrer que la TFQ d'un signal réel présente une symétrie hermitienne au sens des quaternions, ceci grâce à la distribution des différentes symétries spatiales dans le domaine fréquentiel. Cela veut dire que toute l'information fréquentielle peut être déduite à partir de l'un des quadrants de la TFQ. Dans ce cas, Bülow propose tout simplement de reprendre la définition de Hahn pour définir le signal analytique quaternionique 2D.

4.1.3.2 Le signal analytique quaternionique 2D.

D'après les propriétés de symétrie présentes dans le spectre de Fourier quaternionique d'une image niveaux de gris, on constate que toute l'information n'est portée que par un seul quadrant. Nous allons donc pouvoir définir le signal analytique 2-D en ne conservant que le quadrant des fréquences positives.

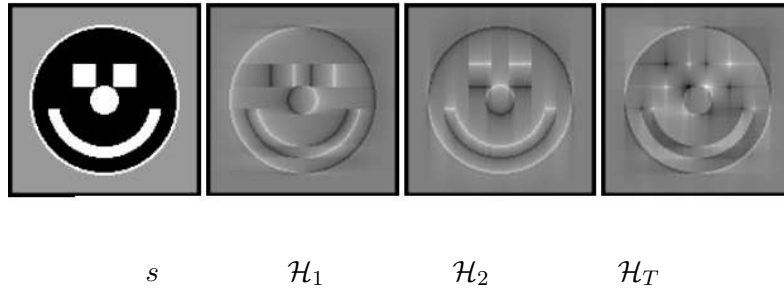


FIGURE 4.1 – Illustration du signal analytique Quaternionique [223]

Soit s un signal réel à deux dimensions et S^q sa transformée de Fourier quaternionique, le signal analytique quaternionique est défini dans le domaine fréquentiel par [35] :

$$S_A^q(\xi_x, \xi_y) = (1 + \text{sign}(\xi_x))(1 + \text{sign}(\xi_y))S^q(\xi_x, \xi_y) \quad (4.35)$$

Si l'on reprend les différentes transformées de Hilbert introduites précédemment, l'expression du signal analytique quaternionique $s_A^q(x, y)$ dans le domaine spatial est ainsi définie par :

$$s_A^q(x, y) = s(x, y) + \mathbf{HT}_x[s](x, y)i + \mathbf{HT}_y[s](x, y)j + \mathbf{HT}_{xy}[s](x, y)k \quad (4.36)$$

Comme pour le signal analytique 1D, la partie réelle du signal analytique f_A est constituée du signal f lui-même, et les fréquences positives suivant les deux axes du spectre du signal analytique quaternionique sont les seules non nulles. Nous avons bien une extension dans le concept du signal analytique 1-D.

Bülow définit à partir de ce nouveau signal analytique quaternionique l'amplitude et la phase instantanée :

$$\begin{cases} |s_A^q(x, y)| = \sqrt{s^2(x, y) + \mathbf{HT}_x[s]^2(x, y) + \mathbf{HT}_y[s]^2(x, y) + \mathbf{HT}_{xy}[s]^2(x, y)} \\ \phi(s_A^q(x, y)) = (\phi(x, y), \theta(x, y), \psi(x, y)) \end{cases} \quad (4.37)$$

On peut constater que la phase est composée maintenant de trois angles, ce qui peut compliquer son interprétation, par exemple comment exprimer le lien avec une notion de fréquence instantanée ? Nous reviendrons par la suite sur ce point.

Nous illustrons sur la figure 4.1 les différentes composantes de ce signal analytique Quaternionique. Nous avons repris pour cela l'image de test construite par R. Soulard, notamment pour une présentation dans le cadre d'une action du GDR [223], qui contient différentes ruptures 2-D : ruptures horizontales, verticales, des coins, des lignes courbes ... Cette image permet de mettre en évidence les particularités des différentes transformées de Hilbert 2-D.

Nous constatons bien évidemment la forte directionnalité des différentes composantes. La première partie imaginaire est associée au calcul d'un signal analytique sur la direction horizontale, donc nous mettons plutôt en évidence des discontinuités verticales. Le raisonnement est inverse en ce qui concerne la deuxième composante imaginaire. Enfin, la dernière partie imaginaire est majoritairement associée aux informations diagonales. Il est intéressant de remarquer que nous avons une lecture similaire d'une décomposition en ondelettes 2-D selon un schéma classique séparable, et donc probablement des critiques similaires à formuler. Nous aborderons ce point dans la deuxième partie de ce chapitre.

Comme dans le cadre 1-D, l'analyse de l'information associée au signal analytique nécessite un support fréquentiel limité. C'est pourquoi, Bülow a proposé une lecture de ce nouveau signal analytique à travers

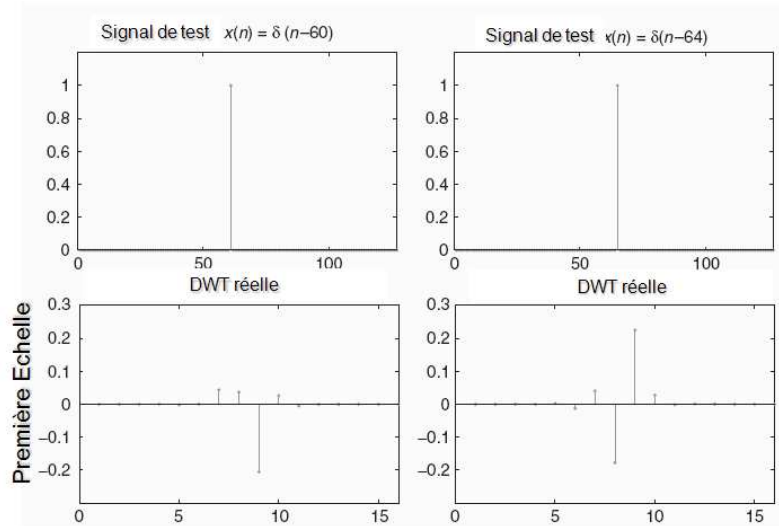


FIGURE 4.2 – Comportement de la transformée en ondelettes décimée face à une translation

une transformée de Gabor Quaternionique, calculée par fenêtrage de la transformée de Fourier. Les liens entre les décompositions en ondelettes et la décomposition de Gabor sont étroits, c'est tout naturellement que nous nous sommes posés la question des ondelettes quaternioniques. Cela fait l'objet de la suite de ce chapitre.

Maintenant que nous avons introduit les concepts qui vont servir de base pour définir les transformées en ondelettes hypercomplexes, nous proposons de revoir ce qui a constitué le point de départ de cet axe, à savoir les ondelettes complexes.

4.2 ondelettes Complexes

4.2.1 Principe

L'une des premières motivations dans la définition d'une décomposition en ondelettes complexes est la recherche de l'invariance par translation. En effet, comme nous l'avons vu, si nous comparons la représentation après transformée en ondelettes classique de deux signaux identiques décalés d'un échantillon, alors nous constatons que les coefficients issus des deux transformées sont différents (figure 4.2). Ce problème bien connu est dû au sous-échantillonnage d'un facteur 2, qui ne conserve pas les mêmes échantillons lorsque l'on décale le signal. Certes la reconstruction est toujours parfaite, et le sous-échantillonnage permet une représentation compacte, mais l'analyse peut être insatisfaisante.

Dans la littérature, l'une des premières solutions à avoir été envisagée a consisté à éliminer ce sous-échantillonnage, et donc de dilater les filtres entre chaque échelle : c'est la famille des décompositions en ondelettes non-décimée (ou "translation Invariante") que nous utiliserons dans le cadre du tatouage. Il est bien sûr tout à fait possible de définir une reconstruction parfaite, d'autant plus que les conditions sur les filtres sont simplifiées puisque nous avons éliminé le sous-échantillonnage. Il en résulte une décomposition

invariante par translation mais aussi redondante, d'un facteur dépendant du nombre d'échelles de décompositions.

Ensuite, comme nous l'avons vu, la décomposition en ondelettes 2-D souffre d'un manque de directionnalité et ces dernières années de nombreuses décompositions ont été proposées afin d'offrir une souplesse d'analyse accrue. Certaines de ces techniques ont une approche spatiale à base par exemple de schéma de décomposition Lifting, d'autres vont se positionner dans le domaine fréquentiel, c'est le cas des ondelettes complexes.

Enfin, l'un des derniers éléments expliquant l'introduction des ondelettes complexes est le caractère oscillatoire des coefficients autour d'une discontinuité. En effet, dans un proche voisinage d'une rupture la décomposition en ondelettes va traduire cette rupture par un certains nombres de coefficients oscillants autour de 0.

La solution proposée à ces différentes limites dans le cadre des ondelettes complexes repose sur une lecture de l'analyse de Fourier. En effet, la transformée de Fourier ne souffre pas des ces différents défauts. Tout d'abord le module des coefficients de Fourier est invariant par translation, le déplacement étant codé dans la phase des coefficients. Ensuite les éléments de base associés à la transformée de Fourier permettent une analyse très directionnelle. Enfin, le module de Fourier n'oscille pas mais se traduit sous la forme d'une enveloppe en général d'allure régulière.

La différence entre ces deux transformées se situe sur la définition des fonctions analysantes, à savoir que la fonction d'ondelettes est une fonction réelle oscillante, alors que la fonction analysante associée à Fourier est une exponentielle complexe constituée de deux composants formant une paire de Hilbert, correspondant à un signal analytique (l'information fréquentielle n'est portée que par la partie positive) :

$$e^{j2\pi\xi t} = \cos(2\pi\xi t) + j \sin(2\pi\xi t)$$

Pour répondre aux différentes critiques formulées sur la transformée en ondelettes, l'idée est donc de construire une transformée en ondelettes définie à partir d'une fonction d'ondelette complexe :

$$\psi_c(t) = \psi_r(t) + j\psi_i(t). \quad (4.38)$$

De plus en reprenant le concept lié à la transformée de Fourier, ψ_r et ψ_i doivent former une paire de Hilbert. Dans ce cas, l'ondelette complexe est un signal analytique, et la fonction d'échelle complexe va être définie de la même façon. Nous avons alors une analyse à partir de fonction analytique qui a priori permet de contenir ou d'annuler les trois défauts cités.

4.2.2 Déploiement numérique des ondelettes complexes

On peut trouver depuis 2000 dans la littérature de nombreux papiers sur le développement numérique d'une décomposition en ondelettes complexes. Nous ne donnons ici que quelques éléments nécessaires à la suite du propos.

Une stratégie consiste à calculer tout d'abord une transformée de Hilbert sur les données à analyser. Ainsi, en calculant la transformée en ondelettes réelle sur le nouveau signal analytique, on obtient une décomposition complexe. Cependant, comme nous l'avons vu dans la première partie, la transformée de Hilbert est associée à un filtre de réponse impulsionnelle infinie. Lors de son utilisation dans une décomposition en ondelettes, il en résulte un accroissement des supports des fonctions analysantes.

La seconde stratégie propose l'utilisation en parallèle de deux projections dans des bases orthogonales ou biorthogonales. La première transformée en ondelettes décimée calcule la partie réelle alors que la seconde fournit la partie imaginaire. Les deux bancs de filtres utilisent des filtres différents que nous noterons h_0, g_0 pour la partie réelle et h'_0, g'_0 pour le banc de filtres associé à la partie imaginaire. Chacun de ces bancs de filtres vérifie les conditions de reconstruction parfaite, c'est-à-dire qu'ils peuvent être utilisés indépendamment l'un de l'autre dans une analyse classique en ondelette. Cependant la contrainte supplémentaire est que la combinaison des deux sorties des filtrages donne des coefficients complexes approximativement analytiques. Cela se traduit par la contrainte que l'ondelette associée au second banc de filtres doit être approximativement la transformée de Hilbert de la première ondelette. C'est l'algorithme de décomposition Dual-Tree développé par Kingsbury [118].

Notons que dans ce cas la transformée présente une redondance de deux puisque nous utilisons deux transformées décimées. L'une des particularités de cette approche est qu'elle ne nécessite pas d'implanter une transformée de Hilbert, puisqu'elle est intégrée en amont dans la construction du banc de filtres. Le principe du dual-tree est facile à implanter et rapide car il utilise le même algorithme que la transformée classique. La seule difficulté réside dans la définition des filtres afin qu'ils soient associés à deux fonctions d'ondelettes approximant une paire de Hilbert. En effet, les bases discrètes proposées dans le cadre de la transformée décimée ne peuvent pas vérifier exactement les contraintes associées à l'arbre double.

Ces dernières années différentes stratégies pour la construction de filtres adaptés à l'arbre double ont été développées, notamment à travers la traduction de la condition sur les fonctions ondelettes en condition sur les filtres. L'exemple le plus connu est le travail de Selesnick [210, 211]. Nous débutons un banc de filtres h_0, g_0 à reconstruction parfaite, à savoir que

$$g_0[n] = (-1)^n h_0[D - n] \quad (4.39)$$

avec D un entier pair. Selesnick [210] a alors montré que si l'on définit le second filtre passe-bas par :

$$H'_0(\omega) = e^{-j0.5\omega} H_0(\omega) \quad (4.40)$$

alors, les deux ondelettes forment une paire de Hilbert. Cependant, par conséquence le second filtre passe-haut vérifie

$$G'_0(\omega) = -j \operatorname{sgn}(\omega) e^{j0.5\omega} G_0(\omega) \quad (4.41)$$

et cette condition se traduit dans le domaine temporel par

$$g_0[n] = h_0[n - 0.5]. \quad (4.42)$$

Puisque g_0 et h_0 sont des séquences uniquement définies sur une grille entière, nous constatons que cette condition doit être adaptée et donc approximée. Il en découle que dans les définitions numériques de l'algorithme Dual-Tree la condition dite de délai sera approchée et donc l'ondelette complexe résultante ne sera que approximativement analytique.

Pour résumer, la transformée complexe présente une redondance de 2, afin d'approximer l'invariance par translation et est associée à deux décompositions parallèles, correspondant à l'utilisation de deux filtres déphasés, qui vont former approximativement une paire de Hilbert, et donc un filtre complexe analytique. La

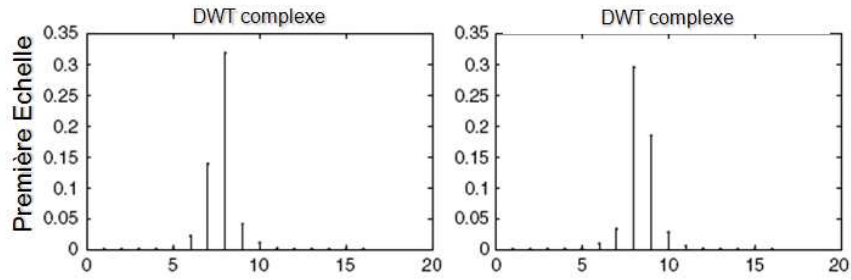


FIGURE 4.3 – Comportement de la transformée en ondelettes complexe face à la translation de la figure 4.2

combinaison des deux sorties des filtrages donne des coefficients complexes approximativement analytiques, dont le module est quasi invariant par translation. Nous avons repris l'expérimentation de la figure 4.2 sur la figure 4.3 pour la transformée complexe. On remarque que les deux DWT sont fondamentalement différentes (figure 4.2), alors que le module des CWT se ressemble fortement, malgré le décalage entre les deux signaux.

4.2.3 Extension 2D : Dual-Tree 2D

Nous décrivons maintenant l'extension 2-D proposée par Kingsbury [212] qui a pour objectif d'adapter la notion de transformée analytique à l'image. Les décompositions ont des propriétés d'analyse directionnelle mais vont échouer en ce qui concerne la construction de vrai paire de Hilbert 2-D, c'est pourquoi nous nous sommes tournés vers la généralisation à travers les quaternions, comme nous le verrons.

Comme pour les ondelettes classiques, l'extension va se construire par une définition des fonctions analysantes avec un produit tensoriel des fonctions 1D. Si l'on reprend le schéma de construction des ondelettes classiques 2-D :

$$\begin{aligned}\phi(x, y) &= \phi(x)\phi(y) \\ \psi_1(x, y) &= \phi(x)\psi(y) \\ \psi_2(x, y) &= \psi(x)\phi(y) \\ \psi_3(x, y) &= \psi(x)\psi(y)\end{aligned}$$

Si nous reprenons maintenant ce schéma de construction avec ψ_c une ondelette complexe (quasi analytique) telle que nous l'avons décrit précédemment, nous obtenons alors la fonction analysante suivante :

$$\begin{aligned}\psi(x, y) &= \psi_c(x)\psi_c(y) = [\psi_r(x) + j\psi_i(x)][\psi_r(y) + j\psi_i(y)] \\ &= \psi_r(x)\psi_r(y) - \psi_i(x)\psi_i(y) + j\psi_r(x)\psi_i(y) + j\psi_i(x)\psi_r(y)\end{aligned}\tag{4.43}$$

Puisque la fonction ψ_c est quasi-analytique, avec un support fréquentiel uniquement sur la partie positive de l'axe fréquentiel, la fonction ondelette 2-D complexe résultant du produit tensoriel ne va avoir un support non nul que sur un quadrant du domaine 2-D. Nous avons alors une ondelette permettant une analyse orientée. Nous illustrons cette construction sur la figure 4.4 en représentant notamment le fenêtrage fréquentiel

associé. Notons que dans cette figure nous reprenons le même mode de schématisation que Kingsbury et Selesnick dans [212] qui nous semble très clair.

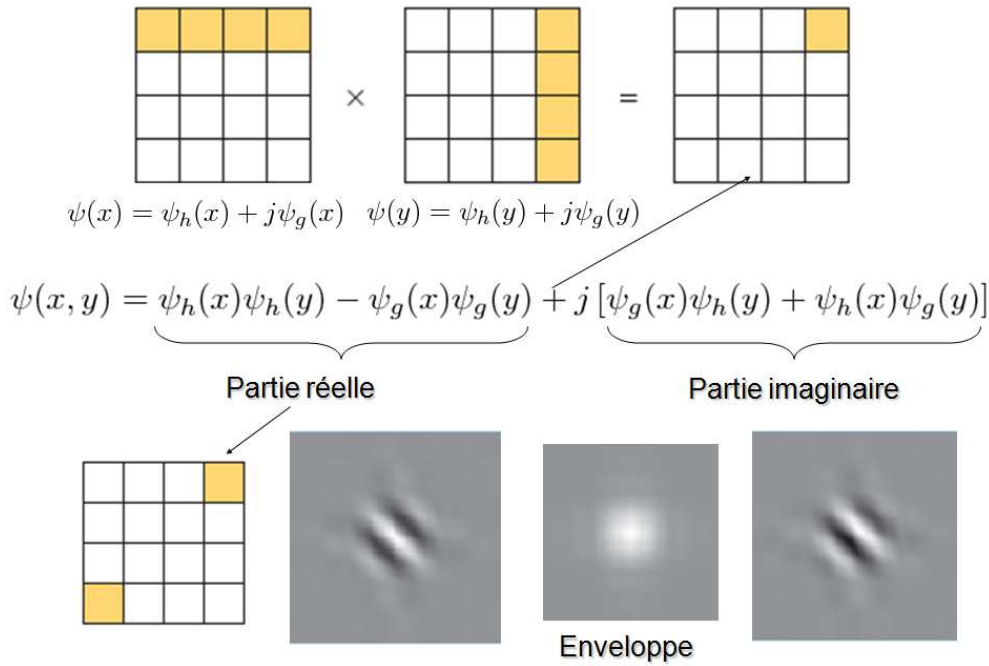


FIGURE 4.4 – Transformée en ondelettes complexe 2-D et découpage fréquentiel associé

Il est intéressant de remarquer que, dans cette construction, si l'on prend la partie réelle, nous obtenons la fonction $\psi_{rel}(x, y) = \psi_r(x)\psi_r(y) - \psi_i(x)\psi_i(y)$ qui bien sûr n'est plus une ondelette analytique mais bien une ondelette à valeurs réelles mais qui propose une analyse directionnelle comme nous l'illustrons sur la seconde partie de la figure 4.4. C'est d'ailleurs un des axes mis en avant par Kingsbury dans sa construction.

En terme de calcul la transformée associée s'obtient, en ce qui concerne $\psi_{rel}(x, y)$, par l'application du filtre g_0 suivant les lignes et les colonnes et l'on obtient $\psi_r(x)\psi_r(y)$, puis du filtre g'_0 suivant les lignes et les colonnes et l'on obtient $\psi_i(x)\psi_i(y)$ et enfin de simplement combiner les résultats des deux filtrages.

Si nous voulons une autre direction d'analyse nous pouvons mettre en place une autre combinaison, comme par exemple

$$\begin{aligned}\psi(x, y) &= \psi_c(x)\bar{\psi}_c(y) = [\psi_r(x) + j\psi_i(x)][\psi_r(y) - j\psi_i(y)] \\ &= \psi_r(x)\psi_r(y) + \psi_i(x)\psi_i(y) + j\psi_r(x)\psi_i(y) - j\psi_i(x)\psi_r(y)\end{aligned}\quad (4.44)$$

De nouveau la fonction ondelette 2-D complexe résultant du produit tensoriel va avoir un support non nul sur un autre quadrant du domaine 2-D. Nous avons alors une ondelette permettant une autre orientation d'analyse. Nous illustrons cette nouvelle construction sur la figure 4.5 en représentant notamment le fenêtrage fréquentiel associé.

Et comme pour la première construction, nous pouvons extraire par exemple la partie réelle afin d'avoir une ondelette à valeurs réelles mais directionnelles $\psi_{rel}^2(x, y) = \psi_r(x)\psi_r(y) + \psi_i(x)\psi_i(y)$.

Si nous voulons compléter la représentation, il suffit maintenant de calculer les combinaisons avec une "partie passe-bas" et donc la fonction d'échelle. Nous obtenons ainsi différentes combinaisons de fonctions d'ondelettes permettant de paver entièrement le domaine fréquentiel 2-D :

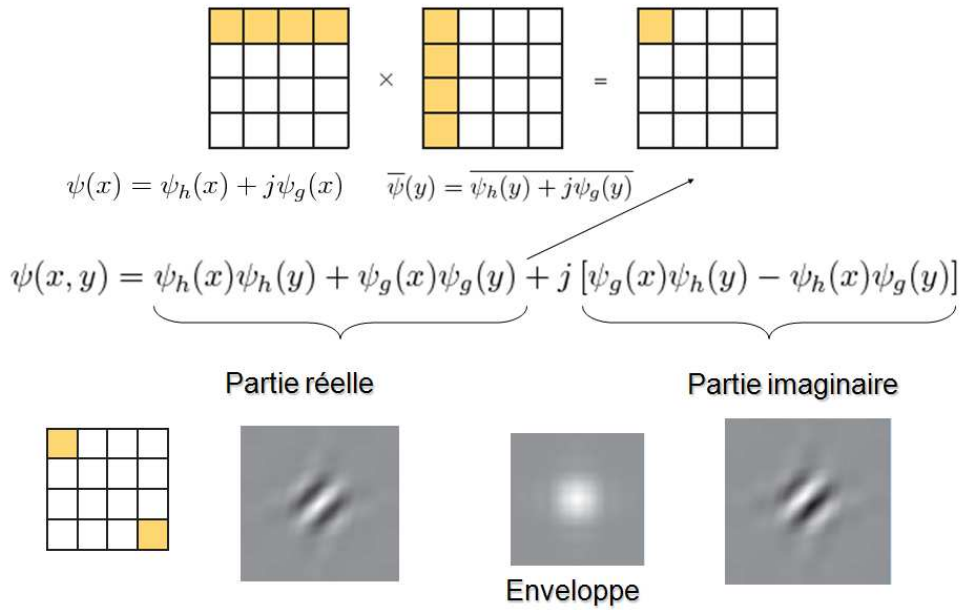


FIGURE 4.5 – Deuxième transformée en ondelettes complexe et découpage fréquentiel associé

$$\begin{bmatrix} \psi_c(x)\psi_c(y); \psi_c(x)\bar{\psi}_c(y); \\ \phi_c(x)\psi_c(y); \phi_c(x)\bar{\psi}_c(y); \\ \psi_c(x)\phi_c(y); \psi_c(x)\bar{\phi}_c(y). \end{bmatrix} \quad (4.45)$$

Nous obtenons ainsi un arbre dual d'ondelettes orientées complexes. Et nous pouvons définir un arbre dual d'ondelettes orientées réelles en prenant la partie réelle de ces différentes fonctions.

Il existe différentes applications au Dual-Tree CWT. En effet, cette nouvelle transformée offre une représentation plus précise en ce qui concerne les orientations d'analyse. De plus apparaît une nouvelle information à savoir la phase des coefficients que l'on calcule en chaque position. Cette phase correspond, selon Selesnick et al. [212] à l'éloignement du détail par rapport à la position du coefficient, perpendiculairement à l'orientation. Selon les auteurs, l'utilisation de la phase permet de calculer efficacement un flot optique entre deux images consécutives d'une séquence [150]. Il existe aussi de nombreuses applications en débruitage car l'invariance par translation permet d'obtenir de meilleurs résultats que la transformée classique décimée, le premier datant par exemple de 1998 [117]. Enfin, comme nous le verrons dans la suite de ce chapitre, du fait de ses bonnes propriétés la CWT 2-D est utilisée dans un contexte de reconnaissance de textures.

Toutefois, en général les travaux utilisent plutôt la notion de module et à notre connaissance, la phase est très peu utilisée dans la littérature, et semble encore une notion assez obscure, à la fois par la difficulté de son interprétation, et par les difficultés liées à la manipulation numérique d'une valeur angulaire.

4.3 Ondelettes quaternioniques

4.3.1 Principe

Bien que l'algorithme dual-tree 2D complexe présente soit associé à des fonctions d'ondelettes présentant de nombreuses qualités comme la directionnalité, la quasi invariance par translation, on peut s'interroger sur la réponse apportée par cette construction par rapport à la problématique posée à l'origine, à savoir la construction d'une fonction d'ondelette analytique 2-D. Ce point n'a pas été abordé dans les articles fondateurs concernant les ondelettes complexes 2-D. Nous avons mené une réflexion sur ce point à partir des travaux de Bülow et la définition d'une ondelette 2-D réellement analytique n'est apparue que très récemment dans les papiers de Chan et al. [49].

Par exemple, si nous reprenons la définition de la première ondelette complexe 2-D

$$\psi(x, y) = \psi_r(x)\psi_r(y) - \psi_i(x)\psi_i(y) + j\psi_r(x)\psi_i(y) + j\psi_i(x)\psi_r(y) \quad (4.46)$$

Il faut alors s'interroger si $\psi(x, y)$ est un signal analytique 2-D. Nous retrouvons les interrogations développées précédemment autour de la définition d'un signal analytique 2-D. Nous constatons, que si nous notons $\psi_{rr}(x, y) = \psi_r(x)\psi_r(y)$ alors la fonction complexe définie ci-dessus peut s'écrire (avec l'intégration des notions de transformées de Hilbert totales et partielles déjà vues) :

$$\psi_1(x, y) = \psi_{rr}(x, y) - \mathbf{HT}_{xy}[\psi_{rr}](x, y) + j\mathbf{HT}_x[\psi_{rr}](x, y) + j\mathbf{HT}_y[\psi_{rr}](x, y) \quad (4.47)$$

Nous constatons que la fonction d'ondelette définie ainsi correspond à la définition de Hahn pour la construction du signal analytique, à savoir l'extension directe de la définition avec comme noyau de la transformation de Hilbert 2D correspondante est égal à :

$$K(\omega_x, \omega_y) = -\frac{1}{2}j(\text{sign}(\omega_x) + \text{sign}(\omega_y)) \quad (4.48)$$

Nous savons alors que cette définition est insuffisante du fait de la symétrie hermitienne, ce qui nécessite pour définir pleinement le signal analytique de considérer la seconde fonction :

$$\psi_2(x, y) = \psi_r(x)\psi_r(y) + \psi_i(x)\psi_i(y) + j\psi_r(x)\psi_i(y) - j\psi_i(x)\psi_r(y) \quad (4.49)$$

et ainsi de couvrir l'autre demi-plan. Ainsi le couple (ψ_1, ψ_2) correspond à l'ondelette analytique associée à ψ_{rr} .

En terme numérique, cela veut simplement dire que pour chaque position, nous collectons deux coefficients d'ondelettes complexes pour définir le coefficient analytique.

Comme nous l'avons abordé, la présence de deux nombres complexes pour définir le signal analytique entraîne une certaine ambiguïté dans les définitions des informations déduites. En effet, du fait de la structure, on calcule deux normes et deux angles dont il est difficile d'expliciter leur lien ou de proposer une stratégie de fusion. C'est pourquoi tout logiquement dans la suite des travaux de Bülow sur le signal analytique quaternionique a été proposée très récemment une décomposition en ondelettes analytique quaternionique par Chan et al. [49].

Pour mémoire, le signal analytique quaternionique est défini par :

$$s_A^q(x, y) = s(x, y) + \mathbf{HT}_x[s](x, y)i + \mathbf{HT}_y[s](x, y)j + \mathbf{HT}_{xy}[s](x, y)k \quad (4.50)$$

Et comme nous l'avons vu, puisque la TFQ d'un signal réel présente une symétrie hermitienne au sens des quaternions, toute l'information est contenue dans s_A^q .

De cette écriture, on peut tout simplement étendre la décomposition en ondelettes en se fixant simplement une fonction d'ondelette réelle 2-D et en calculant les différentes extensions de la transformée de Hilbert. La transformée en ondelettes analytique quaternionique est définie à partir de la fonction analysante suivante [49] :

$$\psi_A^q(x, y) = \psi_{rr}(x, y) + i\mathbf{HT}_x[\psi_{rr}](x, y) + j\mathbf{HT}_y[\psi_{rr}](x, y) + k\mathbf{HT}_{xy}[\psi_{rr}](x, y) \quad (4.51)$$

Puisque $\psi_i(x) = \mathbf{HT}[\psi_r](x)$, nous avons

$$\psi_A^q(x, y) = \psi_r(x)\psi_r(y) + i\psi_i(x)\psi_r(y) + j\psi_r(x)\psi_i(y) + k\psi_i(x)\psi_i(y). \quad (4.52)$$

Cette écriture se place directement dans la continuité des travaux de Bülw qui proposait une transformée de Gabor quaternionique (notamment pour obtenir des signaux analytiques de bande passante réduite) suivant le même principe. Nous reprenons brièvement le paragraphe présenté dans l'annexe concernant la construction d'un banc de filtres quaternioniques.

À notre connaissance actuellement nous pouvons trouver trois propositions de décompositions en ondelettes Quaternionique qui se rattache à l'écriture exposée ci-dessus :

- Bayro [23] a proposé une transformée en ondelettes quaternionique pour les images en niveaux de gris ($s \in \mathcal{L}^2([0..N-1]^2, \mathbb{R})$) avec le principe des filtres de Gabor modulés en quadrature. Les filtres utilisés sont définis par :

$$h[n_1, n_2] = g_\sigma[n_1, n_2]e^{2i\pi\xi_1 n_1}e^{2j\pi\xi_2 n_2}$$

avec σ l'écart-type de la fenêtre Gaussienne et ξ_1, ξ_2 les coordonnées fréquentielles. Nous constatons que cette définition correspond à une généralisation du travail de Bülw. Dans le papier cité, l'auteur utilise ce concept de phase pour estimer le flot géométrique de l'image analysée et que l'aspect reconstruction n'est pas discutée.

- Chan et al [48, 49] proposent une transformée en ondelettes quaternionique pour les images niveaux de gris suivant une extension directe du concept de la transformée en ondelettes complexe développée par Kingsbury. Ce travail est celui détaillé dans la section précédente. Selon les auteurs cette nouvelle décomposition est approximativement une transformée de Fourier quaternionique telle qu'elle est définie par Bülw mais localisée dans le temps. Comme pour le premier travail, les auteurs utilisent la notion de phase pour estimer le flot géométrique.
- Olhede et al [167] proposent la transformée en ondelettes hyperanalytique pour les images niveaux de gris. Ce travail est proche de celui de Chan : les quatre éléments de la base sont définis à partir de la transformée de Hilbert 1-D calculée sur les fonctions d'ondelettes 2-D obtenues par produit tensoriel et la transformée de Hilbert 1-D est appliquée suivant ligne et colonne. Notons qu'une seconde famille est décrite à partir des travaux de Felsberg et al [85], cette famille correspond aux ondelettes monogéniques que nous présenterons à la fin de ce chapitre.

Comme nous le constatons, tous ces travaux sont liés à ceux de Bülw et correspondent approximativement à une Transformée de Fourier Quaternionique de Bülw localisée dans le temps.

4.3.2 Expression numérique de la QWT

Une des grandes qualités de la proposition de Chan et al [48, 49] est de reposer, en plus d'une justification théorique, sur un algorithme numérique de calcul lié à l'algorithme Double Tree. En effet, pour chaque échelle de décompositions, la transformée en ondelettes quaternionique est associée à trois fonctions

$$\begin{aligned} {}^1\psi_A^q(x, y) &= \psi_r(x)\psi_r(y) + i\psi_i(x)\psi_r(y) + j\psi_r(x)\psi_i(y) + k\psi_i(x)\psi_i(y) \\ {}^2\psi_A^q(x, y) &= \psi_r(x)\phi_r(y) + i\psi_i(x)\phi_r(y) + j\psi_r(x)\phi_i(y) + k\psi_i(x)\phi_i(y) \\ {}^3\psi_A^q(x, y) &= \phi_r(x)\psi_r(y) + i\phi_i(x)\psi_r(y) + j\phi_r(x)\psi_i(y) + k\phi_i(x)\psi_i(y) \end{aligned} \quad (4.53)$$

Nous constatons qu'en terme calculatoire, cette nouvelle décomposition est bien similaire à l'arbre double dual complexe, puisque nous retrouvons les mêmes fonctions impliquées dans l'arbre double 2-D (ψ_r , ψ_i , ϕ_r et ϕ_i). Finalement, la transformée en ondelettes quaternionique est simplement une combinaison fine des coefficients issus de l'arbre double complexe 2-D. Prenons par exemple la fonction ${}^2\psi_A^q$, les coefficients associés à cette fonction d'ondelettes sont issus du filtrage de l'image par :

- $g_0(x)h_0(y)$ pour la partie réelle,
- $g'_0(x)h_0(y)$ pour la partie associée à i ,
- $g_0(x)h'_0(y)$ pour la partie associée à j ,
- $g'_0(x)h'_0(y)$ pour la partie associée à k .

D'une manière générale, la figure 4.6 proposée par R. Soulard dans [221] illustre très bien l'algorithme de calcul et le lien avec la transformée complexe. Nous retrouvons ce lien avec la transformée de Fourier quaternionique qui s'exprime comme deux transformées de Fourier complexes à travers une décomposition symplétique.

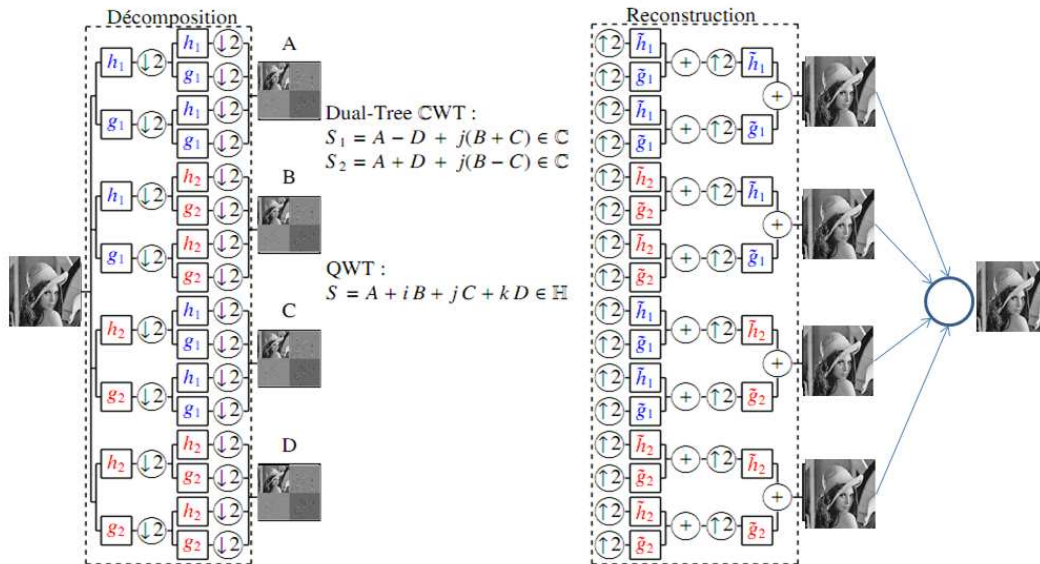


FIGURE 4.6 – Banc de filtres pour le calcul de la QWT et de la CWT

4.3.3 Information associée à la QWT

Tout d'abord, afin d'analyser l'information mise en valeur par la QWT, il faut considérer que chaque sous-bande de la QWT correspond à un signal analytique, associé à un signal réel 2D à "bande étroite" (du fait du filtrage passe-bande associé). Les coefficients d'une sous-bande s'expriment donc suivant les informations de module et de phases. Notons que la trame contient les basses fréquences et à ce titre doit être selon moi analysée différemment, comme nous le verrons dans les applications développées au sein de l'équipe.

La phase 1-D d'un signal analytique peut présenter un lien avec la structure géométrique des signaux. Cependant, les structures 2-D sont plus difficiles à caractériser. Comme nous l'avons vu dans le premier chapitre, une image est constituée de bords, de lignes, de coins, de courbes. En plus de l'instabilité numérique inhérente à l'utilisation d'information angulaire, l'interprétation est donc plus difficile

Tout d'abord, de manière analogue à l'analyse 1-D, le module détecte toujours la présence d'un élément particulier. Et de nouveau lorsque cette valeur sera faible, l'information de phase ne sera pas considérée.

La phase doit décrire sa structure, mais cette nouvelle phase 2-D est composée de trois angles (φ, θ, ψ) . En termes d'information, il faut dissocier les deux premiers termes de la phase (φ, θ) et la troisième phase ψ . Nous avons vu dans le cadre de Fourier que les deux premières phases codent les translations effectuées dans le domaine temporel, chacune dans une direction particulière. Puisque l'on peut créer un lien entre la transformée en ondelettes et une transformée de Fourier fenêtrée, ce principe de décalage donne une information sur les deux premières phases des coefficients de la QWT qui vont coder une information de position et de décalage selon x et y , comme la phase 1-D, mais aussi comme la phase de la transformée complexe.

Selon la littérature le troisième terme ψ va avoir pour fonction de coder les informations plus liées à l'image comme notamment des informations qui peuvent s'apparenter à de la texture.

Pour comprendre ψ , dans le cadre de la décomposition de Gabor, Bülöw a proposé des expérimentations sur des images synthétiques, notamment une image construite à partir de deux ondes, paramétrées par un terme λ :

$$s(x, y) = (1 - \lambda) \cos(2\pi\xi_1 x + 2\pi\xi_2 x) + \lambda \cos(2\pi\xi_1 x - 2\pi\xi_2 x) \quad (4.54)$$

On voit que le paramètre λ permet de régler l'importance de chacune des directions d'oscillation. Bülöw a mis en lumière une relation quasi linéaire entre le paramètre λ et la phase ψ des coefficients de Gabor quaternionique, indiquant ainsi le lien très fort entre cette troisième phase et des structures 2-D. R. Soulard a refait cette expérimentation pour la QWT dans le cadre de son travail de thèse. Nous constatons sur la figure 4.7 que nous retrouvons la même variation dans le cadre de la QWT. Cette expérimentation confirme bien que les constations faites dans le cadre de la décomposition de Gabor quaternionique peuvent s'étendre à la QWT, notamment les différentes pistes d'interprétation des phases soulevées par Bülöw.

Pour terminer cette section, nous proposons de revoir rapidement comment l'information associée à cette nouvelle décomposition est utilisée dans la littérature, à travers l'unique application abordée, hors nos travaux, à savoir l'estimation du flot optique.

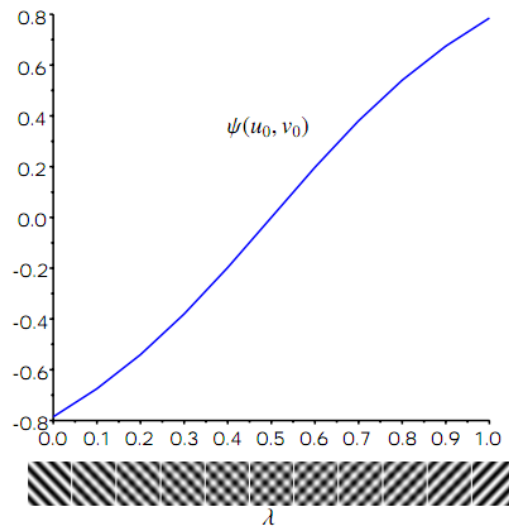


FIGURE 4.7 – Valeur de la phase ψ pour différentes valeurs de λ sur l'image test de Bülow.

4.3.4 Estimation de flot optique par QWT

L'estimation du flot optique consiste à étudier les déplacements locaux $(d_x(x, y), d_y(x, y))$ dans une séquence d'images. Cette estimation est la principale application que l'on trouve dans le cadre des travaux concernant les décompositions quaternionique (que ce soit avec Gabor ou les ondelettes). Le concept d'estimation du flot optique par des méthodes dites spectrales n'est pas nouveau notamment à travers la décomposition de Gabor complexe (par exemple [229]). Par contre, les ondelettes classiques ne pouvaient offrir de solution à cette problématique, à cause de leur variance par rapport aux translations. C'est pourquoi, il a été proposé d'utiliser les ondelettes complexes pour l'estimation de ce flot [150]. Ensuite dans le cadre des nouvelles transformées, l'idée a été d'utiliser la propriété de translation liée aux deux premières phases pour ainsi estimer le déplacement. Cette application a d'abord été proposée dans la thèse de T. Bülow [35] avec des filtres de Gabor quaternioniques (travail repris dans [24, 253]), puis les auteurs de la QWT l'ont adaptée aux ondelettes dans [49]. Notons que la méthode de Chan et al. exploite pleinement la structure de décomposition en ondelettes car l'estimation du flot sur une échelle est utilisée pour guider l'estimation à une échelle plus fine, et les différentes orientations sont exploitées.

Dans la suite de ce chapitre, nous allons maintenant nous attarder sur l'utilisation de ces différents concepts. Les travaux expérimentaux décrits sont effectués dans le cadre de la thèse de R. Soulard (thèse à soutenir en 2012), thèse que j'encadre. Comme nous le verrons, ces travaux sont très récents et pour certains posent plus de questions qu'ils n'apportent de réponses définitives.

4.4 L'analyse de textures par QWT

L'analyse de textures peut alimenter soit des problématiques de segmentation soit de discrimination, où l'on doit attribuer une classe à une texture inconnue. Dans le cadre quaternionique, la segmentation a été abordée par Bülow [35] avec des filtres de Gabor quaternioniques, l'approche proposée consiste simplement

à généraliser les stratégies classiques en ondelettes :

- calcul des différentes bandes spectrales de Gabor ,
- sélection pour chaque position de la bande de plus forte énergie,
- raffinement de la segmentation à partir des variations spatiales de la phase.

La dernière étape est l'originalité de l'approche quaternionique, et part de l'hypothèse que dans une zone homogène, la phase doit évoluer à peu près linéairement.

A travers ces expérimentations, Bülow montre que l'extension quaternionique permet d'améliorer la segmentation grâce aux trois angles de phase qui donnent plus d'information que les filtres de Gabor complexes.

4.4.1 Existant sur la discrimination quaternionique

Dans le cadre de la thèse de R. Soulard, une application de la QWT à la classification de textures a été étudiée [225]. L'objectif de cette problématique était de montrer la supériorité de cette transformée sur les ondelettes classiques, et ainsi de mettre en évidence les particularités des informations liées à la représentation : la robustesse du module invariant par translation et l'information complémentaire portée par la phase. Nous proposons de revoir les grands traits de la méthode.

La principale question que nous avons à trancher est bien sûr la construction du vecteur de mesures. Dans le cadre des ondelettes et de Gabor, il existe depuis de très nombreuses années, différents travaux proposant des solutions, nous pouvons par exemple citer les travaux de Laine et al.[126, 121] qui ont été parmi les premiers dans le cadre des ondelettes. Dans toutes ces approches basées ondelettes, une mesure d'énergie est en général faite à partir des coefficients d'ondelettes pour les différentes bandes et c'est cette mesure qui permet de faire la discrimination. Nous retrouvons une approche similaire dans le cadre des transformées quaternioniques.

En ce qui concerne les transformées complexes, dans la très grande majorité des cas, seule l'information de norme et donc d'énergie est utilisée à travers en général des calculs de moyenne et/ou d'écart-type, les différences entre les méthodes se situent dans leur utilisation : des mesures de moyenne et d'écart-type sont faites sur le module des 6 sous-bandes de la transformée complexe dans [104] (pour chaque échelle, uniquement sur la sous-bande de plus forte énergie), Celik et al. proposent dans [46] de calculer des variances et entropies, dans [109] une somme des énergies par bande est proposée etc. On constate que l'on retrouve une traduction directe des approches développées dans le cadre des ondelettes réelles.

L'étude de la phase de la transformée en ondelettes complexes n'est apparue que récemment. Dans [161], Kingsbury et al. étudient, pour l'estimation du flot optique, les liens entre les phases, ou plutôt les différences de phases, et l'on retrouve ce que nous avons présenté dans le cadre du flot optique.

A notre connaissance, seuls Vo et al proposent dans [246, 247] une modélisation de la phase de la transformée en ondelettes complexes. Les modèles utilisés font partie de la famille des distributions circulaires (Von Mises et Cauchy). Cette identification des distributions à travers l'estimation des paramètres est utilisée dans le cadre de la classification de texture, avec notamment la proposition de mesures de distance. Les résultats mettent en évidence l'apport en termes de qualité de reconnaissance de l'information de phases. Il nous semble que ce sont les travaux les plus aboutis à l'heure actuelle en ce qui concerne l'utilisation de la phase de la transformée complexe.

Enfin un travail très récent de Celik et al. propose lui aussi d'utiliser la phase de la transformée Complexe [47]. Toutefois, dans ce travail la phase est utilisée comme une mesure classique (avec par exemple une

notion de calcul de moyenne) ce qui nous semble discutable (les expérimentations menées par R. Soulard l'ont confirmé). Les résultats présentés dans [47] en termes de classification sont excellents mais une analyse en composantes principales est appliquée sur les mesures initiales, et cette étape doit selon nous gommer la faiblesse de la mesure faite sur les phases.

Dans le cadre de la transformée QWT, il n'existe à notre connaissance aucun travail concernant la classification de texture, si ce n'est ce que nous avons proposé dans [228], que je décris dans la section suivante.

4.4.2 Utilisation des mesures de norme et de phase de la QWT

Norme de la QWT Comme pour les ondelettes classiques, la première mesure que nous utilisons est la norme $M_{ij} = |q_{ij}|$ des coefficients d'ondelettes. Comme nous sommes dans une problématique de classification, nous effectuons une mesure globale sur chaque bande, avec l'hypothèse que la texture soit stationnaire. Nous proposons alors d'utiliser, pour estimer l'énergie par bande spectrale, classiquement les moments d'ordre 1 et 2 :

Moyenne :

$$\mu = \frac{1}{NM} \sum_{i,j} M_{ij}^2 \quad (4.55)$$

avec NM le nombre de pixels dans la sous-bande. Notons que dans les tests menés dans la thèse de R. Soulard, nous proposons d'utiliser une version normalisée $m = \frac{NM\mu}{E}$ avec E l'énergie de l'image moins celle associée à la trame.

Ecart-type :

$$m = \sqrt{\frac{1}{NM} \sum_{i,j} (M_{ij} - \mu)^2} \quad (4.56)$$

Phase de la QWT Nous avons vu que l'un des éléments nouveaux dans la transformée QWT est la présence des trois phases. Les deux premières phases φ et θ vont simplement coder une notion de décalage. A priori elles ne vont apporter aucune information pour la classification. Les expérimentations menées dans le cadre de la thèse de R. Soulard ont confirmé cela, puisque une classification utilisant uniquement les deux premières phases a obtenu des résultats catastrophiques.

Donc comme le suggérait Bülow [35], nous utilisons la troisième phase. Comme pour l'information de norme, seule une mesure globale est recherchée. En revanche, la caractérisation globale d'une mesure d'une phase n'est pas aussi simple. Ainsi que nous l'avons rappelé, les travaux utilisant la phase sont peu nombreux dans la littérature, même si l'on sort du cadre des ondelettes. Le concept que l'on retrouve le plus souvent lié à des mesures d'angle est la notion de "phase congruency" qui propose de détecter les éléments particuliers d'une image en analysant localement la phase de la transformée de Fourier (complexe à cette époque). On peut citer dans ce cadre le travail de Kovécí [122] par exemple. Cependant, il existe un lien entre cette notion de "phase congruency" et l'enveloppe complexe définie à partir de la transformée de Hilbert, et donc la majorité des auteurs ont finalement travaillé sur une mesure d'énergie et non d'angle. On peut encore citer les travaux de Blanchet et al. [25] sur la "Global Phase Coherence" qui se proposent d'intégrer

ce concept de phase de Fourier dans un schéma de variation totale afin de faire ressortir les structures de l'image.

A ce stade, il nous a semblé plus judicieux d'utiliser des mesures simples. Tout d'abord un simple calcul de somme n'est pas adapté pour un angle. De nouveau, les expérimentations menées dans le cadre de la thèse de R. Soulard ont confirmé cela : un paramètre basé sur une moyenne de la troisième phase a obtenu de très mauvais résultats (moins de 1% de reconnaissance). L'information qui est importante ne concerne pas la valeur de la phase mais sa variation car elle va donner une information sur la variation de la forme du signal analysé. La mesure sur la phase est donc :

Variation de la phase :

$$m = \sqrt{\frac{1}{MN} \sum_{i,j} (\psi_{ij} - \mu)^2} \quad (4.57)$$

avec $\mu = \frac{1}{MN} \sum_{i,j} \psi_{ij}$, avec NM le nombre de pixels dans la sous-bande.

Toutefois, nous avons vu que la mesure numérique de l'angle devient instable lorsque la norme est faible. C'est pourquoi, il est intéressant de pondérer le calcul de la variation de l'angle par la valeur de la norme en chaque position :

Variation pondérée de la phase :

$$m = \sqrt{\sum_{i,j} W_{ij} (\psi_{ij} - \mu_W)^2} \quad (4.58)$$

avec $\mu_W = \sum_{i,j} W_{ij} \psi_{ij}$ la moyenne pondérée selon la norme.

Combinaison des mesures de normes et de phases Enfin, tout logiquement il peut être intéressant de combiner les deux types de mesures, à savoir norme et phase. Dans les quelques travaux portant sur la transformée complexe, lorsque la phase est utilisée, elle est toujours associée à sa mesure de norme dans le vecteur de description de la texture. Il reste une difficulté, à savoir que ces deux mesures ont un domaine de variation bien différent. C'est un problème bien connu dans le cadre des algorithmes de discrimination qui se résout par l'introduction d'une métrique autre que la matrice identité dans le calcul des distances. Cependant comme nous le verrons dans les expérimentations la simple concaténation est à notre grande surprise tout à fait compétitive.

4.4.3 Utilisation de la QWT pour la classification de textures : mesures

Protocole Durant la thèse de R. Soulard, plusieurs bases de textures ont été utilisées :

- la base classique de Brodatz [32] qui contient 111 photos. Pour chaque image de cette base de taille $[640 \times 640]$, nous créons des exemplaires de la classe en découpant 25 imagerettes de taille $[128 \times 128]$,
- la base Outex (<http://www.outex.oulu.fi/>) qui est très souvent citée dans des problématiques de classification de texture. Dans cette base, deux groupes d'images ont été utilisés par R. Soulard ; TC12 qui contient 24 classes de 380 exemplaires, et TC14 avec 68 classes de 60 exemplaires.

Notons que pour le groupe "TC12", les différents exemplaires correspondent à la même texture mais suivant différents points de vue et différents éclairages. Pour le groupe "TC14", les images correspondent à un simple changement d'éclairage.

Ensuite puisque l'objectif n'est pas la définition d'un classifieur performant mais l'analyse de la nouvelle transformée, une procédure de discrimination très simple est mise en place : aucune étape de réduction de

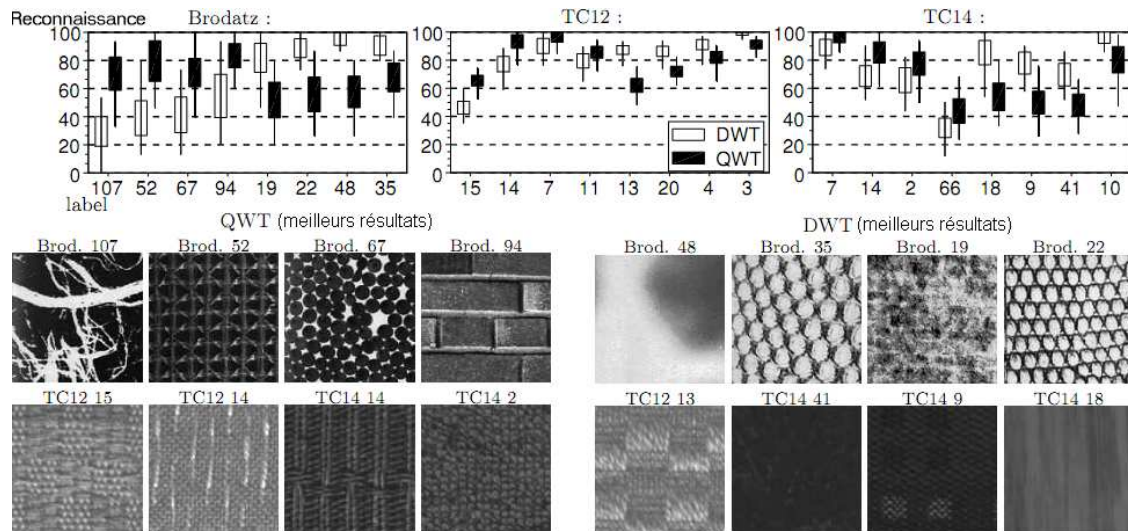


FIGURE 4.8 – Résultats de classification à partir de la mesure de module [228].

données ou de changements d'espace des paramètres n'est effectuée et un discriminateur le plus simple possible en terme de réglage tout en acceptant des frontières de classes irrégulières (nous avons choisi La méthode des k -plus proches voisins).

Enfin, la procédure de test doit être la plus indépendante possible de la procédure d'apprentissage, c'est pourquoi une mesure par cross-validation a été choisie. La base de test est découpée en deux groupes :

- une base d'apprentissage : N_A exemplaires labellisés ;
- une base de tests : N_T exemplaires inconnus .

A partir du second jeu de données, on peut alors estimer le taux de bonne détection pour chacune des classes. Pour obtenir une estimation sans biais du taux d'erreur, ce découpage est répété un certain nombre de fois.

Nous proposons maintenant de revoir les principaux résultats de ces nouveaux paramètres de classification de textures.

Tous les résultats présentés ici sont issus des expérimentations menées dans le cadre de l'article [228].

Utilisation du module de la DWT et QWT La première caractéristique que nous proposons d'étudier concerne le module. La figure 4.8 présente les résultats associés à cette mesure. L'axe des abscisses correspond aux labels des différentes classes, et l'axe des ordonnées correspond à l'indice de reconnaissances. Puisque l'expérimentation est effectuée 100 fois, nous indiquons la moyenne, l'écart-type et les valeurs extrêmes obtenues pour chaque classe des différentes bases. Globalement les deux décompositions avec cette mesure obtiennent des résultats sensiblement identiques. Sont donc affichés sur la figure 4.8, les résultats présentant les plus grandes différences (pour des problèmes de lisibilité). Et dans la même logique, en dessous des tableaux de mesures sont illustrées les images de texture présentant les meilleurs résultats pour les deux stratégies de classification. Les différences sont minimales entre les deux mesures ce qui est logique puisque dans les deux cas, nous sommes sur des mesures d'énergie par bandes de fréquences. Ceci étant pour expliquer les légères différences, les pistes avancées dans [228] sont :

- pour la QWT son invariance par translation donc son absence de sensibilité face au découpage en image et donc la translation du motif de texture,

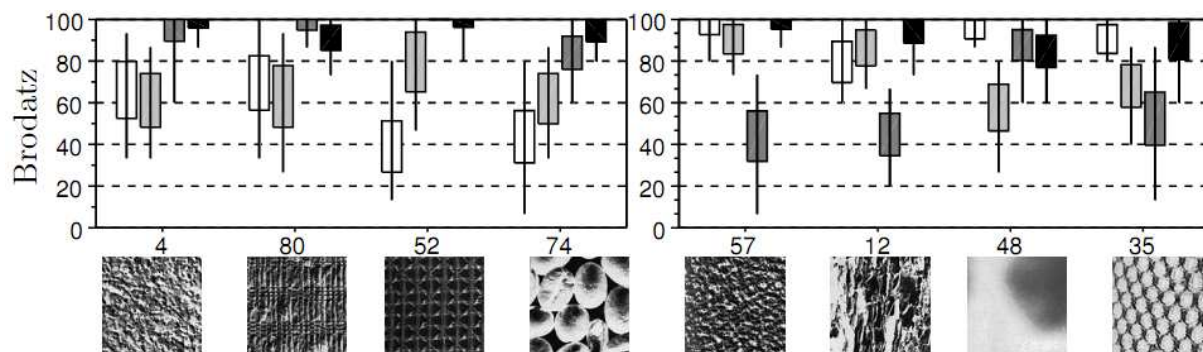


FIGURE 4.9 – Résultats de classification à partir de la mesure de module, de phase et des deux combinées : base Brodatz [228].

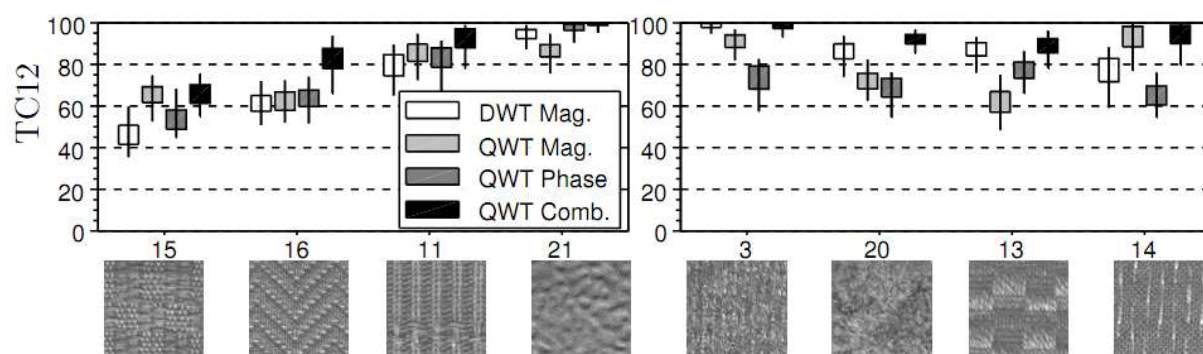


FIGURE 4.10 – Résultats de classification à partir de la mesure de module, de phase et des deux combinées : base TC12 [228].

- pour la DWT, son orthogonalité qui lui permet de mieux caractériser toutes les textures purement aléatoires sans motif géométrique.

Pour conclure, bien que l'invariance par translation ainsi que la non oscillation des coefficients permettent une concentration de l'énergie sur quelques coefficients et donc une meilleure sélection pour certaines textures, l'apport de la QWT lorsque l'on utilise simplement la mesure de norme n'est pas effective. Mais ce résultat était attendu puisque l'un des objectifs de la généralisation de la transformée DWT à travers l'opérateur de Hilbert est l'introduction d'une notion de phase.

Utilisation de la phase de la QWT Nous présentons les résultats à partir de la phase sur les différentes bases sur les figures 4.9, 4.10 et 4.11. Nous constatons que les résultats de la mesure seule de la phase arrivent globalement à concurrencer ceux de la mesure de norme. Bien évidemment, la pondération à partir de la valeur du module augmente la qualité de discrimination de la mesure de phase. Pour la suite, seul l'écart-type pondéré sera pris en compte.

A la lecture attentive des résultats, on constate que les deux informations sont complémentaires, à savoir que certaines textures sont mieux discriminées par l'information de phase, en général celles qui contiennent des informations géométriques très marquées, et d'autres à comportement plus dispersé, plus aléatoire sont mieux discriminées suivant le module. C'est donc tout naturellement que les deux mesures sont combinées.

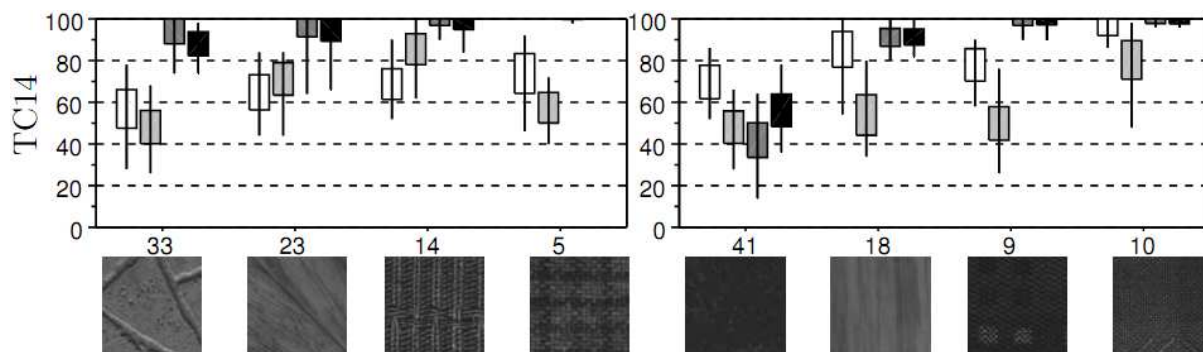


FIGURE 4.11 – Résultats de classification à partir de la mesure de module, de phase et des deux combinées : base TC14 [228].

	Brodatz		TC12		TC14	
	Moy	Std	Moy	Std	Moy	Std
DWT Mod.	64%	9	83%	4	56%	8
QWT Mod.	64%	9	82%	4	51%	8
QWT Phase	65%	8	79%	4	58%	6
QWT Comb.	76%	7	91%	4	63%	6
CWT Mod.	65%	8	87%	4	62%	7

TABLE 4.2 – Mesure de discrimination globale

Utilisation de la phase et de la norme de la QWT Comme attendu, l'utilisation des deux informations combinées permet d'obtenir de très bons résultats, avec une augmentation par rapport à la DWT de 10%. Une analyse détaillée selon les textures est présente dans le papier [228] donné en annexe. Une mesure globale est indiquée dans la table 4.2 qui confirme les résultats de la figure 4.12.

Pour conclure, nous comparons le classifieur QWT avec le classifieur CWT. Pour cela, nous reprenons les paramètres proposés dans les articles les plus récents [46, 47]. Tout d'abord, nous avons constaté comme attendu que la mesure du module de la transformée complexe obtient globalement de meilleurs résultats que la transformée DWT du fait de son invariance.

La seconde expérimentation illustre la difficulté de l'utilisation de l'information de phase. Pour cela, nous reprenons la proposition de mesures faite dans [47]. En revanche, comme nous l'avons précisé, nous n'appliquons par d'étapes de pré-traitement sur les mesures, à la différence de la solution proposée dans [47] où une ACP est calculée. Dans ce cas, la combinaison du module et de la mesure de phase de la CWT obtient des résultats catastrophiques (autour de 5%) [228]. Ceci s'explique par une phase instable numériquement. Ce phénomène n'apparaît pas dans l'article de Celik et al. car la réduction de l'espace des paramètres élimine probablement ces mesures inconsistantes et donc les mesures associées à la phase. Ce point illustre deux choses : la difficulté à mesurer une phase et l'ambiguïté des résultats lorsque l'on étudie un système de discrimination complet.

Nous n'utiliserons que le vecteur de mesures proposé dans [46] qui repose sur la norme. La figure 4.13 illustre les résultats de discrimination obtenus. Comme pour les autres figures, seules les classes les plus marquantes sont présentées. Nous constatons de nouveau que l'information de phase introduite dans la

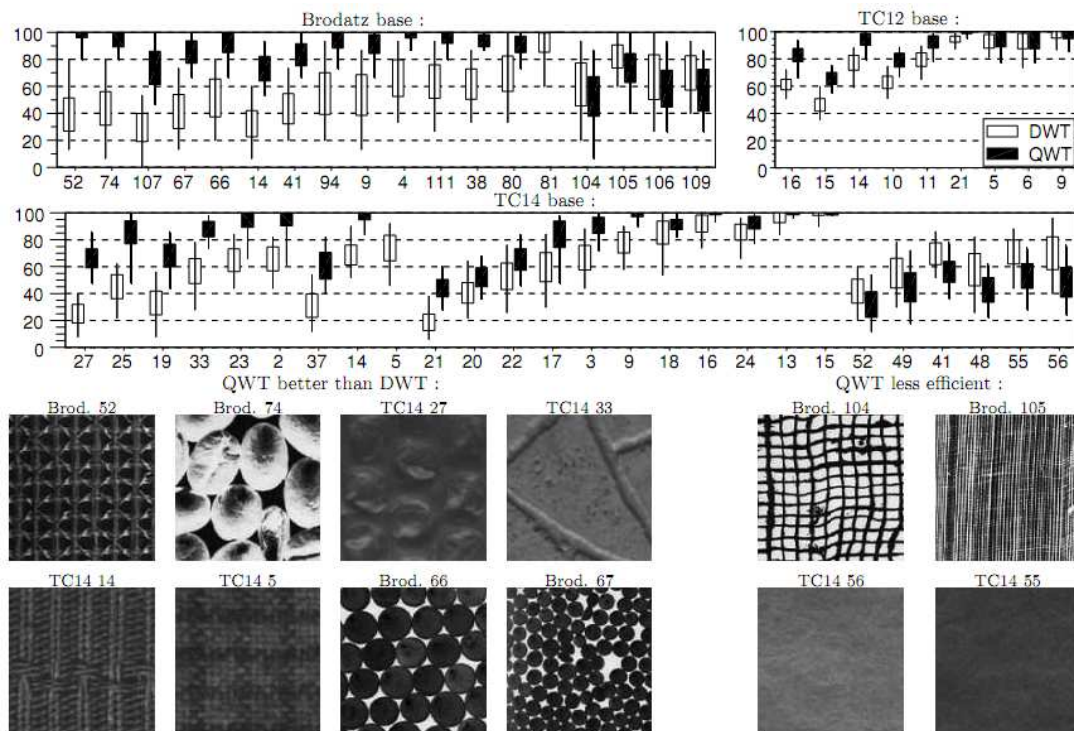


FIGURE 4.12 – Résultats finaux de classification à partir de la mesure combiné de module et de phase [228].

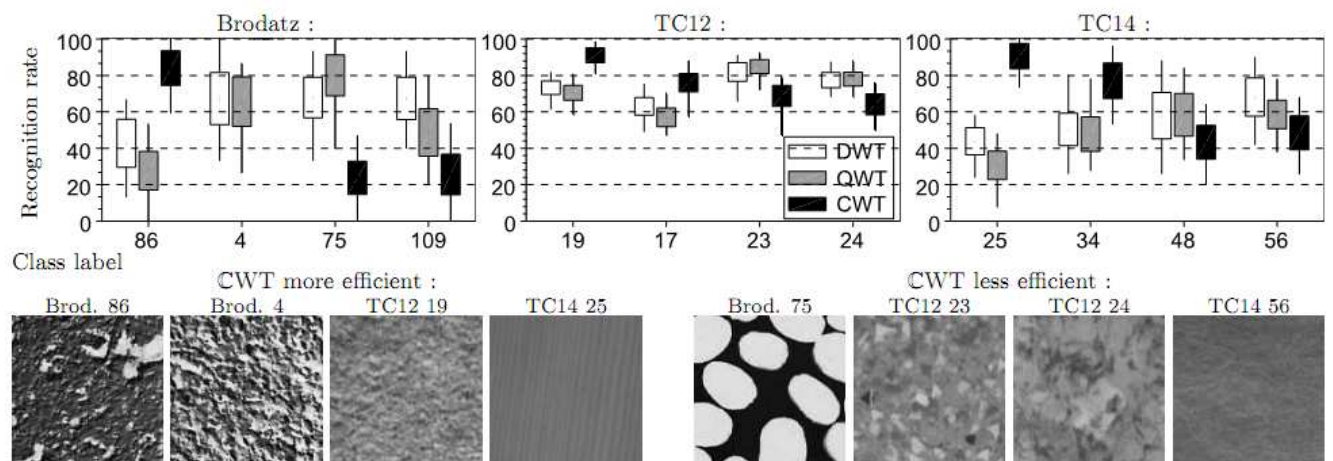


FIGURE 4.13 – Comparaison avec la classification obtenue avec les ondelettes complexes [228].

mesure liée à la transformée QWT permet sensiblement d'améliorer les résultats.

Cette première étude met en lumière la qualité de cette nouvelle représentation. C'est pourquoi un projet est en cours ayant pour but de l'inclure dans une chaîne plus complexe dédiée à la compression d'image.

4.5 Codage d'image par QWT

Nous proposons maintenant de présenter brièvement la seconde application de la transformée QWT, à savoir le développement d'un codeur d'images. Dans le cadre de l'ANR CAIMAN, auquel participe le département XLIM-SIC, nous proposons d'appliquer la transformée QWT pour des problématiques de compression. Ce projet a débuté en décembre 2008 avec comme partenaires Thalès Communications, l'IETR (UMR 6164) et ETIS de Cergy (UMR 8051). Le projet CAIMAN se propose de travailler sur de nouveaux codeurs d'images permettant de retrouver l'usage universel de JPEG, mais avec des fonctionnalités avancées, telles que celles définies au sein du standard JPEG 2000. L'objectif principal de CAIMAN est d'étudier un codeur d'images qui intègre conjointement sécurité, tatouage, robustesse, scalabilité, adaptation au réseau sans sacrifier à l'efficacité de compression et la qualité perçue.

Nous avons mené, notamment à travers le post-doctorat de A.S. Bacquet et la première partie de la thèse de R. Soulard [224], une étude de faisabilité d'un codeur utilisant la QWT. Nous donnons ici quelques éléments.

4.5.1 Principe général du codeur

Nous présentons sur la figure 4.14 le codeur tel qu'il a été imaginé. Bien évidemment nous retrouvons les briques classiques présentes dans un schéma de compression. L'originalité de ce codeur, au delà de l'utilisation de la QWT, est le bloc d'optimisation qui permettra de s'adapter à la fois au canal de transmission et à un critère de qualité. Nous ne détaillerons pas dans le cadre de ce mémoire toute la stratégie envisagée mais l'objectif est tout simplement de ne pas définir le codage source sur seulement des critères d'information mais de prendre en compte l'environnement de la chaîne.

L'optimisation en fonction du canal suppose une organisation du flux binaire pouvant s'adapter à la qualité du canal. La prise en compte de la qualité suivant un critère perceptuel va se faire à travers une stratégie de quantification et de codage adaptatif spatialement en fonction de points d'intérêt ou de focalisation.

En ce qui concerne plus spécifiquement le codage source, le principe est d'utiliser les propriétés de la QWT qui ont déjà été mises en lumière dans le cadre de la problématique de classification. Tout d'abord l'invariance par translation et surtout l'absence de phénomène d'oscillation autour des discontinuités nous laissent espérer une concentration de l'énergie sur un nombre de coefficients plus faible que pour la DWT. Ensuite, l'information de structure présente dans les phases doit se résumer en quelques valeurs.

4.5.2 Quantification

La quantification des sorties de la QWT est conçue afin de respecter l'importance inégale des différentes zones et éléments de la transformée de l'image. Cela signifie que les quantifications des différentes bandes et des différentes informations au sein des bandes peuvent se faire indépendamment.

Quantification de la trame Tout d'abord, comme il est habituel, la composante basse-fréquence sera traitée à part pour deux raisons principales :

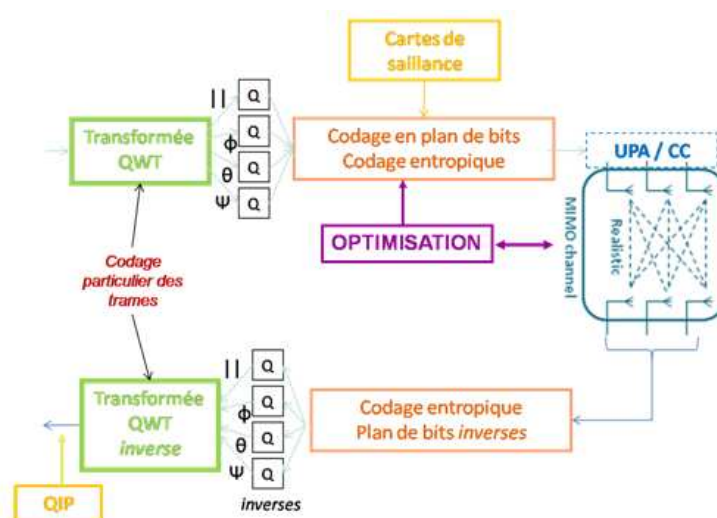


FIGURE 4.14 – Chaîne de compression basée QWT

- elle contient la structure de l'image et donc doit être codée avec attention car une forte dégradation de cette composante entraîne une image compressée très dégradée ;
- elle ne peut pas être considérée comme un signal analytique et donc analysée à travers la notion de norme et de phase du fait de son information "passe-bas".

La trame va être traitée comme dans le cadre d'un compresseur basé ondelette avec une difficulté supplémentaire, à savoir la redondance d'information. Il faut prévoir une stratégie permettant de réduire cette redondance comme par exemple un codage différentiel.

Quantification des coefficients d'ondelettes Comme dans le cadre de la classification, il n'est pas nécessaire de transmettre les phases lorsque le module est nul ou très faible. Ce point là est fondamental car il permet de creuser la représentation de l'image. Les premières expérimentations menées par R. Souillard [224] ont porté sur la quantification module. Pour cela le module de la QWT et celui d'une transformée en ondelette classique ont été quantifiés uniformément, tout en conservant la phase à l'identique (resp. le signe).

On remarque que la reconstruction par QWT est bien meilleure, ce qui confirme la forte présence d'information dans la phase QWT. En particulier, la phase QWT a permis de conserver la structure des bords et lignes de l'image, et le phénomène de pseudo-gibbs est absent de la reconstruction QWT.

Ensuite comme nous l'avons indiqué en préambule, nous avons constaté que nous pouvons quantifier très grossièrement la phase en quelques valeurs tout en conservant presque toute l'information de structure. Cependant, cette quantification doit être dépendante de la sous-bande observée. En effet, puisque les premiers angles φ et θ codent des déplacements dans les directions horizontales et verticales respectivement, ils ne vont contenir une information significative que dans les plans associés à la même direction d'analyse. Un raisonnement similaire peut se faire sur le troisième angle ψ .

Enfin, dans le cadre du travail de thèse de R. Soulard, on suggère que pour la première échelle, qui représente 3/4 des données, une description précise du décalage (φ, θ) est inutile, puisque la résolution de la sous-bande est juste deux fois plus faible que celle de l'image. L'impact d'un mauvais codage du décalage

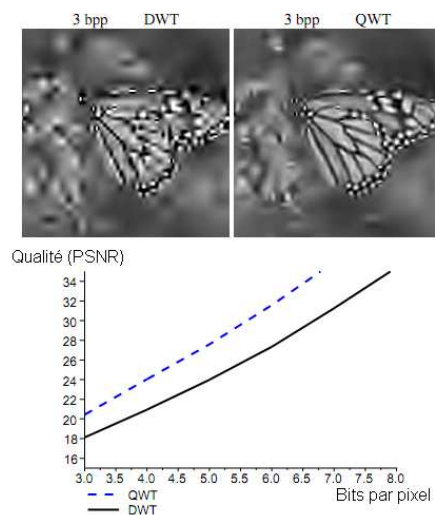


FIGURE 4.15 – Quantification de l'information de module

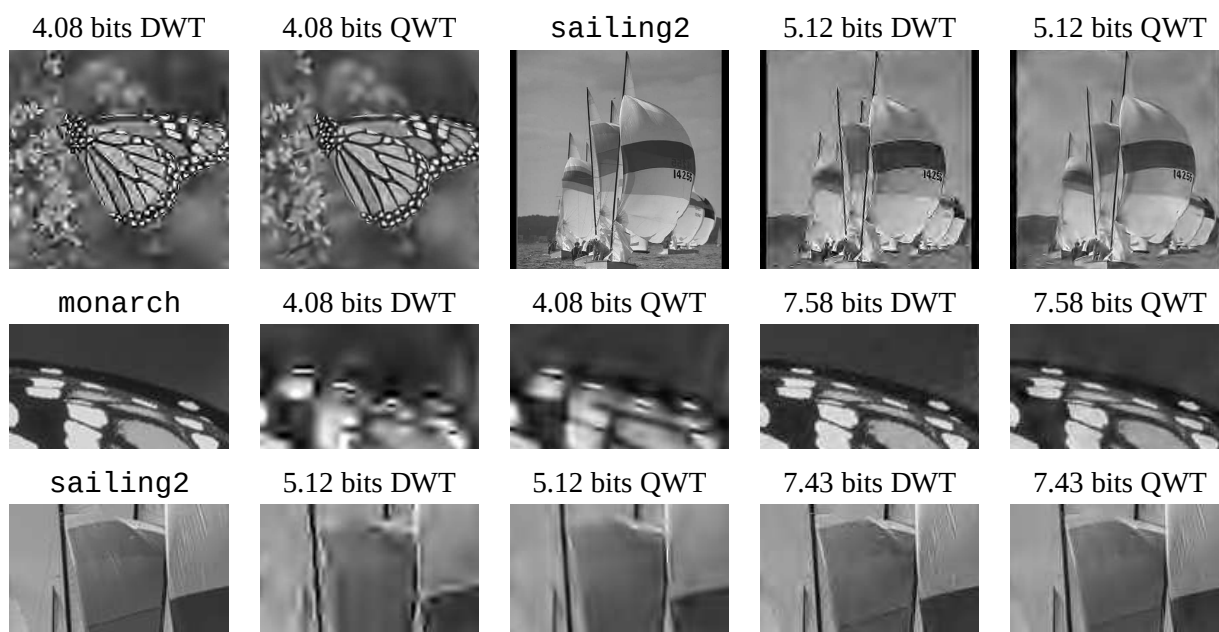


FIGURE 4.16 – Etude de faisabilité du Codage par QWT

est très faible dans ce cas, donc nous quantifions ces phases très grossièrement.

Plus généralement, on peut intuitivement quantifier la phase avec un pas de plus en plus fin, lorsque l'on passe à une échelle de plus en plus grossière.

Nous présentons maintenant les performances avec l'algorithme de codage QWT[224]. Il quantifie uniformément le module avec le nombre de bits en paramètre, et une quantification adaptative de la phase est réalisée selon la description ci-dessus. Pour comparer avec les ondelettes classiques, on force les codages DWT et QWT à produire le même débit pour une même image.

4.5.3 Organisation du flux en fonction du canal

L'idée est d'adapter les schémas de codage utilisés avec les ondelettes classiques dans le cadre de la QWT, comme un codage par plan de bits, la notion d'arbre de 0 etc. Toutes ces notions permettent de structurer le flux et ainsi d'offrir une progressivité au niveau de la qualité, tout à fait adaptée à la transmission MIMO et son découpage en canaux de différentes qualités. Le schéma de codage par plan de bit est assez classique dans son concept mais il est novateur dans le cas de signaux issus d'une transformée telle que la QWT avec 4 éléments (module et 3 phases). La difficulté est de considérer la hiérarchisation présente dans la transformée elle-même. En effet par rapport à la transformée en ondelette classique, il existe un niveau supplémentaire de hiérarchie pour une position donnée à une échelle donnée, à savoir comment se structurent les quatre informations.

A ce stade de l'étude, nous avons abordé le problème d'une manière empirique en estimant la courbe débit/distorsion suivant les combinaisons de paramètres possibles, c'est-à-dire que l'on fixe un certain nombre de bits pour chaque information, et nous évaluons la distorsion associée. Toutefois étant donné le nombre de combinaisons, cette étape, en cours de développement, suit une stratégie selon ces principes :

- Pour chaque coefficient, les quatre informations sont ordonnées (suivant la finesse de codage) avec le module en premier, puis un ordonnancement des angles dépendant de l'orientation de la bande traitée,
- on conserve bien sûr la hiérarchisation propre aux ondelettes, à savoir affectation des bits en priorité aux échelles les plus grossières, puis au fur et à mesure de l'augmentation du débit on s'intéresse aux échelles les plus finies.

Comme nous l'avons dit, l'objectif est de créer un codeur prenant en compte son environnement, et notamment le canal. Dans le cadre de CAIMAN, le canal que nous considérons est un canal MIMO composé de 4 antennes à l'émission et 4 antennes à la réception. Afin d'optimiser la transmission, celui-ci est transformé en 4 sous-canaux virtuels équivalents. Grâce à cette transformation en sous-canaux virtuels, il est possible d'en déduire le SNR de chacun de ces sous-canaux virtuels et donc de déterminer quel taux d'erreur binaire peut être atteint.

Dans cette première version de la chaîne, on ordonne les sous-canaux par robustesse décroissante et on alloue les données les plus importantes aux premiers sous-canaux. Ceci renforce l'importance de la structuration du flux et de la scalabilité dans le cadre d'une stratégie MIMO. De plus, il faut que cette distribution des informations puisse évoluer au cours du temps suivant les modifications du canal.

Enfin l'optimisation en fonction de la modélisation du système visuel humain va se faire à travers l'intégration de carte de focalisation ou carte de saillance. L'intérêt des cartes de saillance est qu'elles permettent un codage adapté aux zones saillantes de l'image. Plus précisément, on doit allouer le débit de codage en priorité aux zones les plus saillantes avant de l'allouer aux zones moins importantes. Toutefois, la difficulté est de traduire dans le domaine transformée une carte de saillance obtenue en général sur l'image dans le domaine spatial. Il nous semble que pour réussir pleinement cette intégration, il faut définir cette carte de saillance directement à partir des coefficients d'ondelettes. On retrouve là peut être les travaux sur "phase congruency" ? Notons qu'il existe des travaux très récents proposant des notions de points de saillance à partir soit des transformées complexes, soit même de la QWT. Ce point constitue l'une des perspectives à court terme de notre travail sur la QWT. D'autant plus que ce point d'étude ne concerne pas seulement des problématiques de compression.

4.6 Discussion sur la transformée QWT

La QWT permet une vraie généralisation du travail concernant la définition d'une transformée en ondelettes analytique 1-D. Elle apporte une réelle amélioration de la transformée en ondelettes en permettant une analyse très fine de l'information 2-D. A travers cette transformée, nous retrouvons le concept de phase qui a alimenté de nombreuses recherches en traitement du signal 1-D depuis cinquante ans. Cependant, le concept de phase locale 2-D est assez peu étudié en image car on ne dispose réellement de définitions 2-D que depuis quelques années. Son utilisation pratique est très limitée car en plus des problèmes d'interprétation, la phase en général est liée à des difficultés pratiques (discontinuité, instabilité numérique).

Le signal analytique quaternionique 2-D et les ondelettes analytiques comme la QWT permettent une nouvelle approche de l'étude de la phase 2-D, et les deux applications développées sous ma direction dans l'équipe Icones montrent l'utilité et la pertinence de ces nouveaux outils.

Comme, nous l'avons dit en introduction de ce chapitre, ces résultats sont très récents mais confirment l'opportunité de poursuivre cet axe de recherche à travers des perspectives pratiques et théoriques. Plus précisément, il me semble opportun de poursuivre la réflexion sur la modélisation de la phase, notamment à travers l'utilisation de lois de probabilité circulaires comme le fait Vo et al. [247] dans le cadre des ondelettes complexes. Ensuite, le travail sur le schéma de compression d'image va se poursuivre afin, comme nous l'avons vu, de définir une chaîne complète adaptative prenant en compte la totalité de l'environnement (canal, système visuel humain).

Nous proposons maintenant d'étudier le dernier et le plus récent des axes concernant les ondelettes que j'anime, à savoir les ondelettes analytiques basées sur une évolution du signal quaternionique de Bülow : le signal monogénique de Felsberg [87].

4.7 Les perspectives : Ondelettes monogéniques couleur

Dans cette dernière section, je vais présenter nos travaux actuels et à venir menés notamment dans la thèse de R. Soulard [227, 226] qui concerne l'étude et la définition de transformée en ondelettes monogéniques, avec une ouverture vers la couleur.

Tout d'abord, nous proposons de revoir la dernière évolution du signal analytique proposé par Felsberg et al. ([87]), mais pour cela nous devons introduire la transformée de Riesz.

4.7.1 Transformée de Riesz

Il existe une autre manière de généraliser la notion de signal analytique 1-D pour les images, c'est le signal monogénique développé notamment en détail par Felsberg [87] dans le cadre des algèbres de Clifford. La définition du signal monogénique utilise une généralisation de la transformée de Hilbert appelée transformée de Riesz. Si l'on applique la transformée de Riesz sur une image niveaux de gris $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ alors d'un signal à valeurs scalaires, nous obtenons un signal à valeurs vectorielles dans $\mathbb{R}^2$:

$$\mathbf{TR}[s](\mathbf{x}) = \begin{pmatrix} \mathbf{TR}_1[s](\mathbf{x}) \\ \mathbf{TR}_2[s](\mathbf{x}) \end{pmatrix} = \begin{pmatrix} h_x * s(\mathbf{x}) \\ h_y * s(\mathbf{x}) \end{pmatrix} \quad (4.59)$$

avec h_x et h_y les noyaux de Riesz qui s'exprime dans le domaine de Fourier complexe par :

$$H_x(\Xi) = -j \frac{\xi_1}{\sqrt{\xi_1^2 + \xi_2^2}} \text{ et } H_y(\Xi) = -j \frac{\xi_2}{\sqrt{\xi_1^2 + \xi_2^2}} \quad (4.60)$$

Nous pouvons remarquer que si nous nous ramenons au cas 1-D, alors le gain fréquentiel de la transformée de Riesz devient

$$H_x(\xi) = -j \frac{\xi}{|\xi|} = -j \text{sign}(\xi)$$

et nous retrouvons la transformée de Hilbert 1-D.

Dans le domaine spatial, les réponses impulsionnelles des deux filtres s'écrivent :

$$h_x(\mathbf{x}) = \frac{x_1}{2\pi [x_1^2 + x_2^2]^{3/2}} \text{ et } h_y(\mathbf{x}) = \frac{x_2}{2\pi [x_1^2 + x_2^2]^{3/2}} \quad (4.61)$$

La transformée de Riesz est parfois appelée transformée de Hilbert isotrope. En effet, on peut noter que dans les deux noyaux de Riesz interviennent les deux variables x_1 et x_2 comme pour la transformée de Hilbert totale, ce qui n'est pas le cas avec les transformées de Hilbert partielles, ce qui assure une analyse ne privilégiant pas une direction.

La transformée de Riesz possède des propriétés d'invariance par translation et changement d'échelle, à savoir que :

$$\forall \mathbf{x}_0 \in \mathbb{R}^2, \text{TR}[s(\cdot - \mathbf{x}_0)](\mathbf{x}) = \text{TR}[s](\mathbf{x} - \mathbf{x}_0) \quad (4.62)$$

$$\forall \alpha \in \mathbb{R}_+, \text{TR}[s(\cdot/\alpha)](\mathbf{x}) = \text{TR}[s](\mathbf{x}/\alpha) \quad (4.63)$$

A notre connaissance, il n'existe que peu d'utilisation de la transformée de Riesz dans le domaine du traitement d'images. Ce n'est que très récemment que différents travaux sont apparus utilisant cette transformée de Hilbert 2-D isotrope. On peut citer par exemple Larkin et al. [127] qui proposent dans le domaine de l'optique la transformée "spiral phase quadrature" qui permet d'avoir des propriétés de démodulation isotrope. Mais ce qui a constitué le point de départ d'une nouvelle manière d'analyser l'image est le travail de Felsberg [87] qui dans le cadre du formalisme des algèbres de Clifford propose une généralisation du signal analytique à l'aide de la transformée de Riesz. Le second apport majeur est venu ces dernières années, de l'équipe de Unser [242] qui a préconisé une relecture de ces travaux sans le formalisme des algèbres de Clifford et en utilisant une approche que l'on pourrait peut-être qualifier de plus "signal". Notamment, ils ont proposé de nombreux développements autour des ondelettes monogènes adaptées pour les images en niveaux de gris.

Dans le cadre de leurs travaux, Unser et al [242] poursuivent l'idée de Larkin, à savoir d'exprimer la transformée de Riesz comme une transformée qui pour un signal à valeurs réelles scalaires fournit un signal à valeurs complexes scalaires, les deux noyaux de Riesz étant tout simplement réunis dans un complexe :

$$\text{TR}[s](\mathbf{x}) = \text{TR}_1[s](\mathbf{x}) + j\text{TR}_2[s](\mathbf{x}) \quad (4.64)$$

ce qui correspond dans le domaine de Fourier à

$$\mathfrak{F}_c[\text{TR}[s]](\mathbf{x}\mathbf{i}) = \frac{-j\xi_1 + \xi_2}{\sqrt{\xi_1^2 + \xi_2^2}} S(\Xi) \quad (4.65)$$

La réponse impulsionnelle h_{TR} associée à la transformée de Riesz complexe vérifie une notion de "steerability" bien connue dans le cadre des bancs de filtres, à savoir que :

$$h_{\text{TR}}(R_\theta \mathbf{x}) = e^{-j\theta} h_{\text{TR}}(\mathbf{x}) \quad (4.66)$$

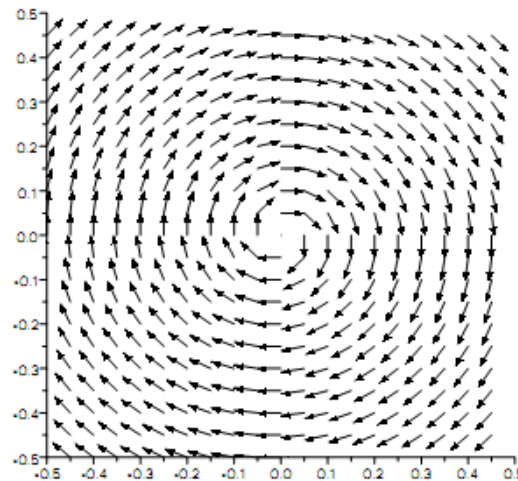


FIGURE 4.17 – Gain fréquentiel associé au noyau de Riesz [223].

avec $R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$ qui correspond à une rotation dans le plan d'un angle θ . Cette propriété est similaire à la propriété de translation pour la transformée de Fourier complexe. Cette propriété est très importante car elle signifie que la transformée de Riesz va permettre de construire des espaces de représentation insensible à la rotation, ce qui est l'un des challenges de ces dernières années en image.

Afin d'illustrer cette notion de généralisation de la transformée de Hilbert à travers Riesz, nous représentons le gain fréquentiel associé au noyau de Riesz. La figure 4.17 illustre la phase de ce gain à travers un vecteur orienté selon le déphasage. Nous constatons que le noyau correspond bien à un déphasage isotrope du signal, avec un déphasage dépendant de l'orientation des composantes.

Nous proposons d'étudier maintenant l'extension du signal analytique à l'aide de cette transformée de Riesz.

4.7.2 Le signal monogène

Tout naturellement, puisque nous avons à disposition une extension de la transformée de Hilbert aux images, nous reprenons le concept du signal analytique en combinant le signal d'origine avec sa transformée de Riesz pour obtenir un signal analytique isotrope, appelé signal monogène.

Soit s une fonction à deux variables x_1 et x_2 et à valeurs réelles, le signal monogène est défini par Felsberg suivant [87] :

$$s_M(\mathbf{x}) = \begin{pmatrix} s(\mathbf{x}) \\ \mathbf{TR}_1[s](\mathbf{x}) \\ \mathbf{TR}_2[s](\mathbf{x}) \end{pmatrix} \quad (4.67)$$

Nous illustrons sur la figure 4.18 la sélection fréquentielle faite suivant les deux définitions du signal analytique, à savoir quaternionique et monogénique. Nous constatons que le spectre du signal monogène n'est pas réduit à un seul quadrant, ce qui renforce l'idée d'une analyse isotrope.

Pour illustrer les qualités de cette définition isotrope et de son influence, nous reprenons le signal synthétique utilisé dans le cadre du signal analytique quaternionique (figure 4.1). Nous indiquons sur la figure 4.19

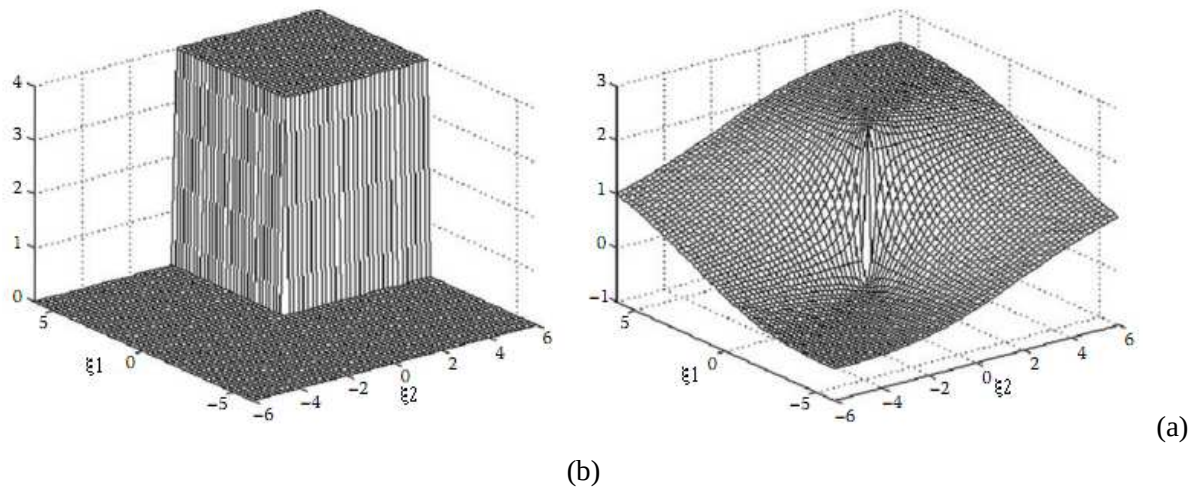


FIGURE 4.18 – Gain fréquentiel des opérateurs définissant : (a) le signal analytique quaternionique (b) le signal monogénique

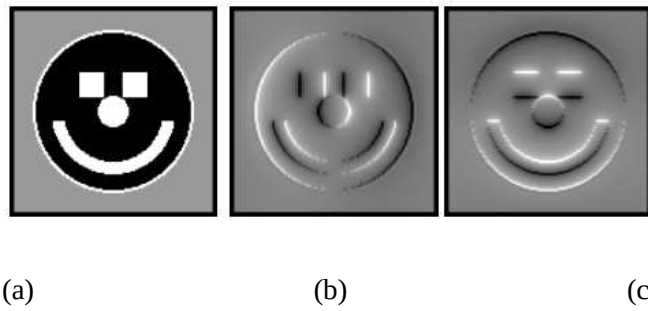


FIGURE 4.19 – Définition du signal monogène : (a) image originale s (b) $\mathbf{TR}_1[s]$ (c) $\mathbf{TR}_2[s]$ [223].

les différentes composantes du signal monogène. Si nous comparons graphiquement les deux versions du signal analytique 2-D, nous constatons que la construction monogène ne présente pas de direction d'analyse privilégiée.

Sur la figure 4.20, sont indiqués les deux modules des deux signaux analytiques 2-D. Du fait des directions privilégiées d'analyse présentes dans les transformées de Hilbert partielles (associées à la définition quaternionique), nous observons par exemple des défauts aux niveaux des coins des objets. Ces défauts disparaissent avec l'approche isotrope associée au signal monogène.

Comme dans le cadre du signal analytique quaternionique nous pouvons définir une notion de phase. Pour cela, nous changeons de système de coordonnées et utilisons les coordonnées sphériques, comme nous l'illustrons sur la figure 4.21. Ces coordonnées sphériques permettent de définir une norme [87] :

$$|s_M(\mathbf{x})| = \sqrt{s(\mathbf{x})^2 + (\mathbf{TR}_1[s](\mathbf{x}))^2 + (\mathbf{TR}_2[s](\mathbf{x}))^2}. \quad (4.68)$$

et deux angles[87]. Le premier angle θ appelé orientation est défini sur la partie Riesz :

$$\theta(\mathbf{x}) = \arctan \left(\frac{\mathbf{TR}_2[s](\mathbf{x})}{\mathbf{TR}_1[s](\mathbf{x})} \right). \quad (4.69)$$



FIGURE 4.20 – Module du signal analytique quaternionique et du signal monogène

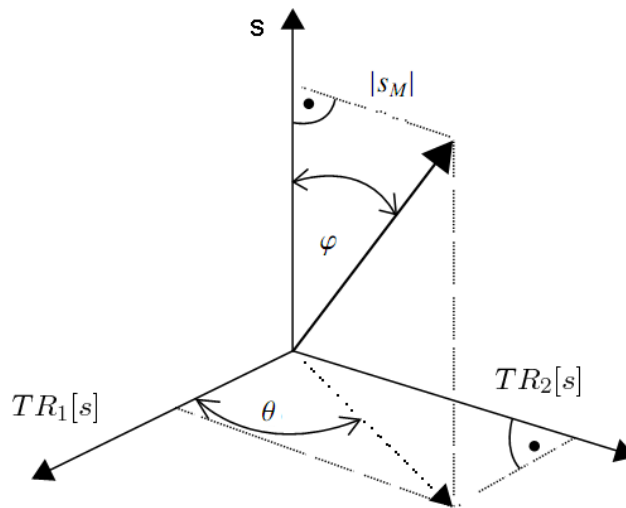


FIGURE 4.21 – Coordonnées sphériques du signal monogène

Le second angle φ appelé la phase est défini par

$$\varphi(\mathbf{x}) = \arctan \left(\frac{\sqrt{\mathbf{TR}_1[s]^2(\mathbf{x}) + \mathbf{TR}_2[s]^2(\mathbf{x})}}{s(\mathbf{x})} \right) \quad (4.70)$$

Le vocabulaire proposé par Felsberg est bien sûr en lien avec l'information portée par ces trois paramètres. Pour illustrer le propos nous reprenons l'image de test (figure 4.22). Comme à chaque fois, le module va être une mesure d'énergie locale permettant de mettre en évidence la présence d'information. Ensuite l'angle θ est, comme nous le voyons sur la figure 4.21, défini suivant les deux noyaux de Riesz. Nous allons mettre en évidence une notion de direction locale dominante de la surface. Enfin, l'angle φ est défini entre le signal d'origine et l'information combinée des deux parties Riesz. Cette lecture est équivalente à l'angle existant entre la partie réelle et la partie imaginaire du signal analytique 1-D. C'est pourquoi il est nommé phase et va porter une information géométrique similaire à celle décrite dans le cas du signal analytique. Nous constatons que nous avons une modélisation de l'information 2-D qui est beaucoup plus claire que la traduction donnée par le signal analytique quaternionique.

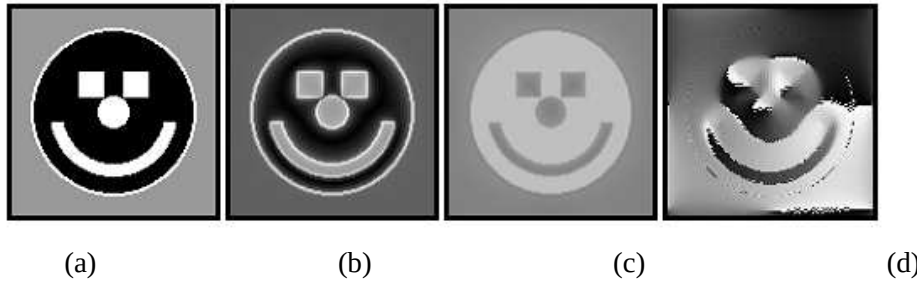


FIGURE 4.22 – Description du signal monogène : (a) image originale (b) Norme (c) Phase (d) Direction

Cependant, il existe un bémol, à savoir que la formulation discrète et numérique n'est pas immédiate. Les travaux de la littérature utilisent un échantillonnage de la version continue du noyau lors de son application sur une image. Nous savons bien que ce type d'approche va forcément limiter les applications puisqu'il n'y a pas de proposition d'algorithme rapide assurant une reconstruction parfaite. Cette voie numérique n'est apparue que très récemment à travers les travaux menés sur les ondelettes monogéniques. Nous nous plaçons clairement dans cette orientation.

4.7.3 Ondelettes Monogéniques

Il est tout naturel d'introduire à partir de la construction précédente une transformée en ondelettes qui permet de définir un signal monogénique associé à différentes bandes de fréquence. A notre connaissance, dans la communauté Signal-Image, il existe deux propositions en ce qui concerne la construction de cette nouvelle transformée en ondelettes monogènes : Olhede et al. [168] qui proposent la construction d'ondelettes Hyperanalytiques et qui se placent plutôt sur un plan continu et l'équipe de Unser [242] qui, à travers un certain nombre d'articles a mis en place une transformée numérique en ondelettes monogéniques. Très récemment de nouveaux articles ont proposé des décompositions en ondelettes monogéniques : Hitzer suivant une approche continue [110] et Held et al. [107] dans un cadre numérique proche de celui de Unser.

Dans tous les cas (sauf le travail de Hitzer), le principe consiste à reprendre l'écriture du signal monogène associée à une fonction s à deux variables x_1 et n et à valeurs réelles :

$$s_M(\mathbf{x}) = \begin{pmatrix} s(\mathbf{x}) & \mathbf{TR}_1[s](\mathbf{x}) & \mathbf{TR}_2[s](\mathbf{x}) \end{pmatrix}.$$

Comme pour les ondelettes quaternioniques, à partir de cette écriture, on peut tout simplement étendre la décomposition en ondelettes en se fixant simplement une fonction d'ondelette isotropique réelle 2-D et en calculant la transformée de Riesz. La transformée en ondelettes monogéniques est définie à partir de la fonction analysante suivante [168] :

$$\psi_A^M(\mathbf{x}) = \begin{pmatrix} \psi(\mathbf{x}) & \mathbf{TR}_1[\psi](\mathbf{x}) & \mathbf{TR}_2[\psi](\mathbf{x}) \end{pmatrix} \quad (4.71)$$

Il faut s'assurer que cette nouvelle fonction ψ_A^M vérifie les différentes conditions liées à une analyse par ondelette, notamment que l'extension monogénique vérifie la condition d'admissibilité. Pour mémoire, on dit qu'une fonction $\psi(x) \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ est une ondelette si elle vérifie la condition d'admissibilité :

$$\int \frac{|\mathfrak{F}_c[\psi](\xi)|^2}{|\xi|} d\xi < \infty \quad (4.72)$$

Dans le cadre des ondelettes Quaternioniques, il existait un lien fort avec les ondelettes complexes et les ondelettes classiques, puisque nous avons une simple combinaison, cela assurait la validité de la construction. Dans le cadre de la construction de l'équation (4.71), on peut montrer que chaque composante issue de la transformée de Riesz d'une fonction d'ondelette vérifie la condition d'admissibilité, donc l'extension monogénique permet bien la définition d'une décomposition en ondelettes. Ceci est notamment discuté dans [168].

A ce stade, nous pouvons alors définir une transformée en ondelettes et calculer pour chaque coefficient obtenu une notion de norme et de phase. Nous avons alors à disposition un outil permettant d'analyser pour différentes bandes de fréquences la présence d'information à travers le module, une direction locale dominante de la surface (l'angle θ) et enfin, l'angle φ va nous donner une information structurale.

Il reste l'épineux problème de l'écriture numérique de cette décomposition. A travers cette formulation monogénique, nous sommes arrivés à un niveau très avancé en termes de description de l'information. En revanche, nous avons fait un retour en arrière de 20 ans en ce qui concerne la formulation numérique. En effet, pour l'instant seul le schéma continu est clairement défini et la transformée de Riesz n'a pas une expression discrète, ainsi nous retournons en quelque sort dans la même situation qu'à l'époque où le lien entre la transformée en ondelettes et les bancs de filtres n'avait pas été fait. Nous pouvons alors trouver les solutions classiques, à savoir une discrétisation directe des fonctions d'analyse, mais nous perdons la rigueur de la définition et en général une partie des qualités.

A notre connaissance, la seule proposition ayant pour vocation un déploiement numérique est la proposition de Unser et al. [242]. Elle repose sur la définition d'une fonction d'échelle Spline, à laquelle on associe une construction d'une fonction d'ondelette anisotrope par un opérateur Laplacien. Nous retrouvons le concept du "chapeau mexicain". L'étape monogénique est amené par une autre fonction d'ondelette complexe liée à la première à travers la transformée de Riesz. Pour chacune des fonctions, Unser et al. proposent les filtres associés, dans le domaine fréquentiel [242].

Il est intéressant de revoir brièvement la définition de ces différentes fonctions car elle apporte un nouveau regard sur la transformée en ondelettes associée à la transformée de Riesz.

L'idée de départ est de construire une fonction d'échelle de lissage isotropique approximant la fonction Gaussienne (pour les propriétés associées à ce type d'espace-échelle). La fonction utilisée dans le cadre des travaux d'Unser est la Spline polyharmonique $\beta_{2\gamma}$ avec γ l'ordre de la Spline. Cette fonction se définit dans le domaine fréquentiel :

$$\mathfrak{F}_c[\beta_\gamma](\omega) = \frac{||\omega||^{-\gamma}}{\sum_{k \in \mathbb{Z}} ||\omega + 2k\pi||^{-\gamma}} \quad (4.73)$$

De cette fonction de lissage, une ondelette primaire isotropique non-séparable est définie suivant l'opérateur Laplacien fractionnaire :

$$\psi(\mathbf{x}) = (-\Delta)^{\gamma/2} \beta_{2\gamma}(D\mathbf{x}) \quad (4.74)$$

avec D la matrice de dilatation.

Nous indiquons sur la figure 4.23 la forme des fonctions analysantes. Nous observons que nous avons bien deux fonctions isotropiques, avec une fonction d'ondelette très proche de l'ondelette historique "chapeau mexicain". A cette analyse multirésolution est associée un banc de filtres restructurable avec un facteur de redondance 4 :3.

Nous avons à ce stade une fonction d'ondelette non-séparable 2-D, nous pouvons utiliser de nouveau le principe de construction exposé précédemment, en calculant la transformée de Riesz sur la fonction d'on-

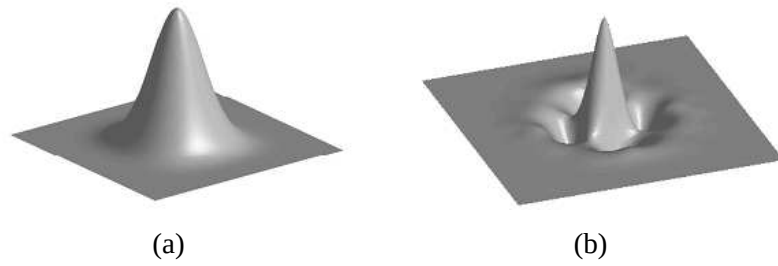


FIGURE 4.23 – Fonction d’analyse associée à une décomposition en ondelette Laplacienne (tirée de [242]) : (a) Spline Polyharmonique de lissage (b) Ondelette définie par un opérateur Laplacien

ondelette primaire. Il apparait alors une nouvelle lecture de la transformée monogénique. En effet, l’opérateur de Riesz est relié à différents opérateurs différentiels, notamment (si l’on reprend la notation complexe) :

$$\mathbf{TR}[s](\mathbf{x}) = -\frac{\partial}{\partial x_1}(-\Delta)^{-1/2}[s](\mathbf{x}) - j\frac{\partial}{\partial x_2}(-\Delta)^{-1/2}[s](\mathbf{x}). \quad (4.75)$$

Etant donné la construction de la fonction d’ondelette primaire à travers l’opérateur Δ , nous constatons que la décomposition selon les deux ondelettes issues de la transformée Riesz est équivalente à :

$$\mathbf{TR}[\psi(\mathbf{x})] = \frac{\partial}{\partial x_1}\Delta\beta_{2\gamma}(D(x)) + j\frac{\partial}{\partial x_2}\Delta\beta_{2\gamma}(D(x)) \quad (4.76)$$

Nous retrouvons la décomposition en ondelettes de Maar proposée en 2008 par Van De Ville et al [243] et qui crée ainsi un lien à la fois avec le concept de Marr et avec les analyses de gradient multiéchelle.

David Marr a proposé une modélisation du fonctionnement du système visuel humain. Ce modèle se structure en trois étapes : construction d’un espace-échelle par lissage Gaussien, calcul du Laplacien et détection des passages par 0 et de leurs orientations pour l’extraction des contours.

Nous constatons dans le schéma exposé ci-dessus qu’avec le premier couple de fonctions analysantes, nous calculons le lissage "quasi" Gaussien ainsi que le Laplacien. Ensuite à travers la transformée de Riesz, et donc le calcul de gradient associé on peut facilement détecter les passages par 0 à travers l’extraction des maxima.

La considération conjointe des deux décompositions forme une analyse monogène de redondance totale de 4 : 1 dont on peut extraire le module et la phase 2-D, suivant les expressions proposées par Felsberg, et nous retrouvons une mesure locale de module, de direction et d’angle. A travers l’utilisation du calcul du gradient, une telle analyse permet aussi une relecture des informations détectées par la transformée monogénique suivant un tenseur de structure et avec un calcul d’énergie d’orientation et de cohérence associé [242]. Nous nous positionnons suivant cet angle de lecture dans l’extension couleur que nous cherchons à construire pour les décompositions en ondelettes mais aussi et surtout en ce qui concerne la compréhension des informations.

Dans le cadre de la thèse de R. Souillard, nous avons proposé récemment une première extension directe de la transformée en ondelettes monogénique à la couleur [227], nous présentons ce point dans la section suivante.

4.7.4 Introduction de la dimension couleur

Cette nouvelle généralisation 2-D du signal analytique reprend la définition en termes d'analyse de la fonction à valeurs complexe, présentée brièvement au début de ce chapitre.

Signal Monogène : signal analytique et algèbre de Clifford Afin d'introduire la construction du signal monogène défini par Felsberg ([87]), nous reprenons la définition du signal analytique à travers l'algèbre de Clifford. Felsberg propose d'associer à s un signal réel, une courbe dans le plan $e_1 \wedge e_2$, donc une courbe définie par : $f : \mathbb{R} \longrightarrow \mathbb{R}_{2,0}$ définie par $f(x) = s(x)e_2$. Comme nous l'avons vu précédemment le signal analytique est l'extension holomorphe de s restreinte à l'axe des réels, donc cela consiste à trouver une fonction f_a qui satisfait $\Delta f_a = 0$ et dont la partie dans la direction e_2 , soit le signal s de départ. La résolution de cette équation mène au signal analytique.

De cette approche, Felsberg propose dans [87] la généralisation aux images niveaux de gris, à savoir une image niveaux de gris $s : \mathbb{R}^2 \longrightarrow \mathbb{R}$, est associée à la "courbe" f de $\mathbb{R}^2$ à valeurs dans la partie vectorielle de l'algèbre de Clifford $\mathbb{R}_{3,0}$ définie par $f(x, y) = s(x, y)e_3$. Pour proposer un signal analytique associé, Felsberg propose une généralisation :

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 & \text{si } z > 0 \\ e_3 \frac{\partial u}{\partial z} = f & \text{si } z = 0 \end{cases} \quad (4.77)$$

où nous retrouvons le laplacien et la partie dans la direction e_3 , doit être l'image de départ. La résolution de cette équation mène au signal monogénique et à la transformée de Riesz.

Suivant cette idée, nos collègues de La Rochelle ont proposé une généralisation du signal monogène aux images couleur [61]. Pour cela, l'image couleur s un signal de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}^3$ va être associé à une "courbe" $f : \mathbb{R}^2 \longrightarrow \mathbb{R}_{5,0}$, une fonction à valeurs dans la partie vectorielle de $\mathbb{R}_{5,0}$ définie par

$$f(x, y) = f_3(x, y)e_3 + f_4(x, y)e_4 + f_5(x, y)e_5 \quad (4.78)$$

De même que précédemment, le problème se ramène à trouver une fonction u solution du système [61]

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \frac{\partial^2 u}{\partial x_4^2} + \frac{\partial^2 u}{\partial x_5^2} = 0 & \text{si } x_3 > 0, x_4 > 0, x_5 > 0 \\ e_3 \frac{\partial u}{\partial x_3} + e_4 \frac{\partial u}{\partial x_4} + e_5 \frac{\partial u}{\partial x_5} = f(x_1, x_2) & \text{si } x_3 = x_4 = x_5 = 0 \end{cases} \quad (4.79)$$

Demarcq et al. ont proposé pour résoudre ce système, de le scinder en trois sous-systèmes suivant les trois canaux. La solution que l'on obtient est une somme de solutions de ces trois sous-systèmes, ce qui revient à appliquer la condition de Felsberg sur chaque canal. Cette généralisation amène la définition suivante pour le signal monogène couleur [61] :

$$s_M(\mathbf{x}) = [s_R(\mathbf{x}); s_G(\mathbf{x}); s_B(\mathbf{x}); \mathbf{TR}_1[s_R + s_G + s_B](\mathbf{x}); \mathbf{TR}_2[s_R + s_G + s_B](\mathbf{x})] \quad (4.80)$$

où $\mathbf{TR}[s_R + s_G + s_B]$ correspond à la transformée de Riesz appliquée à $s_R + s_G + s_B$.

Cette formulation consiste à analyser la structure géométrique sur l'intensité lumineuse de l'image. L'analyse monogène couleur est non marginale puisqu'elle combine l'utilisation des canaux chromatiques avec l'intensité lumineuse.

Maintenant que l'extension couleur du signal monogène de Felsberg est rappelée, nous proposons d'introduire notre premier travail mené dans le cadre de la thèse de R. Soulard concernant la construction des ondelettes monogènes couleur.

Ondelettes couleur Nous avons vu dans la partie précédente que l'extension du signal monogène à la couleur passe par la division du système en trois sous-systèmes liés à l'équation de Laplace, tel qu'il avait été présenté par Felsberg.

Nous pouvons alors définir une transformée en ondelettes dont les sous-bandes sont des signaux monogènes couleur. Pour cela, la fonction d'échelle et d'ondelette isotropique va s'appliquer sur chacun des canaux et l'ondelette complexe issue de la transformée de Riesz est appliquée sur l'intensité lumineuse.

Donc d'un signal à valeurs dans $\mathbf{R}^3$: $s : \mathbf{x} \leftarrow (s_R(\mathbf{x}) \ s_G(\mathbf{x}) \ s_B(\mathbf{x}))$ la transformée en ondelettes monogéniques couleur fournit des coefficients à valeurs dans $\mathbf{R}^5$:

$$(\psi_{rr} * s_R, \psi_{rr} * s_G, \psi_{rr} * s_B, \mathbf{TR}_1[\psi_{rr}] * (s_R(\cdot) + s_G + s_B), \mathbf{TR}_2[\psi_{rr}] * (s_R(\cdot) + s_G + s_B)) \quad (4.81)$$

avec ψ_{rr} une fonction d'ondelette 2-D isotropique à valeurs réelles. On obtient cinq coefficients formant une représentation en ondelettes monogènes couleur.

Etant donné la construction, le schéma numérique est direct et consiste à utiliser la décomposition en ondelette Laplacienne sur chaque canal chromatique, et la pyramide "Maar" associée à la partie "Riesz" sur la somme des trois. Les 5 ondelettes couleur associées sont illustrées sur la figure 4.24 ainsi que la schématisation de la décomposition [227]. La figure 4.25 présente les décompositions en trois échelles de deux images de test. La première image de test permet de simuler un contenu présentant des ruptures couleur alors que la seconde est associée à différents atomes de couleur. Chaque décomposition est composée de 4 sous-bandes issues d'une décomposition sur 3 niveaux. Nous observons une seule sous-bande (anisotrope) par échelle.

En termes d'informations, les 3 composantes c_R , c_G et c_B permettent de mettre en lumière les ruptures suivant les différents canaux couleur, sans orientation spécifique.

Ensuite, afin d'illustrer les composantes "Riesz", nous construisons une image artificielle couleur avec le module codé dans l'intensité de cette image et la direction/phase codée dans la teinte. Les pixels de fortes intensités indiquent les positions des différentes structures de l'image. L'orientation principale de ces structures est traduite par la couleur de ces points.

Nous développons actuellement une modélisation des coefficients, notamment l'information de phase suivant le même principe que les ondelettes complexes avec pour application la restauration de données.

Toutefois, cette construction dans la ligne directe des travaux de Felsberg et de l'extension de Demarcq ne nous satisfait pas complètement car elle ne résout pas le problème de l'interprétation "signalistique" des informations contenues dans la représentation. A ce stade, il nous semble qu'il n'existe pas de lien évident entre cette construction et une lecture physique du signal de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}^3$. En effet, la phase "couleur" est construite entre la partie issue de la transformée de Riesz et le signal réel, ce qui est dans la logique directe de l'extension de Felsberg. Mais dans le cas de l'image couleur, nous obtenons 3 phases associées aux trois canaux, donc difficiles à interpréter. Ceci limite alors les applications possibles. C'est pourquoi, nous menons actuellement à travers la thèse de R. Soulard un travail original proposant une extension couleur des travaux de Unser autour des ondelettes monogéniques avec une intégration des travaux concernant les tenseurs de structure menés dans le cadre de la couleur.

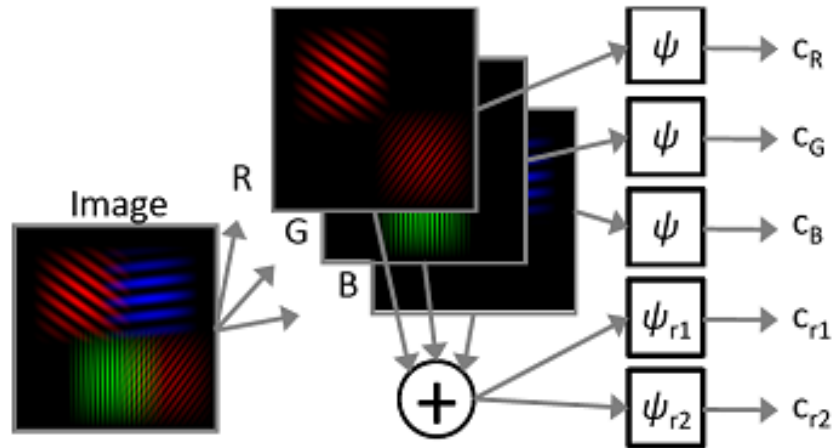


FIGURE 4.24 – Illustration du schéma de calcul de la transformée en ondelette monogénique Couleur

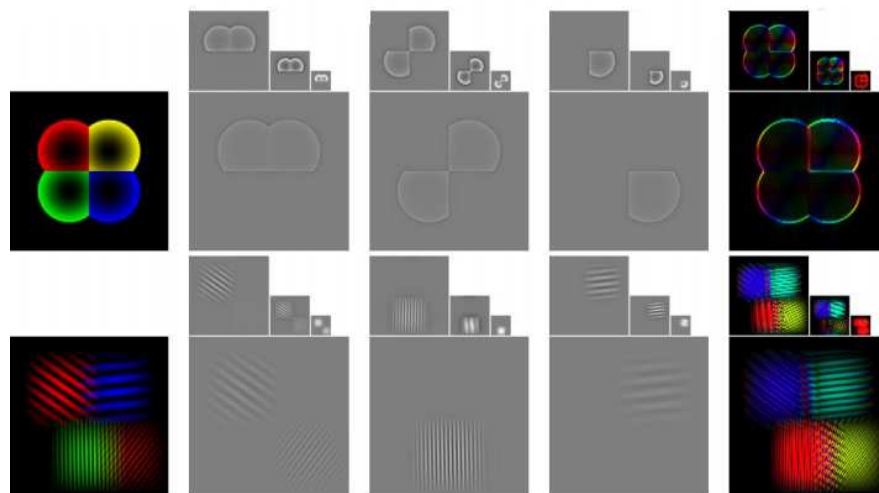


FIGURE 4.25 – Exemple de décomposition en ondelette monogénique Couleur

4.7.5 Monogénique, Radon et Ridgelet

La transformée de Riesz présente des liens très étroits avec la transformée de Radon longuement étudiée au début de ce mémoire. Pour mémoire, soit $R_s(t, \theta)$, le coefficient de la transformée de Radon du signal $s \in L^2(\mathbb{R}^2)$, calculé selon la droite de paramètres (t, θ) , ce coefficient est défini par :

$$\mathbf{Ra}[s](t, \theta) = R_s(t, \theta) = \int_{\mathbb{R}} \int_{\mathbb{R}} s(x_1, x_2) \delta(x_1 \cos \theta + x_2 \sin \theta - t) dx_1 dx_2 \quad (4.82)$$

Différents travaux, notamment la thèse de De Knock [120], ont montré que l'on peut relier la transformée de Radon et la transformée de Riesz :

$$\mathbf{TR}[s](\mathbf{x}) = \mathbf{Ra}^{-1}[\mathbf{TH}[\mathbf{Ra}(\cdot, \theta)[s]]n_\theta] \quad (4.83)$$

avec $n_\theta = [\cos \theta \ \sin \theta]$.

De cette relation, similaire au "Fourier Slice théorème" vu dans le premier chapitre, on observe que la transformée de Riesz peut se calculer par l'intermédiaire d'une transformée de Hilbert 1-D appliquée sur chacune des projections de Radon.

Cette écriture permet tout d'abord d'apporter encore un nouveau regard sur la transformée de Riesz et donc sur le signal monogénique. En effet, Felsberg propose une lecture de la phase du signal monogénique à travers la phase du signal analytique grâce à cette extension par transformée de Radon. Ensuite, nous avons une connexion inattendue avec notre travail présenté dans le premier chapitre concernant la transformée Ridgelet discrète. Cela ouvre deux réflexions que nous menons actuellement :

- puisque nous avons proposé un schéma numérique de transformée de Radon discrète inversible, un objectif consiste à définir un nouvel algorithme de calcul de la transformée de Riesz par ce biais là et par conséquent un schéma original de calcul de la transformée en ondelettes monogéniques. Des premiers travaux menés dans le cadre de la thèse de R. Soulard proposent une construction à partir de notre méthode DART et d'un algorithme Double-Tree.
- il se pose la question du lien entre la transformée Ridgelet et l'analyse par ondelettes monogéniques, notamment peut-on considérer la transformée en ondelettes monogéniques comme une décomposition en Ridgelet complexe ?

De cette relation (4.83) découlent de nombreux points fondamentaux à éclaircir, qui font l'objet de nos travaux actuels.

4.8 Conclusion

Dans ce chapitre, nous avons abordé nos travaux les plus récents concernant la définition de décomposition en ondelettes adaptée à l'image. Cette approche est très séduisante et nous semble prometteuse car elle construit la réflexion à partir de fondements issus du signal et de la physique afin de fournir des outils pour caractériser et comprendre le signal 2-D. Cependant, la communauté ayant alimenté cette réflexion est encore très réduite et issue de domaines différents. Ainsi en plus du travail de recherche sur les outils, nous devons communiquer pour renforcer ce groupe de chercheurs, afin de tenter, peut-être, d'accélérer l'avancée sur certains points. Par exemple, la modélisation et la compréhension de ces informations de phases 2-D ne doivent pas encore rester des impasses. Nos discussions avec différents collègues travaillant sur le traitement d'image et certains sujets de thèses récemment proposés au niveau national indiquent un intérêt croissant pour ce sujet.

A travers nos travaux concernant notamment la transformée QWT, nous avons montré que ces outils qui restaient confinés dans une petite communauté peuvent fournir des réponses applicatives grâce à la qualité de la représentation. Et à travers l'étude d'un codeur QWT qui pose de nombreuses questions, nous proposerons d'autres utilisations comme par exemple l'analyse de points d'intérêts. Enfin notre axe de travail actuel sur les ondelettes monogéniques permet, comme nous l'avons vu, de faire une connexion inattendue avec les travaux présentés au tout début de ce chapitre. Les deux stratégies semblent à l'origine plutôt disjointes, l'une cherche une analyse directionnelle, l'autre reformule une notion fondamentale du signal, et tout logiquement ces deux approches se rejoignent montrant finalement la complémentarité des axes de recherche développés au sein du département XLIM-SIC.

Cet axe très récent portant sur une nouvelle génération d'ondelettes s'est concrétisé à travers le co-encadrement d'une thèse [222]. De nouveau ce sujet se retrouve au niveau national à travers l'action du GDR ISIS ayant pour sujet les outils mathématiques pour l'image couleur, que je co-anime avec D. Tschumperlé. Ces travaux ont été principalement valorisés dans [228, 226, 227, 224, 225].

Le dernier chapitre change quelque peu de sujet, à savoir qu'il va mettre en avant un cadre applicatif que je développe depuis mon arrivée au sein du département XLIM-SIC concernant la sécurité des données.

Chapitre 5

Sécurité

5.1 Introduction

Ce dernier chapitre, qui vient compléter les trois autres chapitres, est associé à une problématique indépendante des travaux menés dans le cadre des analyses multirésolution mais certains de ces développements principaux vont utiliser les décompositions en ondelettes et tenter de profiter de leurs propriétés.

Pour finir ce mémoire, je propose de décrire mes travaux menés dans le cadre de la sécurité des données Multimédia. Cet axe a débuté dès mon recrutement en 2000 d'une manière opportune à travers une thèse CIFRE, réalisée par A. Parisi [171] que j'ai encadrée sous la responsabilité de C. Fernandez, et dont le partenaire industriel était France Telecom R&D. L'objectif de ces travaux était d'étudier les possibilités offertes en termes de Tatouage d'image par l'utilisation des décompositions en ondelettes et la prise en compte de la dimension Couleur. Nous retrouvons donc les thématiques développées dans mon travail mais dans un cadre applicatif précis. Ce premier travail de thèse a permis de mettre en place les bases algorithmiques de nos différentes solutions, présentées dans la première partie de ce chapitre.

Ensuite j'ai pu poursuivre ce travail à travers un important contrat de Recherches toujours avec FT R&D qui avait notamment pour objectif le déploiement logiciel de certaines solutions. Ce contrat a été une opportunité me permettant de proposer différentes évolutions sur la méthode afin de la rendre robuste à certaines attaques géométriques. Tous ces travaux se sont notamment concrétisés par différents brevets ainsi qu'un démonstrateur logiciel, grâce au recrutement d'un ingénieur d'étude ayant en charge cette partie logicielle.

Enfin, à travers le rattachement de notre département à XLIM, se sont ouvertes de nouvelles perspectives. Plus précisément, j'ai co-animé avec P. Gaborit du département DMI de XLIM une action sur la sécurité au sein de la fédération PRIDES. Ce travail s'est concrétisé par la thèse de A. Wadood [1] co-encadrée avec P. Gaborit. Celui-ci mène, entre autres, des recherches théoriques sur la sécurité et notamment les codes correcteurs. Notre objectif, à travers ce travail de thèse, a été d'intégrer différentes familles de codes dans les stratégies de tatouage afin d'augmenter la robustesse de notre schéma, mais aussi d'affiner la procédure de classification de la signature. Ceci fera l'objet de la dernière partie de ce chapitre. Notons que dans ce travail, une réflexion autour de l'invisibilité a aussi été menée en collaboration avec H. Saadane, maître de Conférences HDR, université de Nantes et rattaché au laboratoire depuis 2 ans [4].

Nous détaillerons sous la forme de perspectives des travaux que nous avons débutés cette année à travers l'encadrement d'un stage post-doctoral, toujours en co-encadrement avec P. Gaborit et une collègue de l'équipe Icones AS Capelle. Ces derniers travaux ouvrent un nouvel axe de recherche et concernent une approche particulière en stéga-analyse à travers l'intégration de la dimension psychovisuelle. Cet axe me

semble important car au delà de la dimension scientifique, il est fédérateur car il requiert des compétences dans différents domaines, à savoir la modélisation du système visuel humain, la sécurité et le tatouage et les outils de classification.

5.2 Tatouage vectoriel

Nous proposons, dans un premier temps, de décrire l'algorithme de tatouage vectoriel qui constitue la base de nos différentes solutions de tatouage. Cette présentation comprend une analyse de l'algorithme vectoriel suivie de différentes optimisations comme la détection non décimée, l'analyse du seuil de détection (partie qui nous amènera tout naturellement vers les codes correcteurs) et enfin la définition de procédures complètes ayant pour but la robustesse face aux attaques géométriques. Notons que l'ensemble de ce travail a été associé au développement de codes fournis à FT R&D comprenant :

- une librairie complète de tatouage ;
- une librairie complète concernant le problème des points d'intérêt et du recalage des images ;
- un démonstrateur complet et opérationnel ;
- un module de mesures.

En préambule, nous proposons tout d'abord de revoir très brièvement la notion de tatouage d'images. L'idée n'est pas d'introduire longuement ce domaine très vaste mais plutôt d'introduire les éléments qui nous ont guidés dans le développement de cet axe de recherche.

5.2.1 Introduction au tatouage d'images

Le tatouage d'images peut être perçu comme une branche de la sténographie, qui a pour objectif de cacher un message secondaire dans un message primaire. Le message primaire reste lisible de tous, tandis que le message secondaire n'est lisible que par une ou plusieurs personnes (possédant une information secrète). La sténographie se distingue donc de la cryptographie dans la mesure où l'objectif principal en cryptographie est de rendre illisible le message primaire à toute personne ne possédant pas une information secrète.

Contrairement à un stockage simple d'informations dans l'en-tête du fichier associé à une image, le tatouage est intimement lié aux données. L'information dissimulée a pour but de démontrer l'intégrité du document ou encore d'en protéger les droits d'auteur.

En général, nous avons apprécié les performances d'un tatouage sous les critères principaux suivants :

- La visibilité :

Le tatouage doit être invisible à l'oeil humain, le document marqué doit rester fidèle à l'original. Pour cela, il faut prendre en compte au maximum le système visuel humain. Ce point a fait l'objet de nombreuses réflexions notamment en regard de la dimension couleur.

- La robustesse :

On pourrait séparer cette rubrique en deux parties : la robustesse et la sécurité. Ces deux caractéristiques sont souvent confondues surtout dans le cas du tatouage. On parle de robustesse pour définir la résistance du tatouage face à des transformations de l'image tatouée. Ces transformations peuvent être de type géométrique (rotation, zoom, découpage). Elles peuvent modifier certaines caractéristiques de l'image (histogramme des couleurs, saturation). Il peut aussi s'agir de tous les types de dégradations fréquentielles de l'image (compression avec pertes, filtres passe haut ou passe bas, passage analogique->numérique->analogique, impression de l'image, etc.). Le pirate agit sans réellement savoir ce qu'il

fait. Nous avons travaillé face à ces attaques dès le début de cet axe.

La sécurité caractérise la façon dont le marquage va résister à des attaques "malicieuses". Le pirate va chercher à laver l'image de façon intelligente. Il est sensé connaître l'algorithme et va, en général, chercher la clé qui lit le tatouage. Cela demande souvent une analyse approfondie de la technique de marquage employée. Ce n'est que dans la thèse de A. Wadood et à travers la collaboration avec P. Gaborit que cette dimension sécurité est apparue dans nos travaux [5].

- Le ratio :

Le ratio d'un système de tatouage numérique désigne le rapport : " nombre de données " à dissimuler sur " taille du document hôte ". La question du ratio est apparue lors de l'utilisation des codes correcteurs, lorsque la redondance de l'information cachée augmentait.

- La complexité :

Dans la pratique, la plupart des opérations de tatouage doivent pouvoir s'effectuer en temps réel (surtout la détection, pour des films par exemple). Ceci implique une contrainte supplémentaire sur la complexité des opérations utilisées pour le marquage et pour la détection. Dans notre travail, l'objectif est de définir des méthodes réalisables dans un temps raisonnable, mais sans contrainte de temps de réponse.

- Les informations nécessaires lors de la détection :

Dans ce cadre, on considère le tatouage privé où la donnée originale est nécessaire à l'extraction. Ce système fonctionne en deux étapes : d'abord il compare le support marqué avec le support original, ensuite un algorithme de décision est appliqué pour répondre si, oui ou non, la marque détectée correspond bien à la marque appliquée à l'origine. Il existe ensuite le tatouage semi-privé qui n'utilise pas la donnée originale, mais nécessite la signature lors de l'extraction. Dans ce cas, il s'agit de répondre si telle marque est dans l'image (via un score de corrélation). Nous nous plaçons clairement dans ce cadre. Enfin le tatouage public ne nécessite ni la marque d'origine, ni le support d'origine. Les informations cachées sont extraites par un algorithme particulier, puis sont étudiées dans le cadre de certaines applications, comme le suivi d'intégrité d'un document.

Le tatouage est ouvert à un large champ d'applications. Le travail mené au laboratoire n'a pas été dirigé dans un cadre applicatif précis, cependant il n'est pas raisonnable de déployer une solution de tatouage sans prendre en compte les contraintes de certaines applications. C'est pourquoi nous avons recensé les plus courantes :

- Droits d'auteurs :

L'application la plus évidente du tatouage est le droit d'auteur. Le but est d'insérer une signature, permettant d'identifier le propriétaire, de façon très robuste. Clairement la solution que nous avons développée trouve une implication directe dans ce cadre. D'ailleurs, c'était l'objectif initial de la demande de FT R&D.

- Traçabilité :

Le but est ici de pouvoir contrôler et faire le suivi des copies de document. Cela implique de créer une marque originale pour chaque document distribué. Ces marques doivent être très robustes, pour pouvoir résister à des attaques ayant pour but de détruire la marque. Nous pouvons utiliser la méthode basée ondelettes que nous avons proposée dans un schéma de Fingerprinting, c'est pourquoi des solutions ont été développées dans le cadre de la thèse de A. Waddod. Cependant ce n'est pas actuellement l'objectif principal, et je ne détaillerai pas ce point.

Les deux dernières applications citées ci-dessous n'ont pas été abordées dans notre travail.

- Authentification :

Dans le cadre de cette application, le but est de détecter les modifications effectuées sur une donnée. Ce marquage est qualifié de "fragile". Il doit être résistant à des attaques classiques mais doit être détruit en cas de modification de la donnée. Dans ce cas, la marque peut être intégrée sur les objets principaux de la donnée. Si l'un de ces objets est modifié ou supprimé, la marque est alors détruite.

- Indexation :

Le tatouage doit permettre cette fois de retrouver une image donnée. La marque intégrée à la donnée dépend du contenu de la donnée. L'algorithme utilisé devra permettre d'intégrer une marque de grande capacité. Par exemple, dans le cadre de la vidéo, la marque pourrait contenir la date d'enregistrement, le titre de la séquence, les noms des personnages ou objets principaux, etc.

Ensuite, il faut souligner qu'il existe différents types d'attaques classiques. Nous n'allons pas les détailler puisque nous les introduirons lors des procédures de tests de nos différentes solutions logicielles. Simplement, il est important de bien différencier les deux types d'attaques, à savoir :

- Les traitements bienveillants ou innocents qui sont relatifs à de simples traitements de données.

Dans ce cadre, nous avons les traitements de type compression, filtrage, ajout de bruit, mais aussi les transformées géométriques usuelles et enfin les modifications colorimétriques. Nous nous placerons, comme dans beaucoup de travaux de la littérature, face à ce type d'attaques.

- Les traitements malveillants qui peuvent avoir trois actions : effacement ou désynchronisation de la signature, ou encore utilisation des inconvénients liés à l'algorithme de marquage.

Les actions utilisées peuvent être similaires à la première catégorie, mais le pirate a en général une connaissance de l'algorithme de tatouage et tente de le déjouer. Il est très difficile de contrer ce type d'attaque, concernant la dimension sécurité. Nous n'avons que très peu développé ce difficile aspect du tatouage.

Enfin, nous nous sommes placés dès le début dans le cadre du tatouage des images couleur. Dans ce cadre particulier, les solutions développées dans la littérature peuvent être classées de plusieurs façons en considérant le domaine d'insertion de la marque, les applications visées, les espaces de représentation de la couleur considérés, etc. Dans tous les cas, l'un des objectifs est d'utiliser la modélisation du système visuel humain pour limiter la dégradation visuelle. Sans faire une présentation exhaustive des travaux de la littérature, nous pouvons diviser les techniques spécifiques couleur en trois groupes :

- les techniques de tatouage, basées sur une représentation quantifiée de l'image couleur, qui sont appliquées dans le domaine spatial, en essayant de prendre en compte les propriétés du système visuel humain.
- les techniques dérivées de celles utilisées dans le cadre des images en niveaux de gris, considérant un marquage indépendant d'une ou plusieurs composantes.
- les techniques de tatouage qui prennent entièrement en compte la dimension couleur des images et donc de l'écriture tridimensionnelle de la couleur.

La première technique qui résume les différentes approches, est proposée par Kutter et al. [125] dont l'idée principale est de tatouer la composante bleue, car l'œil est moins sensible aux variations qui correspondent à ces longueurs d'onde. De manière générale, les composantes chromatiques ont un équilibre particulier du point de vue du compromis entre invisibilité et robustesse. Pour une même force de marquage, les composantes basées sur le bleu permettent d'obtenir une meilleure invisibilité de la marque, mais

sa robustesse est alors plus faible. Les composantes basées sur le vert permettent d'obtenir une meilleure robustesse de la marque, mais la dégradation de l'image est alors plus visible. Suivant l'application, l'utilisateur devra alors adapter le choix de ou des composantes marquées en fonction de l'équilibre du compromis invisibilité/robustesse qu'il souhaite obtenir.

Certaines exploitent la couleur de manière vectorielle (par exemple à travers les Quaternions [18, 19], dans d'autres cas les informations de composantes sont intégrées dans d'autres composantes, enfin certaines techniques sont très spécifiques. Par exemple, Battiato et al. [21] proposent de travailler dans un espace de couleur dont les composantes chromatiques représentent des différences de couleur. Le marquage est réalisé par modification des couleurs, par addition d'un même vecteur, pour chaque coordonnée de l'image. L'idée originale de cette méthode est que la marque est générée à partir des couleurs contenues dans l'image originale. W. Puech [189] propose que la matrice de passage entre RGB et l'espace où sera intégrée la marque, soit protégée à l'aide de l'algorithme de cryptage RSA. Ensuite, le tatouage est effectué par modification du deuxième LSB dans ce nouvel espace. Ainsi, une des clés du tatouage est le contrôle de l'espace d'insertion de la marque.

Nous ne souhaitons pas détailler d'avantage la bibliographie concernant le tatouage couleur, et nous renvoyons par exemple à [241] pour un état de l'art plus complet.

Nous présentons maintenant le socle de notre travail en tatouage, à savoir l'algorithme vectoriel basé ondelettes.

5.2.2 Définition générale de l'algorithme vectoriel basé ondelettes

Nous allons décrire l'algorithme issu de nos travaux qui est directement utilisé dans le domaine couleur RGB.

5.2.2.1 Insertion de la marque

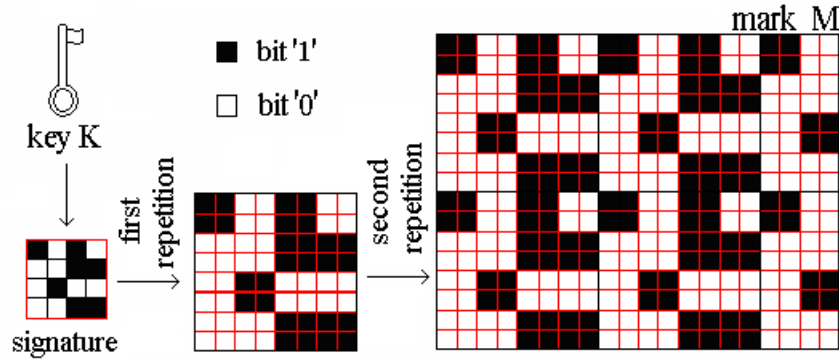
L'insertion de la marque binaire se fait dans le domaine des ondelettes. En effet, la transformée en ondelettes permet d'obtenir une représentation conjointe Espace-Fréquence des détails de l'image. Un tatouage utilisant la transformée en ondelettes bénéficie du caractère localisé dans le domaine temporel des coefficients pour introduire de l'adaptation spatiale. De la même façon, nous pouvons choisir une bande fréquentielle assurant une robustesse correcte avec des dégradations visuelles limitées.

Nous avons proposé de définir alors un ensemble de vecteurs couleur à partir des sous-bandes de chaque composante. Selon les positions des trois vecteurs, l'un d'entre-eux est sélectionné afin d'être marqué. La modification dépendra de la marque à insérer.

Décomposition en ondelettes

La première étape de l'insertion est constituée d'une décomposition en ondelettes décimée de l'image I à tatouer. Cette décomposition est calculée pour chaque composante colorimétrique jusqu'à l'échelle L . Pour cela, nous utilisons l'approche classique par banc de filtres. Le choix de la base d'ondelettes pour une application de tatouage doit satisfaire un compromis entre l'influence spatiale de la modification d'un coefficient (et ceci nous amène à choisir un filtre court) et la qualité de la reconstruction ainsi que la séparation des différentes bandes de fréquences (liée à des filtres longs). Dans cette application, nous proposons d'utiliser une base intermédiaire : la base de Daubechies D8.

Il est bien connu que le système visuel humain a une sensibilité qui dépend des fréquences : la modification de coefficients basse-fréquences provoque en général des dégradations visuelles importantes. D'un

FIGURE 5.1 – Création de la marque M par une répétition adaptative de la signature S

autre côté, un schéma de compression ou un filtrage passe-bas élimine les coefficients de hautes fréquences. Comme pour d'autres algorithmes, nous proposons d'insérer la marque dans une bande fréquentielle intermédiaire qui permet d'être robuste aux attaques de type compression tout en limitant les dégradations visuelles.

Génération de la marque

Afin de définir une signature binaire S , nous utilisons un générateur pseudo-aléatoire contrôlé par une clé K . La marque M est ensuite générée par une répétition adaptative de la signature afin d'augmenter sa robustesse. Comme nous l'illustrons sur la figure 5.1, une première redondance de la signature est introduite par une répétition bit par bit de la signature. Puis une seconde redondance est créée par l'insertion de la précédente marque à de multiples localisations créant ainsi un motif périodique. Cette seconde redondance est utilisée par exemple par Kutter ([123]) et permet d'augmenter la robustesse face à des attaques de type Translation et Cropping. Cette seconde répétition est appliquée jusqu'à couvrir complètement le support à l'échelle L . Par la suite nous ferons évoluer cette étape à travers l'utilisation des codes correcteurs pour faire évoluer la robustesse.

Définition des vecteurs couleur

Tout d'abord, nous calculons une décomposition en ondelettes pour chaque composante R , G et B jusqu'à l'échelle L (échelle associée à une bande de fréquence intermédiaire). De ces coefficients d'ondelette, nous définissons les vecteurs couleur $(\vec{V}_a[n, m])_{0 < (n, m) < \frac{N}{2^L}}$, suivant :

$$\vec{V}_a[n, m] = \{d_{1,L}^a[n, m], d_{2,L}^a[n, m], d_{3,L}^a[n, m]\} \quad (5.1)$$

avec $a = \{R, G, B\}$, $[n, m]$ représente les coordonnées, et $(d_{j,L})_{j=1,2,3}$ les sous-bandes de la décomposition en ondelettes à l'échelle L comme nous l'illustrons sur la figure 5.2. Nous notons P_β l'extrémité du vecteur $\vec{V}_\beta$.

Identification du point à tatouer

Le tatouage consiste à déplacer l'un des trois vecteurs. Le meilleur candidat est défini afin de minimiser la distorsion résultante de la modification des coefficients. Pour chaque coordonnée, nous définissons un vecteur $\vec{V}_M$, qui correspond au vecteur que nous devons marquer et deux vecteurs de références $\vec{V}_{ref1}$ et $\vec{V}_{ref2}$. $\vec{V}_M$, $\vec{V}_{ref1}$ et $\vec{V}_{ref2}$ (figure 5.3) sont sélectionnés selon la règle suivante :

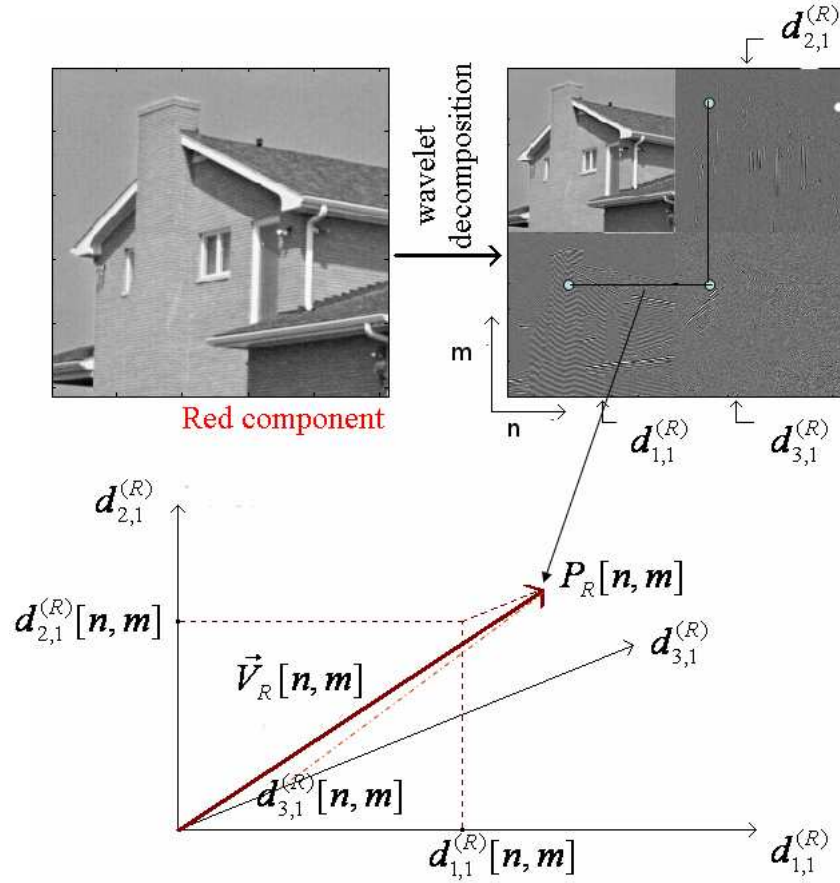


FIGURE 5.2 – Définition du vecteur $V_R[n, m]$ à partir des coefficients d'ondelettes à l'échelle $L = 1$ pour la composante rouge à la position $[n, m]$

$$\|P_{ref1} - P_{ref2}\|^2 = \max_{\substack{(a,b) \in \{R,G,B\} \\ a \neq b}} \|P_a - P_b\|^2 \quad (5.2)$$

$$P_M = P_c$$

avec $c \in \{R, G, B\}$ et $c \neq a$ et $c \neq b$.

Nous constatons que P_{ref1} et P_{ref2} sont les points les plus éloignés parmi les paires de points.

Convention

Nous devons maintenant définir une convention de tatouage comme nous le présentons sur la figure 5.4. Nous notons $\vec{V}_{M,W}$, le vecteur tatoué, qui correspond au vecteur $\vec{V}_M$.

Après tatouage, si $\vec{V}_M$ est associé à $\vec{V}_R$, alors :

- si $M[n, m] = 0$ alors $\vec{V}_{R,W}[n, m]$ doit être plus proche de $\vec{V}_G[n, m]$ que $\vec{V}_B[n, m]$;
- sinon $\vec{V}_{R,W}[n, m]$ doit être plus proche de $\vec{V}_B[n, m]$ que $\vec{V}_G[n, m]$.

Insertion

L'une des possibilités les plus importantes de la stratégie proposée est de contrôler le déplacement de $P_{M,W}$ afin de limiter les dégradations visuelles de l'image. La figure 5.5 illustre les possibles déplacements

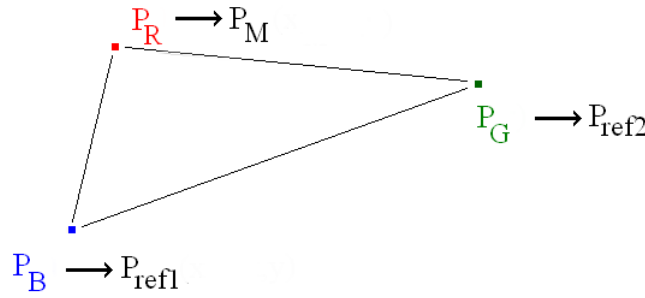


FIGURE 5.3 – Exemple d'affectation

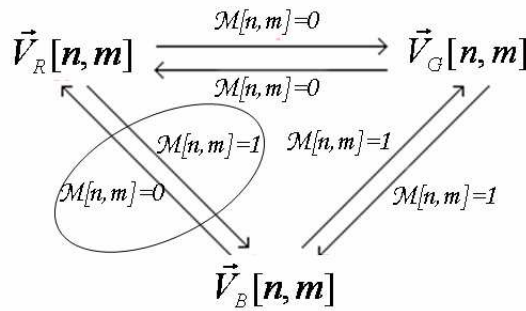


FIGURE 5.4 – Exemple de convention

de $P_{M,W}$. Nous considérons deux cas, notés 'shift 1' et 'shift 2'. A priori la limite de ces deux déplacements est la ligne médiane entre P_{ref1} et P_{ref2} . Toutefois, afin d'être plus robuste, nous avons défini autour de cette ligne une zone de "sécurité" (figure 5.5) telle que, après le tatouage, si $P_{M,W}$ est dans cette zone, il sera déplacé en dehors en augmentant la force de l'insertion. La limite de cette zone peut être par exemple fixée à $\pm 5\%$ de la distance entre P_{ref1} et P_{ref2} .

Nous étudions maintenant les deux cas possibles. Dans le premier cas, P_M est déjà plus proche de P_{ref1} et donc les positions possibles de $P_{M,W}$ après tatouage se trouvent le long du segment $\overrightarrow{P_M P_{ref1}}$ (si P_M est en dehors de la zone de sécurité).

Dans le second cas, P_M est plus proche de P_{ref2} et nous devons alors créer un point intermédiaire P_{int} , défini par :

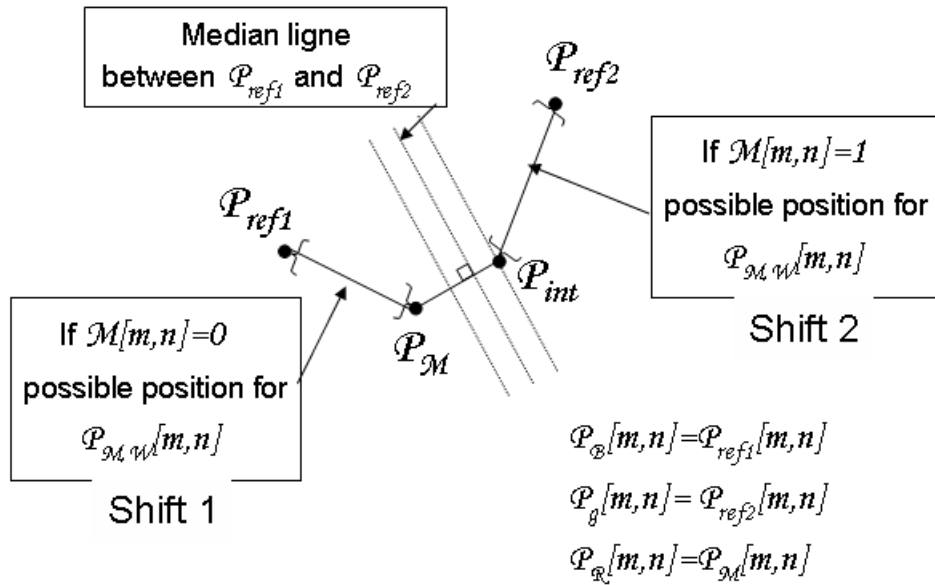
- $\overrightarrow{P_M P_{int}}$ est parallèle à $\overrightarrow{P_{ref1} P_{ref2}}$,
- P_{int} est localisé à la limite de la zone médiane (la distance entre P_{int} et la ligne médiane doit être équivalente à 5% de la distance entre P_{ref1} et P_{ref2}).

Les positions possibles de $P_{M,W}$ sont alors le long du segment $\overrightarrow{P_{int} P_{ref1}}$.

Pour le cas 1, P_M est le point initial du déplacement et pour le cas 2, P_{int} est le point initial du déplacement. Le tatouage est défini par :

$$\vec{V}_{M,W}[n, m] = \vec{V}_{refi}[n, m] - (1 - F_a[n, m]).(\vec{V}_{refi}[n, m] - \vec{V}_S[n, m])$$

avec $i = \{1, 2\}$, $a = \{R, G, B\}$, $0 \leq F_a[n, m] \leq 1$, $S = \{M; int\}$ et F_a représente la matrice des poids

FIGURE 5.5 – Positions possibles de $P_{M,W}$

de tatouage. D'une façon simple F_a peut être constant pour chaque pixel $[n, m]$.

- Si $F_a[n, m] = 0$, alors la force d'insertion est minimale,
- si $F_a[n, m] = 1$, alors la force d'insertion est maximale et $P_{M,W}$ se superpose à P_{refi} .

Reconstruction de l'image

Les opérations exposées ci-dessus sont appliquées à toute l'échelle de la décomposition. La dernière étape de la procédure d'insertion consiste à appliquer une transformée en ondelettes inverse.

La figure 5.6 présente un exemple d'application de l'algorithme de tatouage. Le réglage utilisé est un réglage standard (au vu de toutes nos expérimentations) permettant d'obtenir un compromis tout à fait satisfaisant entre l'invisibilité de la marque et sa robustesse : échelle 4 et forces de marquage $F_r = 0.6$; $F_g = 0.4$; $F_B = 0.9$.

Nous pouvons observer que l'image tatouée n'est que très peu dégradée, les artéfacts ne se voient que si l'on fait un zoom sur l'image, notamment au niveau des frontières entre des zones homogènes. En ces positions les coefficients d'ondelettes sont d'amplitude importante et leurs modifications vont entraîner un phénomène pseudo-gibbs qui peut être visible.

5.2.2.2 Détection

La première étape de l'algorithme de détection consiste à calculer la décomposition en ondelettes de l'image marquée avec la même base d'ondelettes utilisée durant l'insertion. Puis pour chaque coordonnée $[n, m]$, $\vec{V}_M$, $\vec{V}_{ref1}$ et $\vec{V}_{ref2}$ sont déterminés suivant le principe défini par l'équation (5.2). La marque M_D est alors détectée en mesurant la plus grande distance entre $\|\vec{V}_{ref1} - \vec{V}_M\|$ et $\|\vec{V}_{ref2} - \vec{V}_M\|$. Suivant la convention utilisée durant l'insertion (figure 5.4), la marque est alors reconstruite bit par bit.

La signature S_D est alors obtenue par calcul de moyenne (dans la première version la redondance est simplement gérée par moyenne, l'évolution naturelle sera l'introduction des codes correcteurs) et seuillage

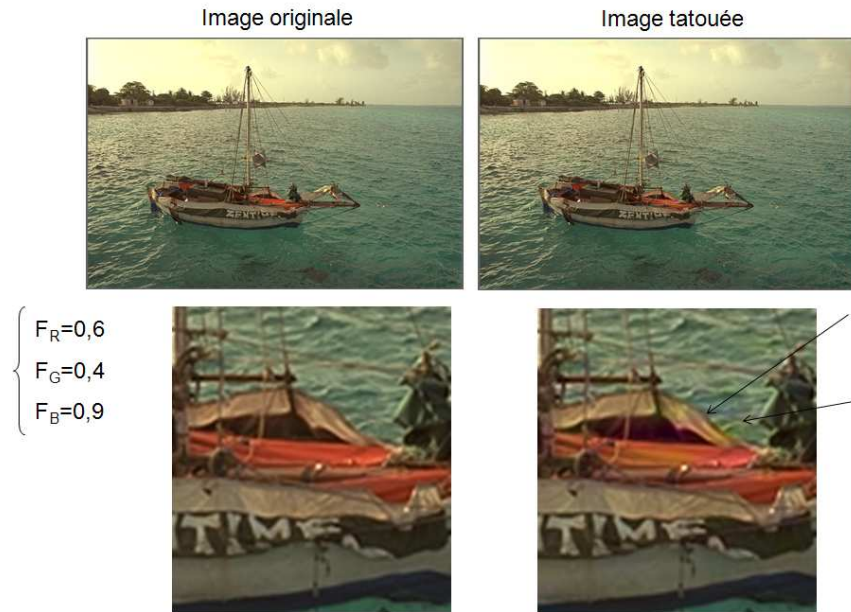


FIGURE 5.6 – Exemple d’application de l’algorithme de tatouage vectoriel : les flèches indiquent quelques défauts

en correspondance avec le processus de répétition utilisé durant la formation de la marque M . Enfin, pour décider si S_D correspond à S , une mesure de similarité est calculée. Cette mesure peut être de deux formes :

- La distance de Hamming notée $d_H(sd, \tilde{sd})$ qui mesure simplement le nombre d’erreurs entre la signature recherchée $\tilde{sd}$ et la signature détectée sd .
- La corrélation qui mesure la ressemblance statistique entre la signature recherchée et la signature détectée.

Bien évidemment, ces deux mesures sont étroitement liées. Nous utiliserons, par la suite, la mesure du nombre d’erreurs (ou de bonnes réponses), plus facile en général à manipuler. Nous présentons donc ci-dessous le lien entre la corrélation et la distance de Hamming permettant ainsi de généraliser les résultats élaborés dans le cadre de la distance de Hamming.

Si $d_H(S, S_D) \leq T_2$ alors nous considérons que S_D correspond à S , sinon la signature originale S n’est pas détectée. Nous reviendrons par la suite sur la notion de seuil.

Cet algorithme tel qu’il vient d’être présenté a été breveté dans le cadre de notre collaboration avec France Telecom R&D [173, 173, 174]. Il offre une grande souplesse d’utilisation, et permet entre autres l’intégration de différents modules ayant pour objectif d’améliorer ses performances. Sa première évolution a concerné le problème de la non-invariance par translation de la transformée qui entraînait une grande fragilité de la méthode face à certaines attaques

5.2.3 Détection non-décimée

La transformée en ondelettes décimée a un désavantage majeur : la représentation n’est pas invariante à la translation dans le temps du fait de la décimation. Cette caractéristique est une limitation importante pour notre détection. Par exemple une simple translation d’un pixel de l’image tatouée, désynchronise la marque et le détecteur est alors incapable d’extraire la signature. De plus les coefficients d’ondelettes support de la

marque peuvent être éliminés durant le processus de décimation.

Contre ce type d'attaque, nous avons proposé dans le cadre du contrat de recherche avec FT R&D d'appliquer une détection non-décimée. La première étape consiste alors en un calcul d'une décomposition en ondelettes non décimée avec la même base utilisée lors de l'insertion. L'algorithme utilisé est l'algorithme à trous [112] dans lequel l'étape de décimation disparaît rendant ainsi la représentation invariante par translation. Comme la décomposition n'est pas décimée, les filtres sont dilatés entre chaque projection. De ce fait, chaque échelle d'ondelette contient $N \times N$ points. Pour chaque échelle l , ces $N \times N$ points correspondent aux $2^l \times 2^l$ décompositions différentes obtenues avec l'algorithme décimé appliqué sur toutes les translations circulaires du signal. Ces décompositions, chacune constituée de $N/2^l \times N/2^l$ points, sont entrelacées.

Durant la détection non-décimée, à l'échelle L , nous extrayons toutes les $2^L \times 2^L$ décompositions différentes potentiellement obtenues par décomposition décimée. Ensuite une marque de taille $N/2^l \times N/2^l$ est formée pour chacune des décompositions. Enfin à partir de ces différents patterns, $N_S \times N_S$ signatures sont reconstruites ($N_S \times N_S$ est la taille de la signature) en utilisant une procédure de moyenne et de seuillage appliquée sur les différentes combinaisons issues de toutes les translations circulaires de la marque.

Nous sélectionnons enfin la signature et la translation qui obtient la meilleure mesure de similarité avec la signature originale. Si cette meilleure mesure de similarité satisfait le seuil fixé alors l'image est considérée comme tatouée.

La détection non décimée est robuste contre les distorsions temporelles mais présente deux désavantages. Le premier concerne le coût de calcul. Toutefois l'augmentation de la complexité algorithmique est limitée car la stratégie proposée ne correspond pas à une recherche exhaustive : le détecteur ne requiert pas l'inversion de toutes les translations possibles. La décomposition en ondelettes est calculée en une seule passe (la complexité augmente de N à $N \log_2 N$) et ensuite nous appliquons $2^L \times 2^L$ extractions de la signature. La recherche exhaustive est seulement présente lors du calcul de la mesure de similarité. Le second défaut concerne l'augmentation de la probabilité de fausse détection du fait du calcul des multiples mesures de similarité. Nous développons ce point dans la section suivante.

5.2.4 Définition du seuil d'acceptation

Il nous faut définir un nombre d'erreurs limite entre la signature d'origine et la signature détectée, c'est ce l'on appelle le seuil d'acceptation. Dans la littérature, nous pouvons trouver différentes études portant sur la définition d'un seuil dans le contexte du tatouage. Ces seuils sont proposés par exemple à partir des approches théoriques de Bayes ou de Neyman (par exemple [158, 108, 188, 159, 22, 16]). Mais pour toutes ces méthodes proposées, nous devons poser différentes hypothèses sur les lois de distributions suivies par les coefficients tatoués. Il nous a semblé difficile de proposer une modélisation à travers des distributions rigides dans le cadre de notre algorithme de tatouage. De ce fait nous proposons de définir le seuil dans le cas le moins favorable à partir de simples hypothèses d'indépendance. Il est évident que le protocole développé à la suite peut sembler naïf, mais il nous a permis d'introduire des notions de bornes dans la configuration de notre détecteur (ce qui constituait notamment une demande de France Telecom R&D).

5.2.4.1 Préambule

L'acceptation d'une signature va se faire dans l'environnement suivant.

Il existe deux hypothèses :

$$\begin{cases} H_0 & \text{la marque n'a pas été insérée ;} \\ H_1 & \text{la marque a été insérée.} \end{cases}$$

Nous avons deux décisions possibles :

$$\begin{cases} D_0 & \text{on considère que la marque n'a pas été insérée ;} \\ D_1 & \text{on considère que la marque a été insérée.} \end{cases}$$

Soit SD le vecteur aléatoire correspondant à la signature extraite. L'ensemble des valeurs que peut prendre SD constitue l'espace d'observation de SD noté Z . Nous notons $sd = \{sd[i]\}_{i=1..N}$ une réalisation du vecteur aléatoire SD .

L'espace d'observation est divisé en deux régions Z_0 et Z_1 tel que si $SD \in Z_0$ la décision est D_0 et si $SD \in Z_1$ la décision est D_1 . L'espace d'observation est l'union de ces deux ensembles : $Z = Z_0 \cup Z_1$. Les frontières de ces deux régions vont être fixées par notre procédure de décision (et plus précisément par notre seuil).

Nous allons devoir manipuler différentes probabilités, notamment :

- $P_{SD/H_0}(sd/H_0)$ la probabilité d'observer la signature sd alors que la marque n'a pas été insérée ;
- $P_{SD/H_1}(sd/H_1)$ la probabilité d'observer la signature sd alors que la marque a été insérée ;
- $P(D_i/H_j)$ la probabilité de prendre la décision i alors que l'hypothèse H_j est vraie ;

Pour définir notre sélection (ou seuil), il faut en général étudier deux probabilités :

- La probabilité de fausse alarme, à savoir $P_f = P(D_1/H_0)$.
- La probabilité de rater la signature, à savoir $P_m = P(D_0/H_1)$.

La probabilité d'erreur associée à notre décision s'exprime en fonction de ces deux probabilités :

$$P(\text{erreur}) = P(D_1/H_0)P(H_0) + P(D_0/H_1)P(H_1) = P_f \cdot P(H_0) + P_m \cdot P(H_1) \quad (5.3)$$

5.2.4.2 Calcul de la probabilité de fausse alarme

Comme nous l'avons dit, la probabilité de fausse alarme se définit par

$$P_f = P(D_1/H_0) = \sum_{sd \in Z_1} P_{SD/H_0}(sd/H_0) \quad (5.4)$$

Nous notons $\tilde{sd}$ la signature parfaite recherchée. Nous allons maintenant étudier P_f suivant la définition du seuil d'acceptation d'une signature en fonction de la distance de Hamming. Nous acceptons une signature en fonction d'un seuil T si $d_H(sd, \tilde{sd}) \leq T$. Nous nous plaçons dans le cadre où les N bits de la signature sont considérés indépendants les uns des autres.

Donc,

$$\begin{aligned} P_f &= P(D_1/H_0) = \sum_{sd \in Z_1} P_{SD/H_0}(sd/H_0) \\ &= P_{SD/H_0}(\tilde{sd}/H_0) + P_{SD/H_0}(d_H(sd, \tilde{sd}) = 1/H_0) + \dots + P_{SD/H_0}(d_H(sd, \tilde{sd}) = T/H_0) \end{aligned} \quad (5.5)$$

A cette étape, une certaine modélisation des mesures doit être faite. Afin de calculer P_f , nous faisons l'hypothèse d'indépendance entre la réalisation $\tilde{sd}$ et l'hypothèse H_0 .

Donc si il y a indépendance entre la réalisation $\tilde{sd}$ et l'hypothèse H_0 :

$$P_f = P_{SD}(\tilde{sd}) + P_{SD}(d_H(sd, \tilde{sd}) = 1) + \dots + P_{SD}(d_H(sd, \tilde{sd}) = T) \quad (5.6)$$

Pour une égalité avec l , nous cherchons la probabilité d'avoir l échec sur N tirages avec remise, donc nous avons une loi binomiale :

$$P_f = P(d_H(sd, \tilde{sd}) = T) = \frac{C_N^l}{2^N} \quad (5.7)$$

Donc si le seuil est $d_H(sd, \tilde{sd}) \leq T$ nous avons :

$$P_f = P(d_H(sd, \tilde{sd}) \leq T) = \sum_{b=0}^T \frac{C_N^b}{2^N} \quad (5.8)$$

Considérer l'indépendance entre la valeur de la signature détectée et l'hypothèse de départ signifie que l'on estime que notre tatouage n'influence pas la valeur de la signature que l'on détecte (si l'on considère par exemple $P_{SD/H_1}(\tilde{sd}/H_1)$). Dans ce cas, nous nous plaçons alors dans le pire cas : pour décider si l'image est marquée, une règle "pile ou face" est appliquée.

Nous avons vu qu'il y a une justification à la définition du seuil suivant le calcul de fausse alarme proposé, et puisque l'on se place dans le pire cas, et donc *a priori* ce seuil est robuste.

En revanche, la critique est que l'on peut avoir tendance à être trop exigeant. En effet, ce qu'il manque dans la définition du seuil, c'est la prise en compte de la probabilité de ne pas détecter des signatures. En effet, pour rappel, la probabilité complète d'erreurs que l'on doit minimiser est :

$$P(\text{erreur}) = P_f \cdot P(H_0) + P_m \cdot P(H_1) \quad (5.9)$$

avec la probabilité $P(D_0/H_1)$ qui mesure, pour un seuil donné, la probabilité de manquer une signature.

La difficulté réside dans le fait que si l'on poursuit avec les mêmes approximations, à savoir indépendance de la valeur de la signature selon l'hypothèse initiale, nous obtenons le calcul suivant :

$$P(D_0/H_1) = \sum_{sd \in Z_0} P_{SD/H_1}(sd/H_1) = \sum_{sd \in Z_0} P_{SD}(sd) \quad (5.10)$$

Donc dans ce cas,

$$P_m = 1 - P_f \quad (5.11)$$

ce qui donne des valeurs très importantes qui n'ont pas de sens.

C'est pourquoi il est essentiel de faire une campagne importante de mesures afin d'affiner un seuil empirique, comme nous pouvons le voir dans certains travaux. Faire ce type de procédure avec un grand nombre d'images permet d'estimer $P(D_0/H_1)$ et $P(D_1/H_0)$ et ainsi d'en déduire un seuil.

Toutefois, si nous conservons l'hypothèse d'indépendance, nous pouvons alors proposer un seuil variable suivant la valeur de P_f indiquée par l'utilisateur.

5.2.4.3 Mise en place du seuil en fonction de P_f

Nous voulons construire une fonction qui nous fournisse une distance de Hamming maximale en fonction d'une probabilité de fausse détection et d'une taille de signature donnée. Comme nous l'avons dit, nous considérons que P_f est égale à :

$$P_f = P(d_H(sd, \tilde{sd}) \leq l) = \sum_{b=0}^l \frac{C_N^b}{2^N} \quad (5.12)$$

Nous recherchons donc le seuil l tel que :

$$P_f = \frac{1}{2^N} \sum_{b=0}^l \frac{N!}{b!(N-b)!} \leq P_f^{max} \quad (5.13)$$

avec N la taille de la signature et P_f^{max} un seuil fixé par l'utilisateur.

Pour fixer le seuil, il suffit d'évaluer les expressions suivantes :

```
Num=N; Den=1; Cumsum=1+N; seuil=2;
Tant que (Cumsum/2N) < Pfmax
    Num=Num*(N-b+1); Den = Den*b; Cumsum=Cumsum+Den/Num;
    seuil=seuil+1;
return(seuil-1)
```

Avec ce développement très "scolaire", nous introduisons une solution pour fixer un seuil en fonction d'une contrainte sur la probabilité de fausse détection. Maintenant, adaptons ce seuil lors d'une détection non-décimée.

5.2.4.4 Évolution avec la détection non-décimée

La détection non-décimée est conçue pour résister aux attaques de types rotation circulaire ou "cropping" qui désynchronisent la marque par translation. Pour cela, une décomposition en ondelettes non-décimée est calculée afin d'obtenir toutes les décompositions décimées possibles, puisque une détection est appliquée sur chaque décomposition décimée. La signature finale choisie est celle qui contient le moins d'erreurs avec la signature recherchée. Bien sûr cette stratégie modifie le calcul de fausse détection. Nous proposons de reprendre le calcul dans le cas de la détection non-décimée.

Considérons directement le cas général où le seuil est $d_H(sd, \tilde{sd}) \leq T$. Nous introduisons l'évènement E qui correspond à l'évènement que la signature étudiée pour une des décompositions décimée soit acceptée (donc qu'il y a au plus T échec sur les N tirages correspondant à la taille de la signature). Si l'on conserve les hypothèses d'indépendance, nous avons vu que la probabilité d'avoir l'évènement E est :

$$P(E) = \sum_{b=0}^T \frac{C_N^b}{2^N} \quad (5.14)$$

Maintenant si nous considérons la détection non-décimée dans sa totalité, si la détection est effectuée à l'échelle L alors le nombre de signatures testées est égal à $M = 2^L \times 2^L \times N_S \times N_S$. Nous faisons de nouveau une hypothèse forte à savoir que les signatures sont indépendantes les unes des autres. Dans ce cas, l'évènement qui correspond à l'acceptation de la signature se traduit par le fait d'avoir au moins une fois sur les M tentatives l'évènement E . De nouveau ce calcul de probabilité est lié à une loi binomiale. Ainsi la probabilité d'avoir k fois l'évènement E est égale à :

$$P(\text{number}(E) = k) = C_M^k P(E)^k (1 - P(E))^{(M-k)} \quad (5.15)$$

La probabilité de fausse détection va donc s'écrire

$$P_f = 1 - P(\text{number}(E) = 0) = 1 - (1 - P(E))^M = 1 - \left(1 - \sum_{b=0}^T \frac{C_N^b}{2^N}\right)^M \quad (5.16)$$

Pour bien quantifier l'augmentation de risque de fausse détection, nous proposons la simulation suivante : nous calculons P_f dans le cas décimé et non décimé pour une signature de 64 bits insérée à une échelle 4. La figure 5.7 illustre les résultats de cette simulation. Pour une même valeur de seuil, nous constatons que la probabilité de fausse détection est supérieure dans le cas non-décimé. Cette estimation "non-décimée" est une borne dans le pire cas car il est faux de dire que les différentes signatures sont non corrélées. La réalité pour la détection non-décimée se trouve entre les deux seuils.

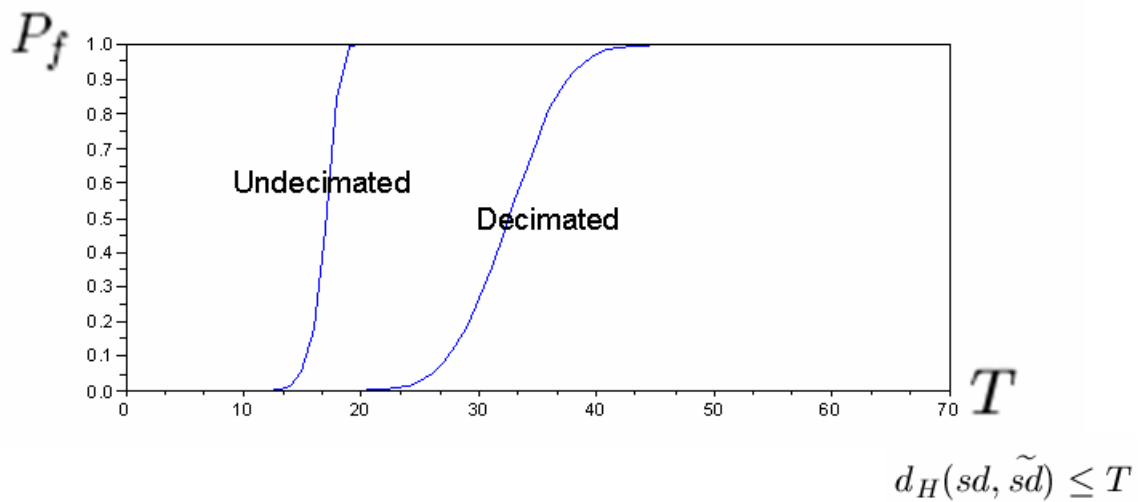


FIGURE 5.7 – Évolution de la probabilité de fausse détection en fonction du seuil

Nous allons maintenant étudier le comportement de cette première version de l'algorithme de tatouage dans différents cas de figure.

5.2.5 Mesures de robustesse de l'algorithme vectoriel basé ondelettes

Je présente ici une campagne de mesures réalisée dans le cadre du contrat de recherche avec France Telecom R&D. Les tests ont été réalisés grâce à l'outil "WIN-SIC-TESTING", développé au sein du laboratoire XLIM-SIC. Ce logiciel de test, dont l'interface est illustrée sur la figure 5.8, permet d'effectuer une série de mesures sur un ensemble d'images.

Nous avons effectué nos tests sur une base de 50 images extraites d'un flux vidéo (voir figure 7.2 en annexe) fourni par France Telecom. Les 50 images illustrent les différents types d'éléments que nous pouvons retrouver sur une chaîne de télévision, que ce soit des logos, du texte, des extraits de films, etc .

Afin d'assurer une indépendance de la mesure par rapport à la marque insérée, pour chaque image, nous utilisons 5 clés de tatouage.

Dans le cadre de la mesure de robustesse face aux attaques nous tatouons une signature de 64 bits avec les réglages permettant a priori d'obtenir une robustesse maximale. Ces paramètres correspondent à une insertion de la marque à l'échelle 4 avec une force de 0.9 pour les trois composantes couleur. Les forces ont été réglées à ce niveau car, étant données les images, le niveau de détérioration était acceptable (selon notamment les critères de FT R&D) et la robustesse était dans cette étude l'objectif prioritaire.

Pour chaque type d'attaque, les résultats sont présentés sous forme d'une courbe qui illustre l'évolution

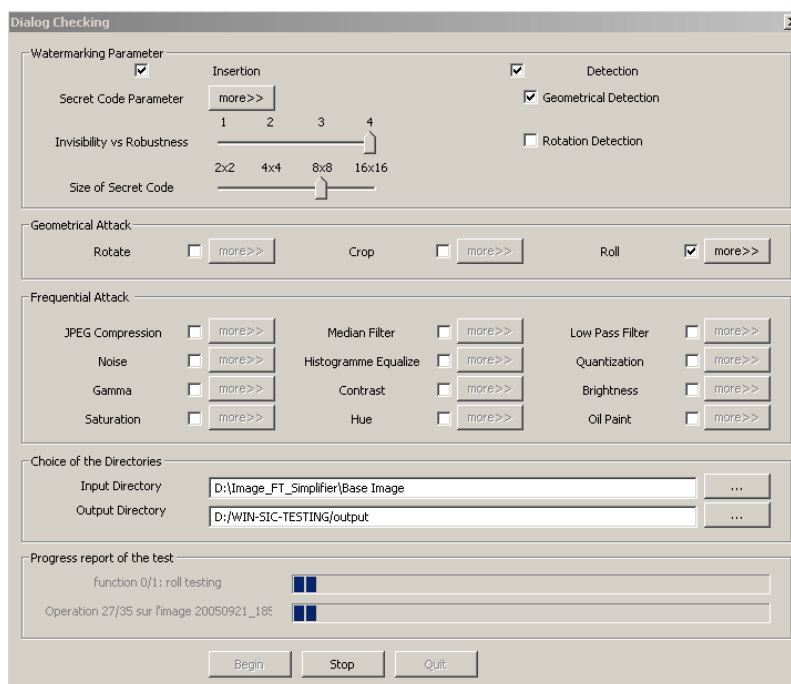


FIGURE 5.8 – Présentation de l’interface du logiciel de test

du nombre de bits retrouvés moyen, maximal et minimal en fonction des paramètres de l’attaque testée. Toutefois, avant de tester la robustesse de l’algorithme, nous proposons d’illustrer empiriquement les problèmes de fausse détection discutés dans le paragraphe précédent.

5.2.5.1 Mesure de fausse détection

D’après les développements menés dans le paragraphe précédent, nous avons fait l’hypothèse que la probabilité de fausse détection se résume à un simple calcul de probabilité d’obtenir une signature donnée constituée de N éléments binaires aléatoires équiprobables indépendants. Il est intéressant de confronter cela avec les mesures, notamment dans le cadre de la détection utilisant l’algorithme non-décimé. En effet, comme nous l’avons vu, cette détection est conçue pour résister aux attaques entraînant une désynchronisation de la marque et donc elle incorpore une recherche de la meilleure configuration.

Nous avons effectué deux campagnes de mesures :

- la première concerne une détection d’une marque associée à une clé impaire sur une image n’ayant pas été tatouée ;
- la seconde concerne une détection d’une marque associée à une clé impaire sur une image ayant été tatouée avec une clé paire.

Dans les deux cas, l’expérience est reproduite avec 10 clés différentes sur l’ensemble de la base de données d’image.

	4x4			8x8			16x16		
	Moyen	Min	Max	Moyen	Min	Max	Moyen	Min	Max
Sans Tatouage									
Nombre erreurs "Décimé"	8	12	3	32	42	22	54	72	42
Nombre erreurs "Non Déc."	12	16	6	46	52	32	75	82	64
Avec Tatouage									
Nombre erreurs "Décimé"	8	12	2	30	41	21			
Nombre erreurs "Non Déc."	14	16	11	46	53	39			

Remarque : les mesures pour des images tatouées avec des marques de $16 * 16$ n'ont pas été indiquées car certaines images de la base sont trop petites pour cette taille de marque.

A partir de ce tableau de mesures, trois constatations s'imposent :

- Le fait que l'image soit tatouée ou non ne change bien sûr en rien la mesure d'erreur,
- Dans le cadre de la détection utilisant la décomposition décimée, l'erreur est d'environ 50% ce qui confirme l'hypothèse utilisée dans le premier paragraphe, à savoir que la probabilité d'erreur d'obtenir un bit juste alors qu'il n'a pas été inséré est tout logiquement d'une chance sur deux.
- Enfin la détection non décimée, du fait de sa recherche de meilleure configuration, diminue le taux d'erreurs lors de la recherche de la meilleure configuration. En fait, pour chaque bit la probabilité d'obtenir la valeur juste reste de 0.5 mais comme ensuite l'algorithme cherche à minimiser l'erreur entre la signature recherchée et la signature détectée, le taux d'erreur diminue environ de moitié (ce que nous avons constaté sur la figure 5.7).

Toutes ces remarques étaient attendues, mais ces mesures permettent de vérifier que l'algorithme de tatouage n'entraîne pas de biais. La dernière remarque nous oblige à fixer le seuil de détection selon l'algorithme de détection que l'on choisit. Si nous utilisons la détection non-décimée, le seuil doit être revu à la baisse.

Pour la suite de notre analyse, pour fixer le seuil de validation sur les erreurs binaires, nous avons proposé de prendre la plus grande valeur du coefficient de taux de réussite observée lors de notre test, à savoir une valeur en pourcentage de 0.8127. Par la suite nous considérons que pour que la détection soit acceptée, il faut que la signature soit retrouvée à plus de 85 %. Ce résultat est en accord avec l'étude menée dans le premier paragraphe. En effet, si nous prenons une signature de 64 bits cela correspond à un seuil de 10 ce qui est bien en accord avec la simulation de la figure 5.7, assurant ainsi une probabilité de fausse détection quasi nulle.

A ce stade cette détermination du seuil d'acceptation reste discutable. C'est pourquoi, dans la suite de cet axe, j'adapterai cette étape de l'algorithme par l'intégration du formalisme des codes correcteurs.

5.2.5.2 Robustesse aux différentes attaques

La campagne de mesures a examiné les effets des attaques suivantes :

- La compression JPEG (le paramètre est la qualité de l'image compressée) ;
- Le bruit (le paramètre est le ratio Signal sur Bruit) ;
- Le filtrage passe-bas (le paramètre est la taille du masque) ;
- Le filtre médian (le paramètre est la taille du masque) ;
- La modification du contraste (le paramètre est l'accroissement ou la réduction de l'intervalle de luminosité : un changement positif dilate l'histogramme de la luminosité, un changement négatif le

contracte) ;

- La modification de la teinte (le paramètre est l'accroissement ou la réduction de la teinte ce qui déplace les couleurs dans le spectre) ;
- La modification de la luminance (le paramètre est l'accroissement ou la réduction de la luminance : -100% produit une image noire ; +100% produit une image blanche car cela fonctionne par translation de l'histogramme) ;
- La modification de la saturation (le paramètre est l'accroissement ou la réduction de la saturation : -100% convertit une image couleur en une image niveaux de gris) ;
- la quantification (le paramètre est le nombre de couleur après quantification).

Toutes ces attaques font parties de la catégorie des attaques non malicieuses.

Notons qu'un pirate voulant par exemple récupérer l'information cachée doit :

- connaître la méthode ;
- récupérer les différents paramètres utilisés (échelle d'insertion, base d'ondelettes ...).

Et ensuite il lui suffit d'appliquer la procédure de détection. Nous n'avons pas intégré, dans cette version, la dimension sécurité avec par exemple un cryptage de la signature.

Nous proposons de détailler chacune de ces attaques car il nous semble très important de définir le comportement de notre algorithme face aux différentes attaques, mais aussi de caractériser ces différentes attaques. En effet, une attaque couleur peut se déployer numériquement de plusieurs manières (selon notamment les logiciels) et les choix faits ont bien sûr une incidence dans la compréhension des résultats.

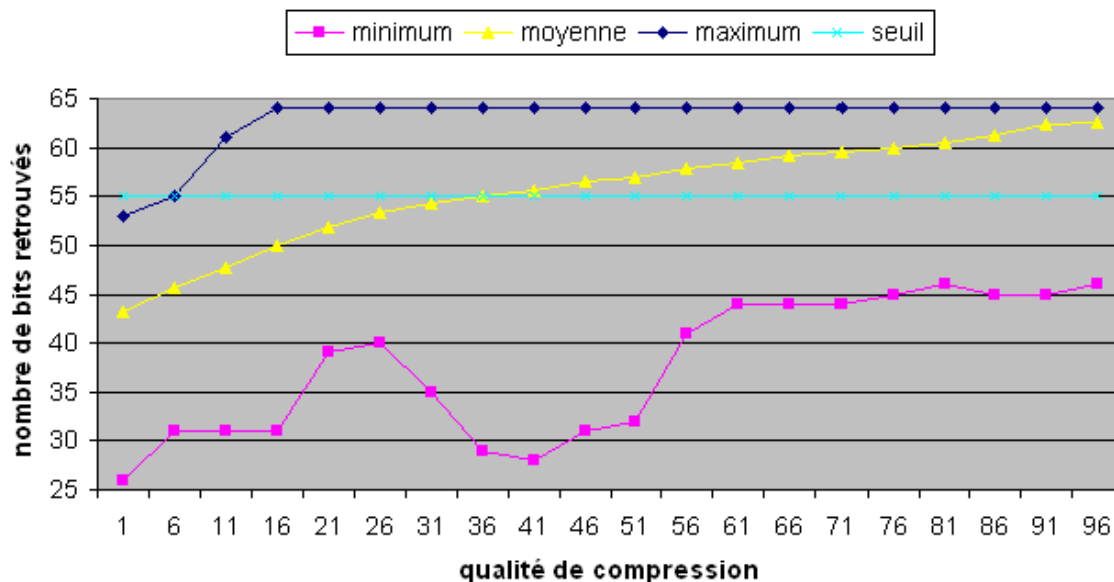


FIGURE 5.9 – Evolution du coefficient de détection correcte en fonction de la qualité de compression

Compression JPEG. Pour tester la résistance face à la compression JPEG, nous partons d'une qualité de compression faible 1% (cf figure 7.5) puis nous augmentons la qualité par pas de 5% jusqu'à obtenir la qualité quasi parfaite de 96% (cf figure 7.5). Les mesures sont présentées sur la figure 5.9. Nous pouvons voir que la compression n'affecte la détection que lorsque l'image est véritablement dégradée. En effet, lorsque la qualité est supérieure à 40%, le taux maximal de bits retrouvés est à 64 et le coefficient moyen est

proche de cette valeur maximale. Ceci signifie que la valeur faible du taux minimal n'est due qu'à un petit nombre d'images sur lesquelles la détection est en difficulté. En général, la mauvaise détection est le résultat de la suppression des moyennes fréquences qui sont les sites d'insertion de la marque, suppression résultant du lissage ou de la dégradation de grandes zones uniformes. Cependant, nous considérons que notre marque est robuste à la compression JPEG pour des taux raisonnables.

Ajout de bruit blanc Gaussien. Nous reprenons l'attaque classique simulant l'ajout d'un bruit blanc Gaussien additif. Le paramètre, dans ce cas là, va être le niveau de bruit. Dans le domaine des ondelettes, le bruit blanc gaussien va le rester et va donc venir s'ajouter à chaque coefficient d'ondelettes.

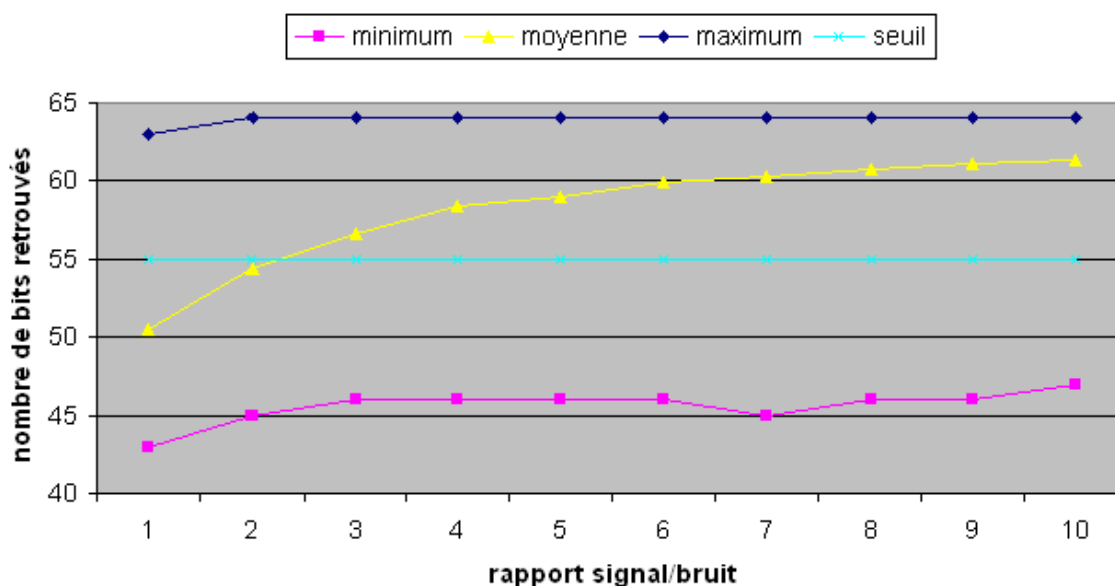


FIGURE 5.10 – Evolution du coefficient de détection correcte en fonction du niveau de bruit

En moyenne, l'application d'un bruit sur l'image ne modifie pas la marque de manière significative. Nous pouvons voir sur le graphique 5.10 que même dans le cas d'un bruit important (ratio signal/bruit égal à 1 dB) la marque est reconnue à plus de 80%. Nous pouvons donc dire que notre méthode de tatouage est résistante à l'ajout de bruit. En terme d'influence, le bruit est présent en majorité dans les hautes fréquences et il ne modifie que légèrement les zones d'insertion de la marque, c'est pourquoi la détection ne présente pas de difficulté.

Modification du gamma. La modification du paramètre gamma de l'image entraîne une mauvaise détection de la marque pour des valeurs de correction gamma supérieures à 3 ou pour de très faibles valeurs, comme 0.1. Le changement important de la valeur du gamma entraîne une modification importante des couleurs des pixels et notamment un phénomène de dépassement soit vers le blanc soit vers le noir. Dans ce cas, par exemple les petites ruptures et oscillations contenant la marque sont "lissées", ces zones ont pour valeurs 0 ou 255 sur les trois canaux, ce qui explique la dégradation de la marque. Nous retrouvons ce problème de dépassement dans différentes attaques couleur.

Modification de la saturation. Comme nous l'avons vu dans le chapitre sur les algèbres géométriques, la saturation mesure l'aspect coloré du pixel (plus ou moins fade ou vive) et correspond à la distance entre la

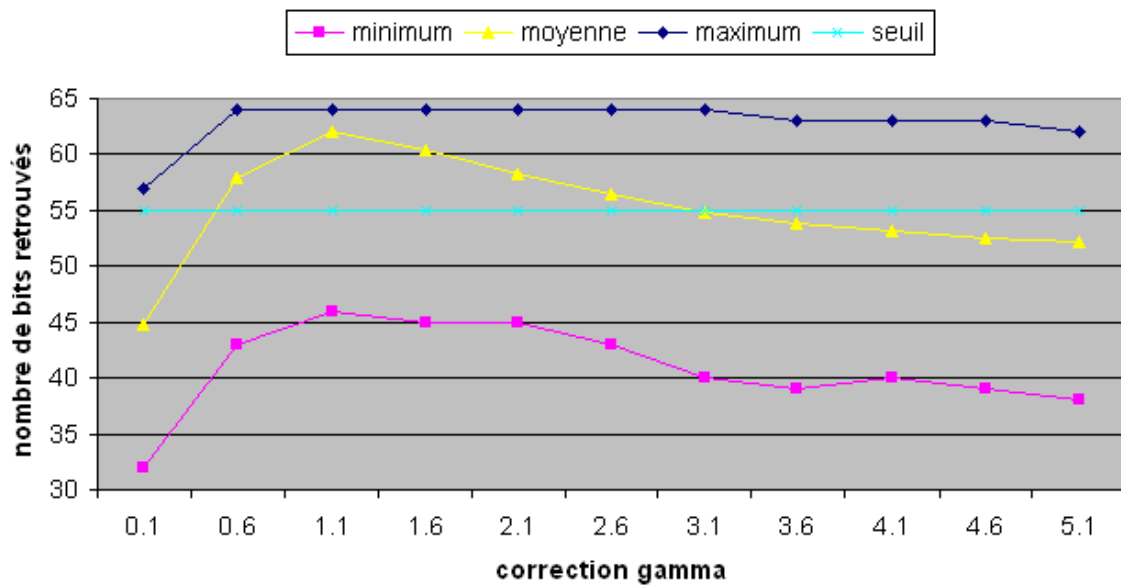


FIGURE 5.11 – Evolution du coefficient de détection correcte en fonction de la modification du gamma de l'image

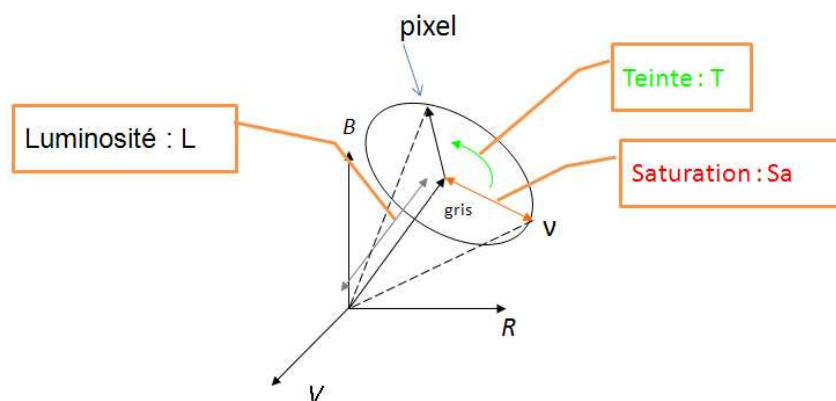


FIGURE 5.12 – Représentation HSV

couleur et l'axe des gris. Le paramètre de l'attaque de saturation va réduire ou augmenter cette distance en conservant la couleur elle-même ainsi que son intensité. Nous rappelons sur la figure 5.12 la représentation HSV des couleurs. Le paramétrage de l'attaque correspond à une simple dilatation/contraction suivant l'axe formé par le pixel couleur dans l'espace RGB et sa projection sur l'axe des gris (attention nous ne sommes pas sur le principe d'une translation suivant cette axe, mais bien un facteur multiplicatif). La valeur du paramètre correspond à un pourcentage de modification.

Lorsque le paramètre est négatif, l'image perd de son aspect coloré jusqu'à -100 pour devenir grise (figure 7.8). A l'inverse, lorsque le paramétrage de l'attaque est positif l'aspect coloré est renforcé voir saturé jusqu'à devenir artificiel (figure 7.8).

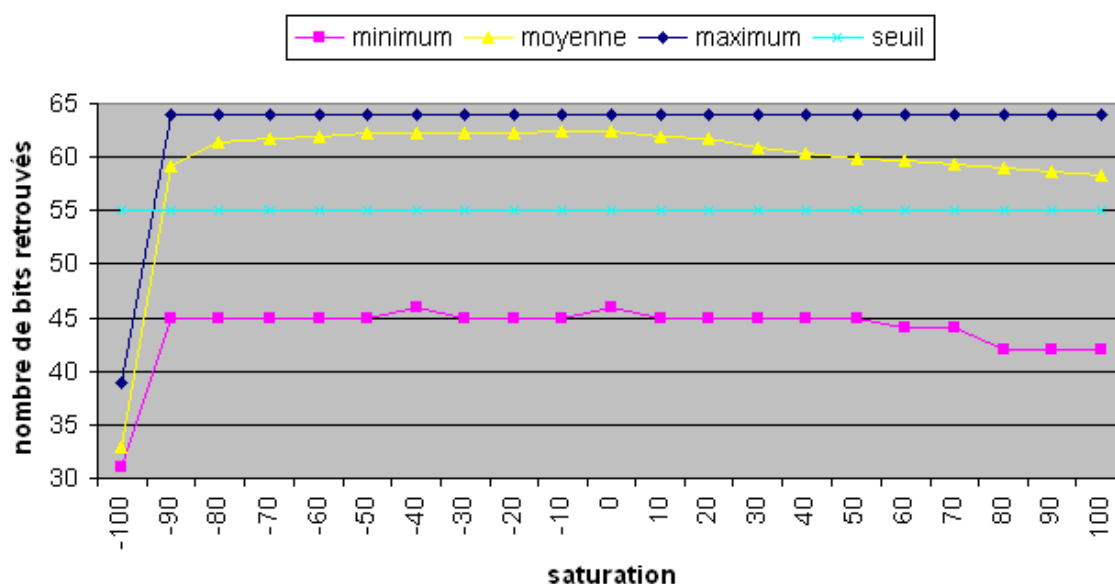


FIGURE 5.13 – Evolution du coefficient de détection correcte en fonction de la modification de la saturation

Cette attaque est fortement dépendante du contenu colorimétrique de l'image de départ. Dans le cadre des flux vidéo, nous pouvons voir sur la figure 5.13 que, sauf dans le cas extrême où l'image perd totalement son aspect coloré et ainsi n'est plus qu'en niveaux de gris, la modification de la saturation n'influence pas ou très peu la détection de la signature. La marque reste reconnaissable à près de 90%. Dans ce cas, le tatouage est donc robuste à la modification de la saturation, sous réserve d'une part de disposer d'un contenu d'image adapté et d'autre part d'utiliser des forces d'insertion qui permettent de lutter contre la dégradation.

Enfin, l'écrasement de la marque par le passage en niveaux de gris est prévisible puisque dans ce cas les trois vecteurs couleurs définis lors du tatouage vont se confondre et il n'est plus possible de retrouver la marque.

Modification du contraste. Les variations de contraste que nous avons fait subir à chaque image n'ont pas eu d'influence notable sur la détection de la marque comme nous pouvons le constater sur la figure 5.14. La modification de contraste se contente de dilater ou de réduire les trois histogrammes associés aux trois canaux, c'est pourquoi la marque continue à être détectable. Bien sûr, de nouveau, si cette modification de contraste entraîne un phénomène de dépassement de la capacité de codage (c'est le cas pour des valeurs positives), les résultats se dégradent.

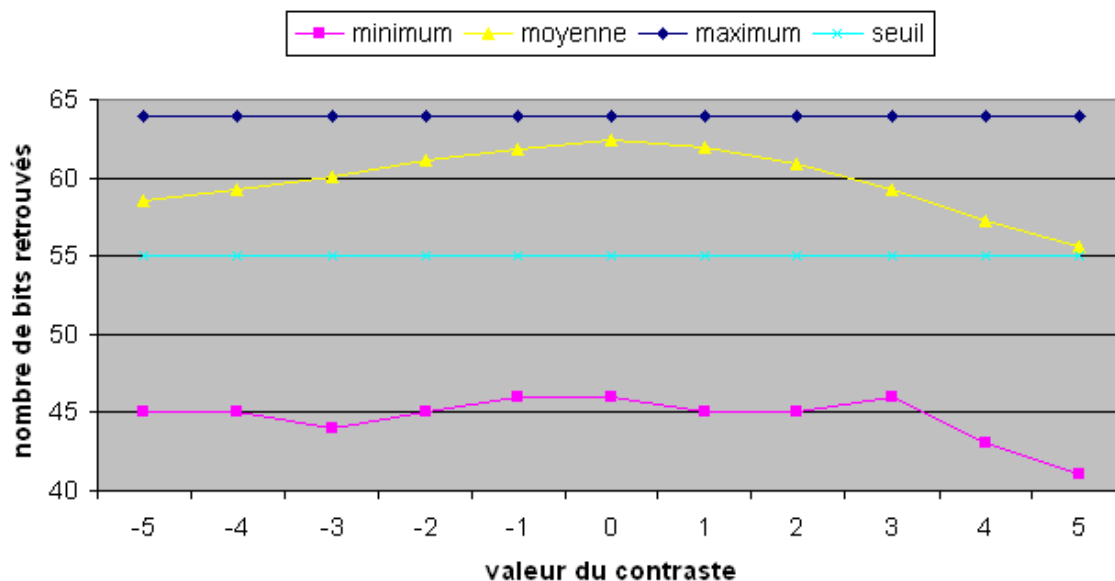


FIGURE 5.14 – Evolution du coefficient de détection correcte en fonction de la modification du contraste

Modification de la teinte La modification de la teinte va changer les couleurs présentes dans l'image par un mécanisme de rotation. Nous illustrons l'effet de cette attaque sur la figure 7.10 qui illustre une rotation de -100° et de $+100^\circ$. La stratégie de tatouage que nous avons mise en place est très fragile face à cette attaque, puisque la modification des couleurs va modifier l'affectation des vecteurs couleur telle qu'elle a été faite lors de l'insertion. Contrairement à la plupart des attaques que nous avons étudiées précédemment, la modification de la teinte va être destructrice pour notre tatouage sur certaines images. La détection reste acceptable dans le cas où la modification de la teinte n'excède pas un coefficient 10 comme nous pouvons le constater sur le graphique 5.15. Pour certaines images, cette attaque entraîne une perturbation qui peut être rapidement importante visuellement et dans ce cas le tatouage est détruit. Si l'image est très colorée la rotation des couleurs provoque une forte modification des coordonnées RGB (notamment pour certains angles) du pixel et donc une suppression de la marque. Le résultat est de nouveau fortement dépendant du contenu de l'image.

Filtrage passe-bas de type moyenne. Le filtrage passe-bas de type "calcul de moyenne" ne modifie pas suffisamment l'image pour influencer la détection de façon notable. En effet, il dégrade principalement les hautes fréquences et d'une manière moindre les moyennes fréquences (surtout pour une petite taille de masque). Etant données les forces de marquages utilisées, la marque est donc conservée, ce qui nous permet de dire que le marquage est robuste au filtrage du type passe-bas dans les limites de notre expérimentation.

Filtrage passe-bas de type médian. Contrairement au filtrage passe-bas de type calcul de moyenne, le filtrage médian modifie la marque de manière importante lorsque l'on atteint des tailles significatives. La détection est déjà compromise pour une taille de 7×7 . En effet, le processus est non linéaire et dans certaines zones avec peu de variations, le traitement va entièrement "lessiver" la marque. Cependant, pour ces tailles de masques, nous nous apercevons que l'image n'est plus véritablement reconnaissable. Nous pouvons considérer que le tatouage reste donc robuste face au filtrage médian pour des tailles de filtres raisonnables.

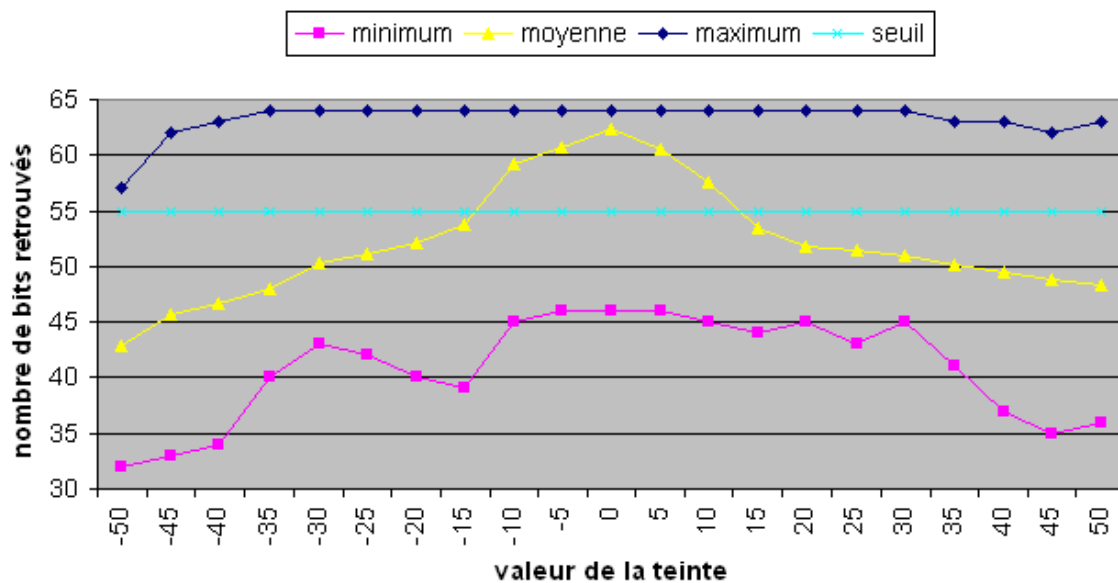


FIGURE 5.15 – Evolution du coefficient de détection correcte en fonction de la modification de la teinte

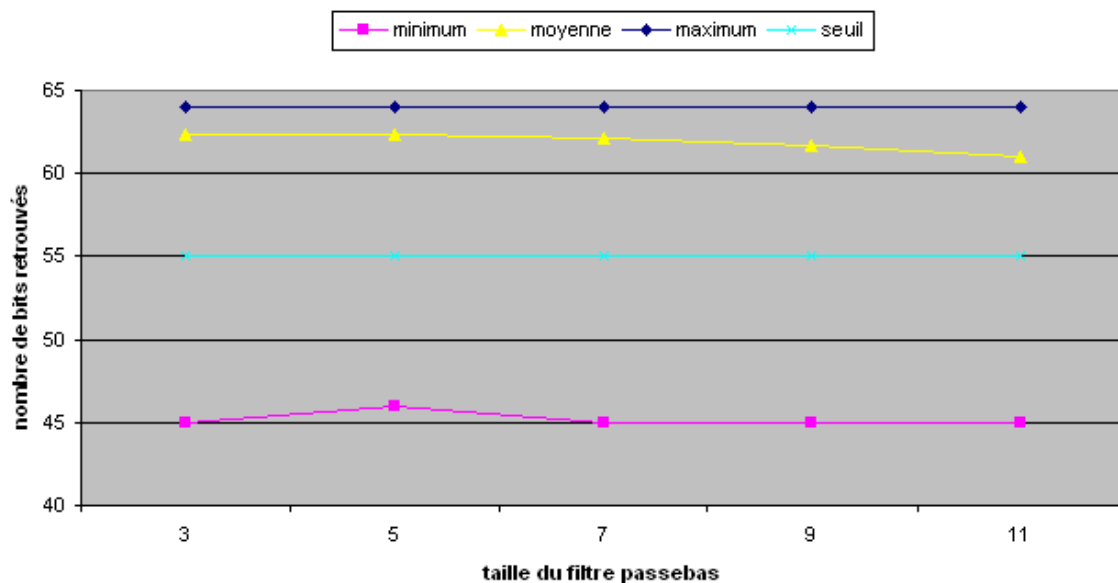


FIGURE 5.16 – Evolution du coefficient de bonne détection en fonction de la taille du filtre passe-bas

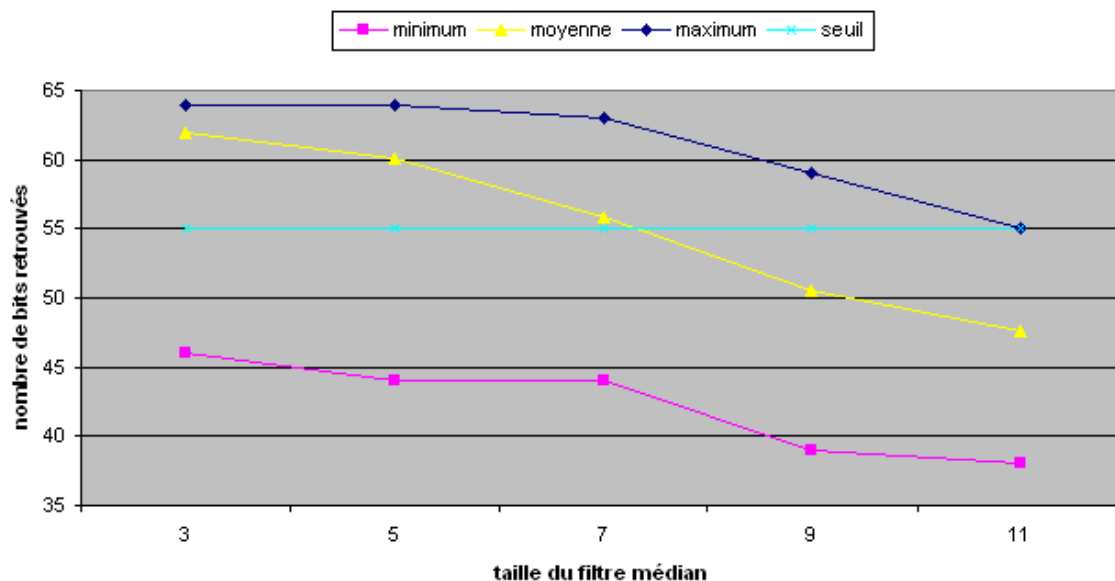


FIGURE 5.17 – Evolution du coefficient de la détection correcte en fonction de la taille du filtre médian

Quantification des couleurs.

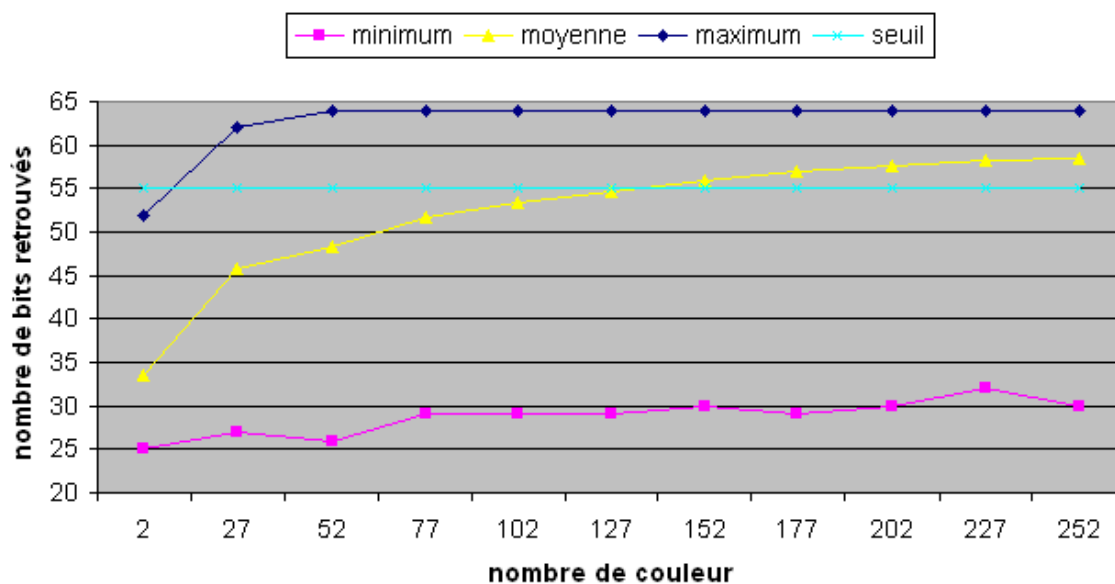


FIGURE 5.18 – Evolution du coefficient de détection correcte en fonction du nombre de couleurs

La réduction du nombre de couleurs ne modifie la détection que lorsque le nombre de couleurs après transformation est faible. Ceci est corrélé avec l'impression visuelle. Par exemple pour l'image "*lenna.bmp*" qui comporte 148279 couleurs au départ, la perturbation ne devient destructive que lorsque la palette de couleur est limitée à environ 128 couleurs. Dans ce cas, lorsque le nombre de couleurs devient faible, cela signifie à la fois que la couleur d'un pixel peut changer d'une façon importante, mais aussi que les ruptures (et donc les informations hautes et moyennes fréquences) sont réduites et enfin dans certains cas, des discontinuités parasites sont créées. Dans ce cas la marque est dégradée.

Modification de l'intensité. Comme nous l'avons vu, l'intensité correspond à la valeur de la projection du pixel sur l'axe des gris, et mesure ainsi une notion de clarté de la couleur. Le paramétrage va là aussi intervenir d'une manière multiplicative suivant l'axe formé entre le pixel à modifier et l'origine. De nouveau la transformation ne correspond pas à une translation du pixel, dans ce cas, parallèlement à l'axe des gris vers le haut ou le bas dans le cube RGB. La figure 7.14 illustre l'influence de la transformation. Lorsque l'attaque utilise un paramètre de -100 les pixels se confondent avec l'origine et l'image est entièrement noire. A l'inverse un paramètre positif provoque un éloignement de l'origine et donc à partir d'une certaine valeur de nouveau un phénomène de dépassement de capacité du codage.

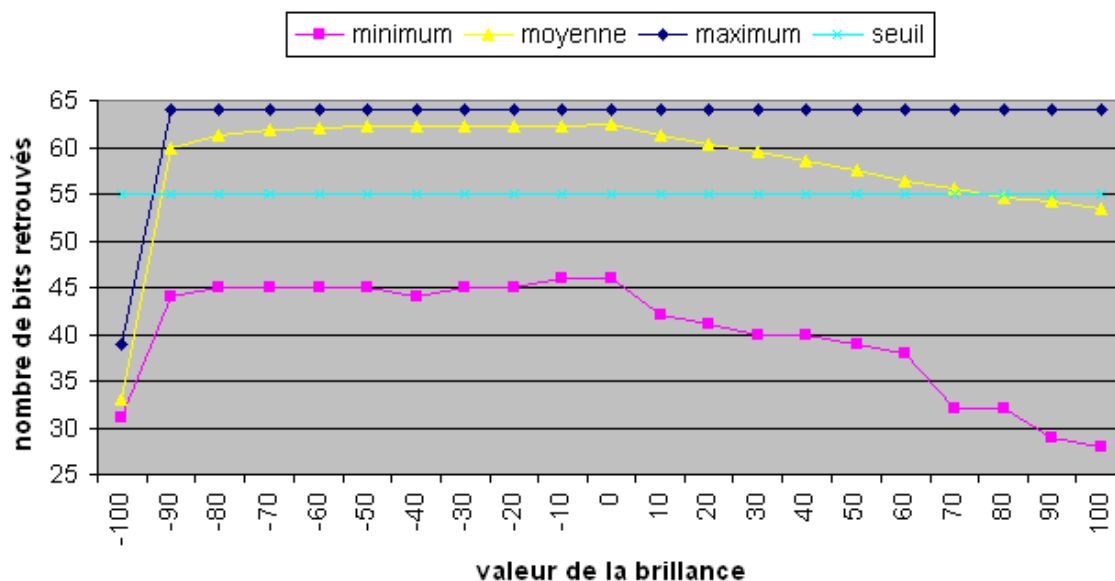


FIGURE 5.19 – Evolution du coefficient de détection correcte en fonction de la modification de la luminance

Comme nous pouvons l'observer sur le graphique 5.19, c'est la courbe des valeurs minimales qui est très fortement altérée lorsque le paramètre de l'attaque évolue. Ceci indique que la modification de la luminosité n'affecte que les images les plus fragiles.

Toutefois, nous pouvons constater que l'affaiblissement de la luminosité ne modifie pas la marque (sauf bien sûr avec un paramètre -100), alors que l'augmentation de cette luminosité a tendance à dégrader les résultats. Ceci s'explique simplement par l'algorithme utilisé pour modifier la luminosité. Lorsque l'on diminue la luminosité on réalise simplement une contraction des trois histogrammes RGB ce qui provoque bien une diminution de la luminosité mais sans modification de l'information. Lorsque l'on augmente la luminosité on réalise une dilatation des trois histogrammes et dans ce cas apparaît un phénomène de dépassement de capacité de codage qui lui est destructeur.

Pour conclure, nous avons déjà indiqué que le résultat de certaines attaques dépendent du contenu de l'image attaquée. Nous reverrons ce point dans la dernière partie de ce chapitre. De plus, certaines des images que nous avons sélectionnées ne sont pas acceptables pour notre tatouage. Ces images vont toujours obtenir les scores minimaux quelles que soient les attaques testées. Ces images sont des images globalement uniformes et pour certaines d'entre-elles, à dominante blanche ou noire (images 09, 11 et 12) avec une saturation nulle. Ces images ne permettent pas l'insertion de la marque dans de bonnes conditions (avec l'algorithme proposé). C'est pourquoi, les résultats obtenus sont toujours faibles par rapport aux autres images. Dans cette première partie, les attaques dites géométriques qui provoquent en général une désynchronisation

n'ont pas encore été abordées. Nous proposons de décrire maintenant le comportement de notre algorithme dans ce cadre là, et notamment les évolutions nécessaires que nous avons développées dans le projet de recherche France Telecom R&D, qui nous ont amenés à développer une étude sur les points d'intérêt.

5.3 Résistance aux attaques géométriques

Comme nous l'avons précisé, les attaques géométriques ont pour conséquence de désynchroniser la marque. A priori, la méthode non-décimée permet de résister aux attaques désynchronisant la marque par translation, puisque l'espace transformée est invariant par translation. Nous allons vérifier ceci dans le paragraphe suivant.

5.3.1 Attaques géométriques provoquant une translation



FIGURE 5.20 – Exemple d'un fenêtrage vertical effectué sur l'image "lenna.bmp"

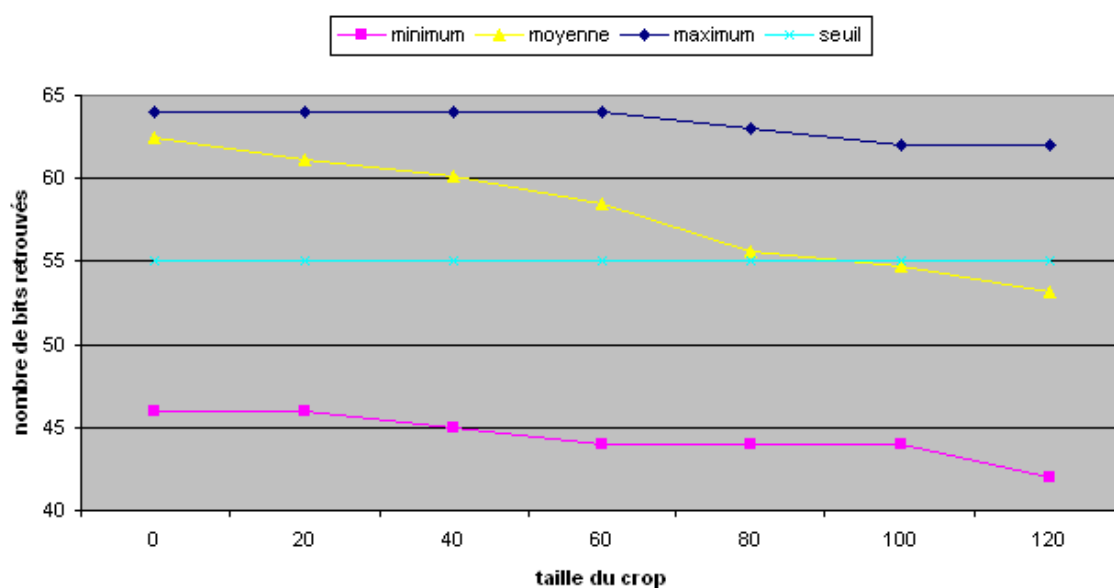


FIGURE 5.21 – Evolution du coefficient de détection en fonction de l'importance du fenêtrage

Découpage. L'attaque par fenêtrage (ou cropping) consiste à éliminer une partie de l'image. La figure 5.20 est un exemple de fenêtrage sur l'image "lenna.bmp". Il faut, pour récupérer la marque, que la taille de l'image après cropping soit encore suffisamment grande. C'est-à-dire que l'échelle 4 de la décomposition en ondelettes puisse contenir au moins une fois l'ensemble des coefficients de la signature. Comme nous pouvons l'observer sur le graphique 5.21, lorsque le cropping reste dans des proportions raisonnables, la détection non-décimée, associée à la duplication de la signature permet de retrouver en moyenne la signature.

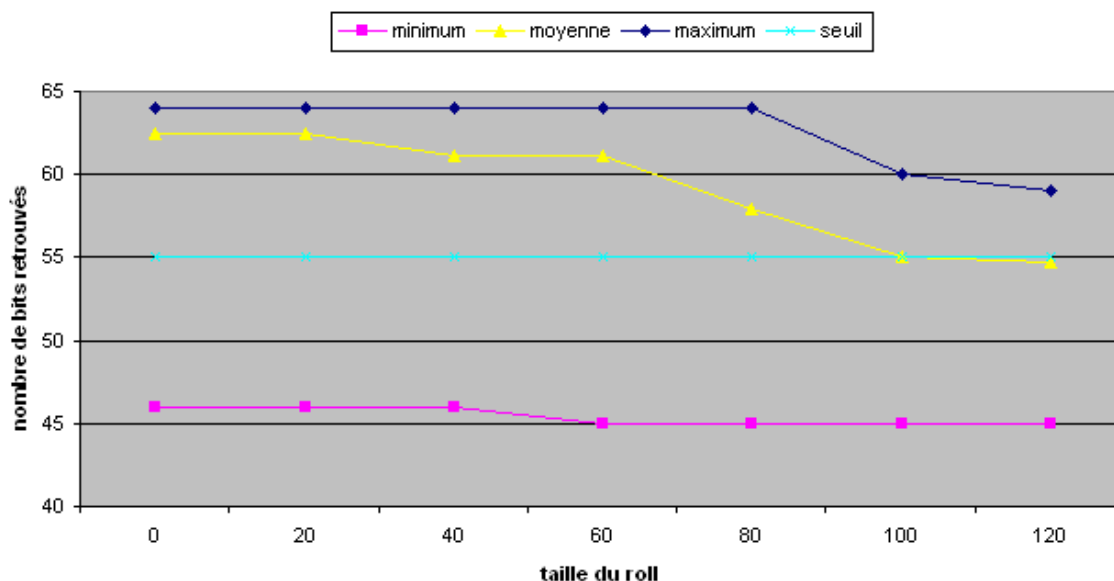


FIGURE 5.22 – Evolution du taux de détection en fonction du décalage

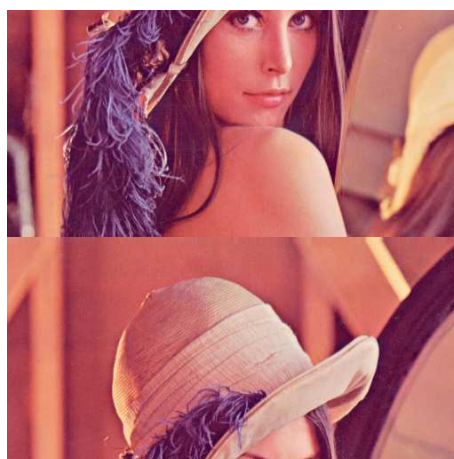


FIGURE 5.23 – Exemple d'une attaque de type roll sur l'image "lenna.bmp"

translation circulaire L'attaque par translation circulaire (ou rolling) présentée sur la figure 5.23, décale les coefficients de la marque par rapport à l'emplacement choisi initialement. Mais de nouveau cette attaque ne perturbe pas la détection puisque celle-ci va tester les différentes possibilités de positionnement de la marque et ne retient que celle dont la corrélation avec la marque d'origine est la plus forte.

Afin de conclure ce paragraphe nous proposons de confronter notre algorithme de détection face à une attaque géométrique qui désynchronise la marque d'une autre manière, à savoir une rotation.

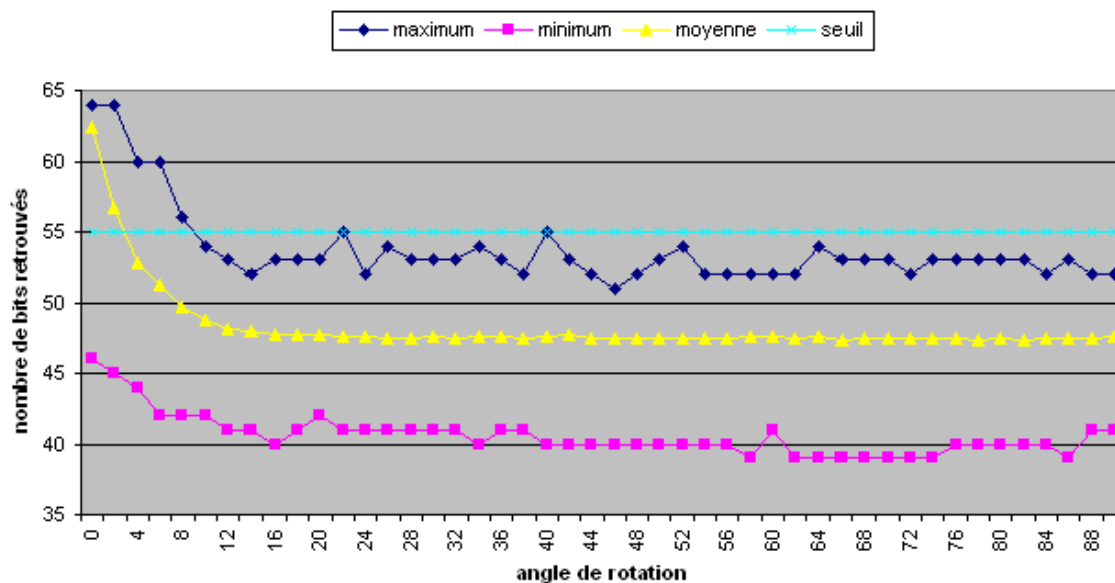


FIGURE 5.24 – Evolution du coefficient de la détection correcte en fonction de l'angle de rotation

Comme prévu, nous pouvons remarquer sur la figure 5.24 que les rotations sont problématiques pour la détection. En effet, l'espace d'insertion est invariant à la translation mais pas à la rotation ou aux changements d'échelle.

Afin de permettre une détection face à ce type d'attaque, il faut soit modifier l'algorithme de tatouage pour se positionner dans un espace invariant aux transformations considérées, soit effectuer un prétraitement qui va repositionner l'image tatouée dans son orientation d'origine. Nous avons orienté notre thématique de recherche vers la seconde catégorie, et les développements associés ont été faits dans le cadre de notre collaboration avec France Telecom R&D ; celle-ci a débouché sur la fourniture d'une solution logicielle complète.

5.3.2 Solution de la littérature contre les attaques par rotation et changement d'échelle

Une solution simple afin de contrer des attaques géométriques comme la rotation ou une opération de changement d'échelle consiste à calculer toutes les transformations inverses possibles et à garder le meilleur taux de détection. Dans le cadre des transformations provoquant une translation, les possibilités sont réduites et si nous intégrons directement cette recherche exhaustive dans la méthode de détection, comme cela est fait dans notre méthode non-décimée, le coût est réduit.

Dans le cadre par exemple d'un changement d'échelle, la solution n'est pas réaliste, ou alors le choix de la transformée inverse est guidé par une connaissance supplémentaire sur les données de départ (le cas le plus simple est bien sûr de posséder l'image d'origine).

Afin de prendre en compte les possibles modifications d'orientation ou de changements de taille, on peut trouver dans la littérature principalement deux stratégies :

- la première consiste à utiliser *a priori* un espace d'insertion invariant aux changements d'échelle ou aux rotations (par exemple la transformée Fourier-Mellin) ;

- la seconde consiste à estimer *a posteriori* les modifications subies par l'image et à appliquer la procédure inverse avant d'extraire la marque.

Les différentes solutions évoquées dans la littérature vont varier selon l'espace proposé ou le moyen d'estimer la transformation de l'image. Nous donnons quelques éléments de bibliographie permettant d'illustrer ces approches.

Le principe des techniques basées sur une transformation invariante va être par exemple d'utiliser des transformations pour lesquelles la notion de position spatiale n'est pas présente. C'est le cas notamment de la transformée de Fourier, souvent utilisée [93, 220, 219]. Une marque binaire est par exemple additionnée à l'amplitude résultante de la transformée de Fourier, marque consistant en une séquence aléatoire de la forme d'un anneau (dans le domaine de Fourier). D'autres transformées sont utilisées, comme la transformée 2-D de Radon-Wigner [231], ou alors la transformée de Fourier-Mellin [198]. Alors que la transformée de Fourier est sensible aux variations géométriques (rotation, translation, changement d'échelle), l'adaptation vers un système en coordonnées polaires logarithmiques permet de neutraliser cette sensibilité et le module de la transformée de Fourier-Mellin est alors invariant aux changements d'échelles et à la rotation. Toutefois, il faut noter que la réalisation numérique de cette transformée n'est pas simple (la notion de repère polaire est problématique en discret), comme nous l'avons vu dans le premier chapitre. Elle va nécessiter une procédure d'interpolation. Lin et al. [136] présentent, quant à eux, un algorithme de tatouage robuste à la rotation, à la translation et au changement d'échelle, en utilisant une fonction invariante à ces attaques, correspondant à une projection 1-D des amplitudes du spectre de Fourier, nous sommes proches de la transformée de Fourier-Mellin.

Dans la seconde catégorie de méthodes, nous trouvons une multitude de stratégies très différentes. Cependant une grande famille ressort qui consiste à insérer un motif possédant une certaine périodicité, à travers par exemple l'utilisation de la redondance de la marque. Par exemple, Delanay et Macq [60] proposent une méthode de génération de la marque ayant des propriétés statistiques cycliques afin de répondre aux problèmes de changement d'échelle et de fenêtrage. Dans ce cas, la signature n'a pas un support prédéfini et la détection de la signature s'effectue par le calcul de l'intercorrélation entre la signature et les différents pixels réordonnés. Hartung et al. [103] appliquent une insertion périodique de la marque dans le but de permettre une synchronisation et contrer ainsi l'attaque Stirmark. De nouveau un calcul de la corrélation de la marque est effectué en utilisant un corrélateur coulissant sur quatre dimensions, permettant ainsi de répondre aux attaques par rotation, translation et changement d'échelle. Kutter et al. [125, 123] insèrent périodiquement la marque dans l'image. La détection permet d'extraire une marque contenant des éléments redondants, qui, corrélée avec la marque initiale, permet d'évaluer les déformations géométriques linéaires appliquées à l'image marquée. La technique définie par Fleet and Hegger [89] est basée sur l'insertion d'une somme de sinusoides. Grâce à l'utilisation de ce signal périodique, il est possible d'extraire des informations sur les transformations géométriques opérées lors de la transmission de l'image. Cette technique illustre la stratégie présente dans les méthodes utilisant des motifs périodiques, à savoir la construction d'une grille dans le domaine fréquentiel.

Ensuite certaines techniques s'intéressent au contenu de l'image. Kutter [124], propose de tenir compte des points caractéristiques de l'image. De manière générale, l'ensemble de points caractéristiques permet de définir des zones particulières. Ces zones sont donc *a priori* robustes aux attaques géométriques. Elles porteront chacune la marque. Cette notion de tatouage de seconde génération est ainsi présentée comme une méthode compatible à toute technique de tatouage, permettant de la rendre plus robuste face à certaines at-

taques géométriques. Cette stratégie s'illustre pleinement dans la méthode proposée par P. Bas et al. [17] qui suggèrent d'utiliser les points saillants de l'image pour obtenir des points de référence nécessaires pour permettre la re-synchronisation de l'image. Nous pouvons citer d'autres méthodes utilisant la même stratégie. Par exemple, Izquierdo [113] combine des techniques d'insertion dans le domaine spatial et dans le domaine fréquentiel, afin de profiter de leurs caractéristiques de robustesse respectives contre les opérations classiques de traitement d'images et les transformations géométriques. Dans ce cadre, est proposée l'utilisation du détecteur de Harris [102] permettant ainsi d'effectuer le recalage de l'image tatouée, par comparaison avec l'image originale. Nikolaidis et Pitas [165] [166] sélectionnent quant à eux, à l'aide d'une segmentation, des zones larges et homogènes de l'image, schématisées, après traitement, par des ellipsoïdes. La marque sera intégrée dans les rectangles englobant ces ellipses. Enfin, comme nous l'avons dit certaines techniques utilisent l'image originale, pour identifier les transformations géométriques : F. Davoine [59], propose de comparer le maillage triangulaire de l'image attaquée avec celui de l'image initiale. Cette technique a pour but de rendre des algorithmes de tatouage robustes à l'attaque Stirmark. Nous remarquons que l'avantage de cette technique est de permettre de compenser les déformations géométriques locales de l'attaque Stirmark. En revanche, nous sommes dans le cadre d'une technique non aveugle de re-synchronisation.

Notre objectif était de faire cohabiter notre tatouage robuste à la compression, au filtrage, etc. avec une stratégie apportant de la résistance aux transformations géométriques. La solution que nous avons proposée repose sur une technique permettant d'identifier les déformations géométriques afin de les compenser (en particulier dans le cas de la rotation et du changement d'échelle). Nous avons choisi de ne pas intégrer d'information dans l'image, mais les paramètres nécessaires pour réaliser une détection synchronisée seront transmis avec la clé gérant la marque. Nous proposons de décrire maintenant notre stratégie.

5.3.3 Evolution de la méthode de tatouage vectorielle

Jusqu'ici, nous avons développé un algorithme de tatouage semi-aveugle c'est-à-dire que la clé et l'image tatouée sont les seuls éléments dont nous avons besoin lors de la détection pour retrouver la marque. A ce stade, la marque résiste correctement aux attaques par filtrage, par compression, aux attaques couleurs et même à certaines attaques géométriques de type translation mais elle est fragile par rapport aux transformations géométriques de type rotation et changement d'échelle. Nous avons alors introduit une méthode de recadrage géométrique destinée à resynchroniser la marque après ce type d'attaque.

Notre technique de recadrage est basée sur l'extraction de points stables face à la rotation ou aux changements d'échelle c'est-à-dire qu'après par exemple la rotation nous devons être capables de retrouver les mêmes points et que ces points ont subi la même transformation que l'image. Pour cela, nous proposons d'utiliser les points d'intérêt.

Cette technique se décompose en trois grandes étapes. La première étape consiste à choisir un détecteur de points d'intérêt géométriquement stable, c'est-à-dire qu'il doit permettre de retrouver les mêmes points après transformation géométrique. Une fois le détecteur choisi, nous devons caractériser le voisinage des points d'intérêt de manière robuste pour faciliter l'appariement qui constitue la troisième étape. Ces différentes parties nous donnent la possibilité de calculer les paramètres de la transformation géométrique utilisée pour attaquer le tatouage.

Cette technique a également été développée dans le cadre de la collaboration avec France Telecom. Les premières pistes avaient été proposées dans les perspectives de la thèse de A. Parisi [171]. Mais le cœur du travail a été réalisé dans le cadre du contrat de recherche en 2005-2006, évoqué en introduction de ce

chapitre, et notamment à travers le recrutement d'un ingénieur de Recherche car cette partie nécessitait une forte composante logicielle à travers la réalisation d'une librairie de détecteurs de points d'intérêt.

5.3.4 Les détecteurs de points d'intérêt

Nous proposons ici une synthèse du travail fait sur les points d'intérêts que nous avons étudiés en lien avec notre algorithme de tatouage, en insistant sur la dimension numérique car étant donné le rôle de ces points, à savoir une robustesse face à des attaques numériques géométriques, il fallait que les choix d'implantation soient clairs.

Les points d'intérêt sont des caractéristiques bas niveau qui ont un fort potentiel informatif. Dans la littérature, les termes de coins et de points d'intérêt sont souvent utilisés de manière équivalente bien que la notion de point d'intérêt est plus générale que la notion de coin. Les points d'intérêt, dans une image, correspondent à des changements bidirectionnels du signal. Les coins, les jonctions en T ou encore des variations rapides de la texture en sont des exemples.

La plupart des méthodes d'extraction des points d'intérêts sont basées sur l'utilisation des dérivées de l'image. Pour approximer les dérivées, une des approches classiques consiste à calculer la convolution de l'image par des filtres basés sur la fonction gaussienne 2D et ses dérivées. Nous utilisons ce type de filtres car ils permettent une localisation plus stable des points d'intérêts.

5.3.4.1 Différents détecteurs mono-échelle

Il existe différents types de méthodes pour effectuer l'extraction des points d'intérêt, comme par exemple :

- Les méthodes basées contours qui cherchent des points de courbure maximale le long des chaînes de contours, des points particuliers (intersection, inflexion, ...) dans une approximation polygonale de l'image.
- Les méthodes basées signal qui ne dépendent pas des contours ni d'un modèle du signal. La mesure qui indique s'il y a un point d'intérêt à un endroit donné est calculée directement à partir du signal.

Nous avons choisi de nous positionner dans cette seconde famille et nous rappelons quelques détecteurs de points d'intérêt faisant partie des méthodes basées signal, que nous avons testés dans le cadre du tatouage.

Le détecteur de Beaudet. Le détecteur de Beaudet est défini sur l'image I suivant :

$$DET = |I_{xx}I_{yy} - I_{xy}^2|$$

où I_{xx} , I_{yy} et I_{xy} représentent les dérivées secondes de l'image par rapport à x , y , et xy . Cette mesure est théoriquement invariante en rotation. Les points d'intérêt correspondent aux maxima locaux de DET. Il faut noter que cet opérateur détecte les points d'intérêt près des coins et non sur les coins.

Le problème de cette méthode est sa sensibilité au bruit, comme en général les méthodes basées sur des calculs de dérivées. Nous pouvons effectivement voir sur la figure 5.25 que l'ajout de bruit modifie la position des points d'intérêt.

Le détecteur de Kitchen-Rosenfeld. Le détecteur de Kitchen et Rosenfeld [119] est défini sur l'image I suivant :

$$K(x, y) = \frac{I_x^2 I_{yy} - 2I_x I_y I_{xy} + I_y^2 I_{xx}}{I_x^2 + I_y^2}$$



FIGURE 5.25 – Résultat de la détection des points d'intérêt de Beaudet pour l'image "house.bmp" et sa version bruitée.

où K représente la courbure d'une courbe plane perpendiculaire au gradient de la fonction image. Les coins sont extraits depuis les maxima locaux de K . De nos tests et des lectures bibliographiques, nous avons conclu que cet opérateur est très imprécis en localisation.

Le détecteur de Harris Moravec a proposé un détecteur basé sur la fonction d'auto-corrélation du signal. Cette fonction mesure les différences entre une fenêtre du signal et ses quatre fenêtres voisines. Lorsque le minimum de ces quatre différences est supérieur à un seuil, cela indique la présence d'un point d'intérêt.

Harris [102] a amélioré l'approche de Moravec en calculant une matrice qui décrit la distribution du gradient dans le proche voisinage du point étudié :

$$\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (5.17)$$

avec σ la variance d'une gaussienne qui permet le calcul de la somme pondérée des dérivées sur le voisinage et I_x, I_y les dérivées du signal par rapport à x et y . Un point est considéré comme un point d'intérêt si les deux courbures associées aux valeurs propres de cette matrice sont grandes. Pour éviter les calculs des valeurs propres, Harris utilise une mesure reposant sur le déterminant et la trace de la matrice. Cette mesure est supérieure à zéro dans le cas d'un coin. Cette approche est optimale pour les coins en forme de "L".

5.3.4.2 Détections des points d'intérêt multi-échelles

Les différents détecteurs présentés précédemment, bien qu'ils soient utilisables dans le cadre des rotations ne tiennent pas compte du changement d'échelle. Il nous faut donc les adapter au changement d'échelle ou utiliser d'autres types de détecteur. De plus l'ajout de l'aspect multiéchelle permet en général d'augmenter la robustesse du détecteur.

Nous avons choisi d'utiliser un détecteur basé sur la construction d'une pyramide de différences de gaussienne appelé détecteur DOG (Différence Of Gaussian).

Détecteur DOG

La détection de points invariants à un changement d'échelle de l'image peut être effectuée en recherchant

des points stables dans l'ensemble des échelles possible, en utilisant une fonction continue d'échelle appelée espace échelle. Il a été montré par Lindeberg [137] que sous certaines conditions le seul noyau d'espace échelle possible est la fonction Gaussienne. C'est pourquoi, l'espace échelle d'une image est défini comme une fonction $L(x, y, \sigma)$ obtenue par la convolution d'une image $I(x, y)$, avec un ensemble de filtres gaussiens d'écart type différents, $G(x, y, \sigma)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5.18)$$

Dans ce cadre, une version multiéchelle du détecteur de Harris a été proposée dans [160].

Afin de détecter les points stables dans l'espace échelle Lowe [138] a proposé d'utiliser les extrema de l'espace échelle dans le résultat de la convolution d'une fonction différence de gaussienne avec l'image $D(x, y, \sigma)$ qui peut être réalisée à partir de la différence de deux échelles proches séparées par un facteur multiplicatif constant k :

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (5.19)$$

Il y a plusieurs raisons pour choisir cette fonction. Premièrement, c'est une fonction particulièrement efficace en terme d'implantation puisque l'image lissée est nécessaire pour la description de l'espace échelle et que D peut être calculée par simple soustraction d'image. De plus, cette fonction est une approximation satisfaisante du Laplacien de gaussienne normalisé en échelle étudié par Lindeberg [137]. Lindeberg a montré que la normalisation du Laplacien avec le facteur σ^2 est nécessaire pour une véritable invariance en échelle.

Des choix numériques Pour commencer, nous devons construire l'ensemble des images lissées. Pour cela, nous devons définir la manière la plus efficace de discrétiser l'espace échelle. Lowe [138] a proposé de structurer cette discrétisation sous forme d'une pyramide contenant des octaves (le facteur de lissage σ double d'une octave à l'autre) qui sont subdivisées en étapes de lissage plus fines. En pratique, nous avons fixé le nombre de subdivision de chaque octave à 3 et le nombre d'octave est directement lié à la taille de l'image à traiter ($\log_2(dim) - 2$).

Pour initialiser la création de la pyramide, l'image initiale est sur-échantillonnée par deux puis lissée avec une gaussienne d'écart type σ_{init} . Ensuite, nous créons une octave en lissant l'image obtenue par des gaussiennes dont l'écart type croît d'un facteur constant (fixé par le nombre d'images par octave). Dans notre cas nous calculons donc 3 images par octave. La première image de l'octave suivante est obtenue en effectuant un sous échantillonnage d'un facteur deux de l'image lissée pour $2 * \sigma_{init}$. Puis le reste de l'octave est construit de la même manière que décrit précédemment. La figure 5.26 résume cette construction. Une fois la pyramide construite, nous pouvons calculer la différence de gaussienne par simple soustraction de deux échelles adjacentes.

Afin de détecter les extrema locaux de $D(x, y, \sigma)$, chaque coefficient de la pyramide est comparé à ses huit voisins dans l'image courante et ses neuf voisins dans les échelles adjacentes (cf figure 5.27). Le coefficient est retenu comme point d'intérêt s'il est extremum dans le cube ainsi défini.

Pour améliorer la qualité de la localisation des points d'intérêt, Lowe a proposé d'interpoler la position exacte de l'extremum dans la pyramide, à travers un modèle continu et pour cela, il utilise le développement

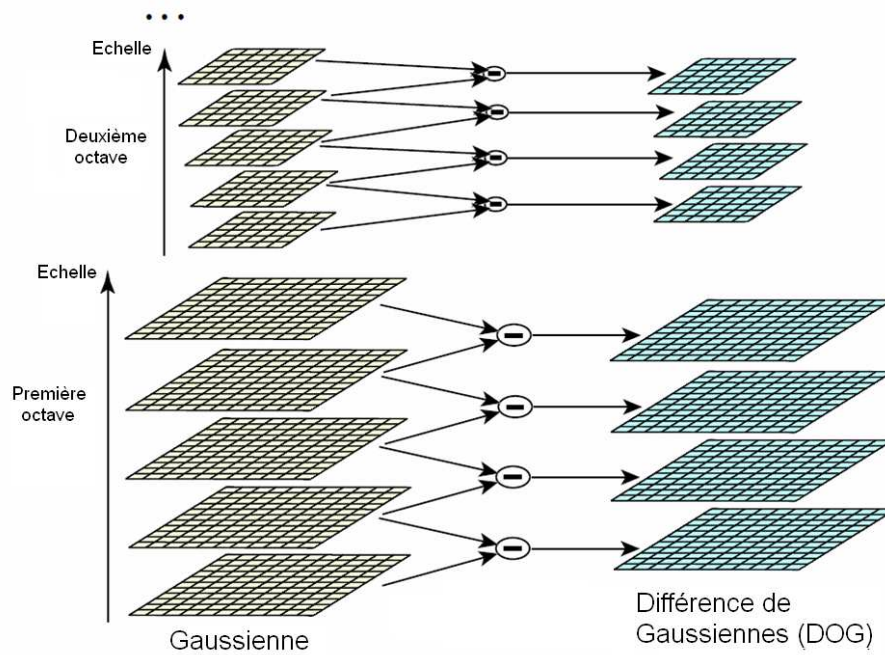


FIGURE 5.26 – Illustration de la construction des pyramides gaussiennes et différences de gaussienne (Image issue de l'article de Lowe)

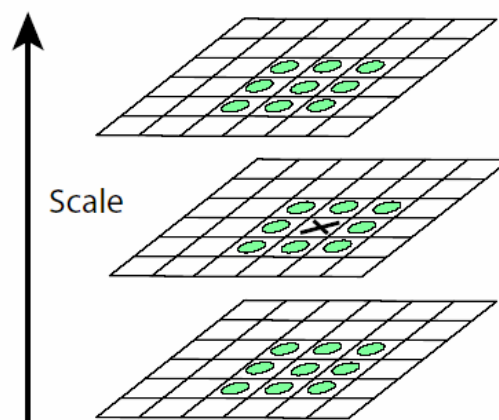


FIGURE 5.27 – Illustration de la détection d'un extremum dans la pyramide DOG

de Taylor à l'ordre 2 de la fonction d'espace échelle $D(x, y, \sigma)$ au point analysé

Enfin, pour augmenter la stabilité des points obtenus, il est proposé d'éliminer les points dont le rapport des courbures principales est trop grand. Cela se produit par exemple pour les points de contour qui ont une courbure principale forte perpendiculairement au contour mais faible dans la direction du contour. Pour détecter ces points, nous nous servons de la matrice Hessien évaluée à l'échelle et à la position du point d'intérêt. La matrice Hessien H contient les dérivées d'ordre deux de l'image lissée :

$$H = \begin{bmatrix} I_{xx,\sigma} & I_{xy,\sigma} \\ I_{xy,\sigma} & I_{yy,\sigma} \end{bmatrix} \quad (5.20)$$

avec par exemple $I_{xx,\sigma}$ la dérivée seconde selon x de l'image lissée par une Gaussienne d'écart-type σ .

En effet, il est connu que les valeurs propres de cette matrice sont proportionnelles aux courbures principales. Nous souhaitons supprimer tous les points dont le rapport entre les courbures principales est trop grand, supérieur à un seuil r . Pour éliminer ces points, il suffit de vérifier l'inéquation suivante, pour r fixé :

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (5.21)$$

De nos tests nous préconisons pour des images "classiques" (comme sur la figure 5.3.4.2) une valeur $r = 10$.

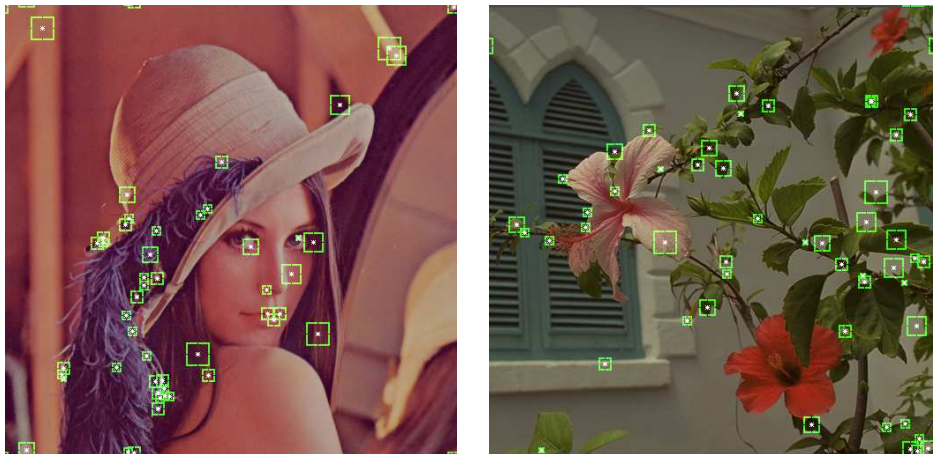


FIGURE 5.28 – Exemples de détection des points d'intérêt extraits de la pyramide DOG

Deux exemples de résultats de la détection des points d'intérêt dans la pyramide DOG sont présentés sur la figure 5.3.4.2. Un point d'intérêt est symbolisé par un carré dont le centre est le pixel le plus proche du véritable point d'intérêt et dont la longueur du côté est proportionnelle à l'échelle d'extraction du point.

La détection des points d'intérêt reste un véritable challenge. Nous l'avons déjà abordé dans le chapitre précédent dans le cadre d'une méthode de compression, simplement nous remarquons que les contraintes sont différentes : dans le cadre d'un algorithme de compression, nous cherchons des points de focalisation au sens du système visuel humain, alors que pour l'algorithme de tatouage nous cherchons des points "géométriques" qui sont stables face à des modifications de l'image.

Avec notre algorithme de tatouage, notre objectif était d'introduire un détecteur robuste ayant déjà fait ses preuves d'où l'utilisation d'une pyramide DOG. Il nous faut maintenant caractériser ces points afin de permettre un appariement lors de la détection.

5.3.5 Description du voisinage des points d'intérêt

Maintenant que nous avons extrait des points d'intérêt, il nous faut décrire le signal au voisinage de ce point d'intérêt de la manière la plus précise et complète. De plus il faut que cette description soit résistante aux transformations usuelles afin de permettre un appariement quelle que soit la transformation effectuée sur l'image. Nous avons testé deux types de descripteur : le Jet Local et le SIFT (Scale Invariant Feature Transform).

5.3.5.1 Le Jet Local

Le Jet Local correspond à l'estimation des dérivées localement autour du point étudié à une échelle donnée. A partir de ces dérivées différents auteurs ont proposé de les combiner pour obtenir un vecteur invariant au rotation [90].

$$V_{invariant} = \begin{bmatrix} L \\ L_x^2 + L_y^2 \\ L_{xx}L_x^2 + 2L_{xy}L_xL_y + L_{yy}L_y^2 \\ L_{xx} + L_{yy} \\ L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2 \\ L_{xxx}L_y^3 + 3L_{xyy}L_x^2L_y - 3L_{xxy}L_xL_y^2 - L_{yyy}L_x^3 \\ L_{xxx}L_xL_y^2 + L_{xxy}(L_y^3 - 2L_x^2L_y) + L_{xyy}(L_x^3 - 2L_xL_y^2) + L_{yyy}L_x^2L_y \\ L_{xxx}L_x^2L_y + L_{xxy}(2L_xL_y^2 - L_x^3) - L_{xyy}(2L_x^2L_y - L_y^3) - L_{yyy}L_xL_y^2 \\ L_{xxx}L_x^3 + 3L_{xxy}L_x^2L_y + 3L_{xyy}L_xL_y^2 + L_{yyy}L_y^3 \end{bmatrix} \quad (5.22)$$

où L est la convolution de la composante de luminance par un filtre gaussien d'écart-type σ et L_x L_y ses dérivées par rapport à x et à y . Notons que les auteurs soulignent la nécessité d'implanter le calcul des dérivées de manière stable pour pouvoir calculer ces invariants à un ordre élevé.

Ce vecteur est en théorie invariant à la rotation ainsi qu'à un changement d'échelle de maximum 20%. Afin de le rendre plus stable face au changement d'échelle, nous devons utiliser un espace multi-échelle. Comme nous ne pouvons pas calculer le vecteur pour toutes les échelles possibles il nous faut discrétiser l'espace. De plus, nous avons vu que la caractérisation proposée résiste à un changement d'échelle de maximum 20% donc nous utilisons une discrétisation qui garantit le fait qu'entre deux échelles consécutives, le changement est inférieur à 20%. C. Schmid [208] dans sa thèse utilise neuf échelles qui ont pour valeur 0.48, 0.58, 0.69, 0.83, 1, 1.2, 1.44, 1.73, 2.07 (ces valeurs sont en fait les différentes tailles du paramètre σ de la gaussienne).

Une fois l'extraction et la caractérisation des points d'intérêt effectuées, il nous faut faire la mise en correspondance des points d'intérêt de l'image originale avec ceux de l'image transformée. Il existe de très nombreuses stratégies dans la littérature, de nouveau nous avons fait le choix de la simplicité. Cette phase d'appariement va s'effectuer à plusieurs niveaux. La première étape consiste à calculer la plus petite distance entre un vecteur de l'image d'origine et un vecteur de l'image transformée. Pour cela, nous utilisons la distance de Mahalanobis. Cette distance, doit aussi tenir compte des différentes échelles, c'est pourquoi la recherche est exhaustive à travers les échelles. Cette méthode élimine beaucoup de points non valables mais n'est pas encore suffisante pour obtenir une majorité de mises en correspondances correcte, c'est pourquoi nous utilisons une contrainte supplémentaire qui porte sur le voisinage du point considéré. En effet, nous

pouvons estimer que les points d'intérêt voisins du point considéré doivent être similaires dans l'image de référence ou dans l'image transformée. Toutefois, comme il peut y avoir du bruit dans la détection, nous ne nous obligeons pas à avoir exactement le même voisinage. Nous nous fixons un seuil sur le nombre de voisins correctement appariés pour que le point étudié soit conservé. Nous imposons que 50% des voisins se correspondent pour que la correspondance globale soit validée.

Nous présentons sur la figure 5.29 des exemples d'appariement pour l'image Lenna. On constate que l'appariement ne fonctionne pas correctement. En effet, si pour des rotations simples 0, 90, 180 et 270 degrés, l'appariement est quasi-parfait, pour les autres rotations, nous obtenons encore de trop nombreuses erreurs pour permettre le calcul de l'angle de rotation. Nous avons constaté les mêmes taux d'erreur pour les changements d'échelle.

Les tests menés sur d'autres images nous ont conduits au même échec. Nous avons alors proposé, pour améliorer notre détection, d'utiliser le descripteur SIFT.

5.3.5.2 Le descripteur SIFT

Le descripteur SIFT est une description locale reposant sur l'utilisation des orientations du voisinage du point considéré.

Pour un point d'intérêt, nous étudions son voisinage 16×16 dans la pyramide gaussienne. Pour chaque coefficient du voisinage, l'orientation et la norme du gradient sont calculées, puis nous construisons un histogramme des orientations sur 16 imagerie de taille 4×4 , histogramme qui est défini par pas de 45 degrés. Les 16 histogrammes obtenus, concaténés les uns avec les autres, nous donnent le vecteur SIFT. Pour que ce vecteur soit le plus possible invariant aux changements de luminosité qui modifient le module du gradient, nous normalisons le vecteur suivant l'intensité. Nous pouvons voir un exemple de construction du vecteur SIFT (pour une orientation donnée) sur la figure 5.30.

Nous avons introduit quelques choix numériques. Tout d'abord nous diminuons l'influence des orientations du bord du voisinage par rapport aux orientations du centre en utilisant un facteur d'éloignement. Ensuite nous proposons de lisser l'histogramme, c'est-à-dire qu'au lieu d'ajouter un élément à un seul intervalle de l'histogramme on le répartit dans au moins deux intervalles.

Nous avons maintenant la possibilité d'extraire et de caractériser les points d'intérêt dans une image. L'étape suivante consiste à appairer les points extraits dans l'image d'origine et les points extraits dans l'image après l'attaque. Il existe de nouveau une littérature importante sur ces problèmes d'appariement dans le cadre du descripteur SIFT. Nous avons déployé une méthode simple, reposant sur la sélection du plus proche voisin dans l'espace des descripteurs (recherche des paires de points pour lesquelles la distance euclidienne est minimale dans les deux sens). Nous pouvons observer sur la figure 5.31 quelques exemples d'appariements effectués pour des attaques géométriques variées. Nous constatons que nous avons un taux de réussite satisfaisant dans la majorité des cas, sauf pour l'attaque la plus désynchronisante (rotation de 45 et changement d'échelle). Mais comme nous allons le voir le nombre correct d'appariements nous suffit.

Pour déterminer les paramètres de l'attaque utilisée, nous allons calculer ces paramètres pour chaque couple de points. Nous faisons l'hypothèse que les coordonnées des points sont exprimées dans un référentiel commun dont l'origine est le centre de l'image que l'on suppose conservé lors des différentes attaques. Nous allons construire deux histogrammes : l'un contenant les angles d'attaques possibles suivant des intervalles de 2 degrés, l'autre contenant les facteurs de changement d'échelle entre 0 et 20 avec un pas de 0.01. Les pics de l'un et de l'autre des histogrammes nous donnent les valeurs de l'angle et du changement d'échelle

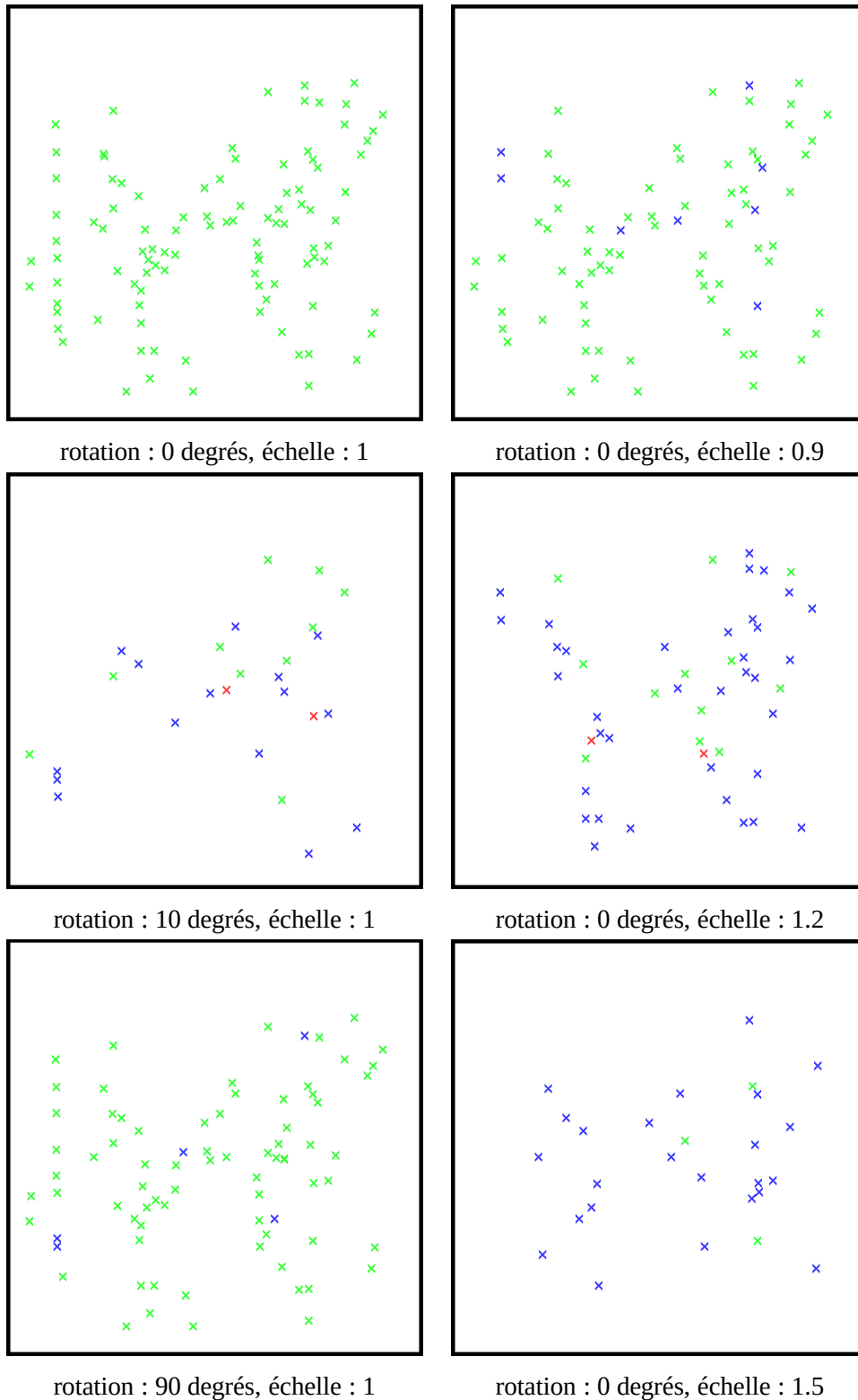


FIGURE 5.29 – Résultat de l'appariement pour des rotations de 0, 10 et 90 degré de l'image *lena.bmp* (vert : points bien appariés, rouge : points mal appariés, mais avec possibilité d'appariement valide, bleu : points pas ou mal appariés sans appariement valide)

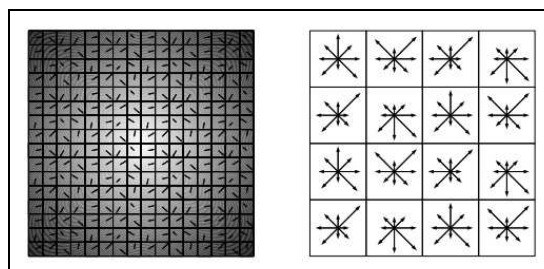


FIGURE 5.30 – Détection des directions principales sur l'histogramme

de l'attaque.

Nous avons vu sur les exemples de la figure 5.31 que le nombre d'appariements corrects est généralement supérieur au nombre de faux appariements, sauf dans le cas de l'image *mandrill.bmp* pour une attaque dont le facteur d'homothétie est de 1.2 et l'angle de rotation de 45 degrés. Mais même dans ce dernier cas, notre méthode permet une bonne détection car, comme nous pouvons le voir sur l'histogramme de la figure 5.32, l'angle d'attaque est mesuré en majorité. Cela s'explique par le fait qu'il n'y a aucune cohérence dans les angles mesurés pour les couples de points associés à un appariement faux. Ces valeurs vont donc se disperser sur tout l'histogramme. Le même phénomène est présent pour les facteurs de changement d'échelle.

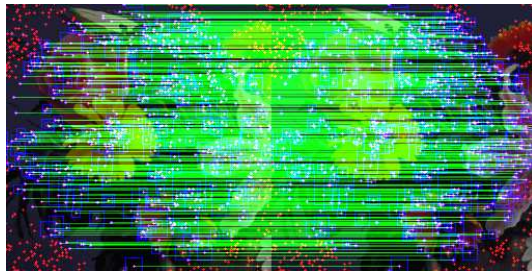
Les expérimentations menées dans le cadre des différentes images de tests ont montré une efficacité importante de cette procédure de recalage pour des attaques de paramètres raisonnables, notamment en ce qui concerne le changement d'échelle. Nous avons donc intégré cette étape dans notre méthode de tatouage.

5.3.6 Intégration de la mesure SIFT dans le cadre du tatouage

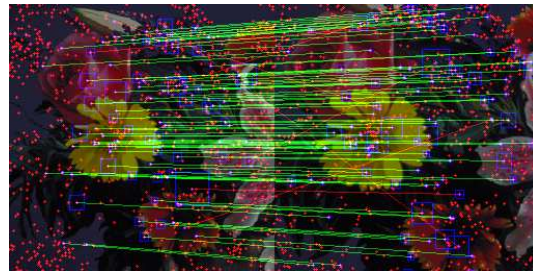
Pour illustrer la méthode de resynchronisation dans le cadre du tatouage vectoriel, nous insérons une marque 8*8 de manière robuste dans l'image *mandrill.bmp*. La figure 5.33 montre le résultat de la détection de la marque dans l'image ayant subi une rotation de 45 degrés et un changement d'échelle d'un facteur 2.0. Nous pouvons voir que la détection des paramètres d'attaque a été suffisamment précise et que la marque est correctement extraite (6 bits d'erreur sur 64 bits insérés), pour une attaque entraînant une forte désynchronisation.

Ce résultat s'est généralisé sur les différentes images de test pour des attaques qui conservaient des facteurs raisonnables (par exemple pour le changement d'échelle entre 0.8 et 1.2), c'est-à-dire que pour les différentes attaques testées nous pouvions extraire la signature avec un taux d'erreur inférieur au seuil fixé dans la section précédente. Nous ne présentons pas ici de résultats exhaustifs sous forme d'un tableau car les résultats de ce type de campagne de mesures dépendent principalement du contenu de l'image à travers la qualité des points détectés et moins de la méthode elle-même.

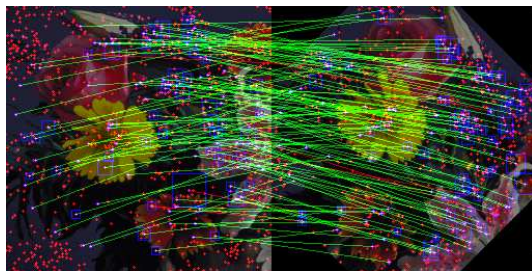
A ce stade, nous concluons que nous avons construit une méthode permettant de corriger une image ayant subi des rotations d'angle quelconque et des changements d'échelle. Cette méthode a prouvé son efficacité dans le cadre que nous nous sommes fixés. Cependant il reste encore des défauts, tels que le temps de calcul et la taille de fichier importante pour pouvoir stocker l'ensemble des points avec leur description. Nous avons ainsi établi une méthodologie de tatouage et de détection qui permet de parer à de nombreuses attaques. Tous ces développements ont donné lieu au développement d'une solution logicielle. Nous avons ainsi à notre disposition un démonstrateur, qui a été fourni à France Telecom R&D avec une étape de suivi de produit.



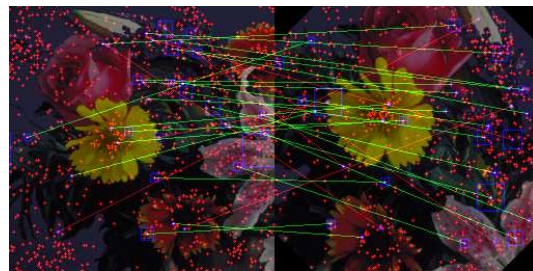
échelle 1.0 rotation 0
appariements justes : 609
appariements faux : 0



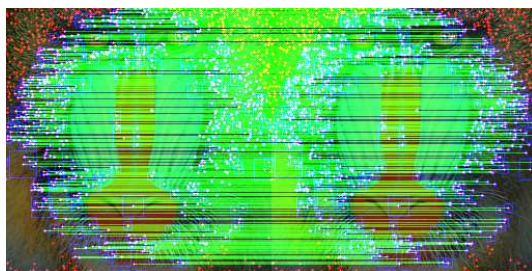
échelle 1.2 rotation 0
appariements justes : 84
appariements faux : 7



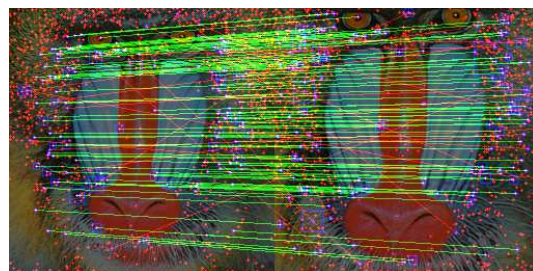
échelle 1.0 rotation 45
appariements justes : 116
appariements faux : 6



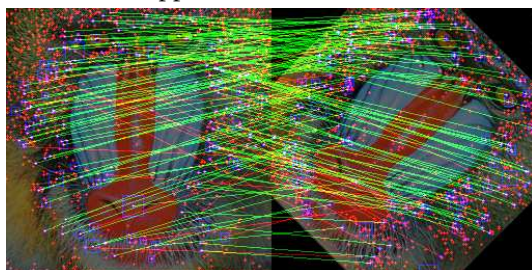
échelle 1.2 rotation 45
appariements justes : 23
appariements faux : 9



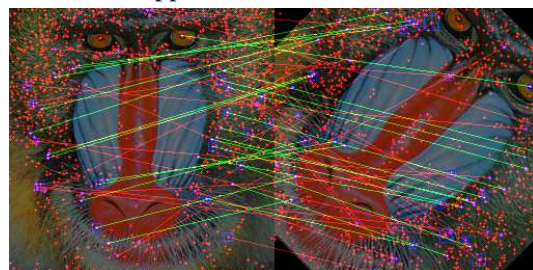
échelle 1.0 rotation 0
appariements justes : 677
appariements faux : 0



échelle 1.2 rotation 0
appariements justes : 116
appariements faux : 26



échelle 1.0 rotation 45
appariements justes : 125
appariements faux : 27



échelle 1.2 rotation 45
appariements justes : 25
appariements faux : 27

FIGURE 5.31 – Les points symbolisent les points d'intérêt (bleu : apparié, rouge : non apparié) ; Les lignes symbolisent les appariements (vert : appariement correct, rouge : mauvais appariement)

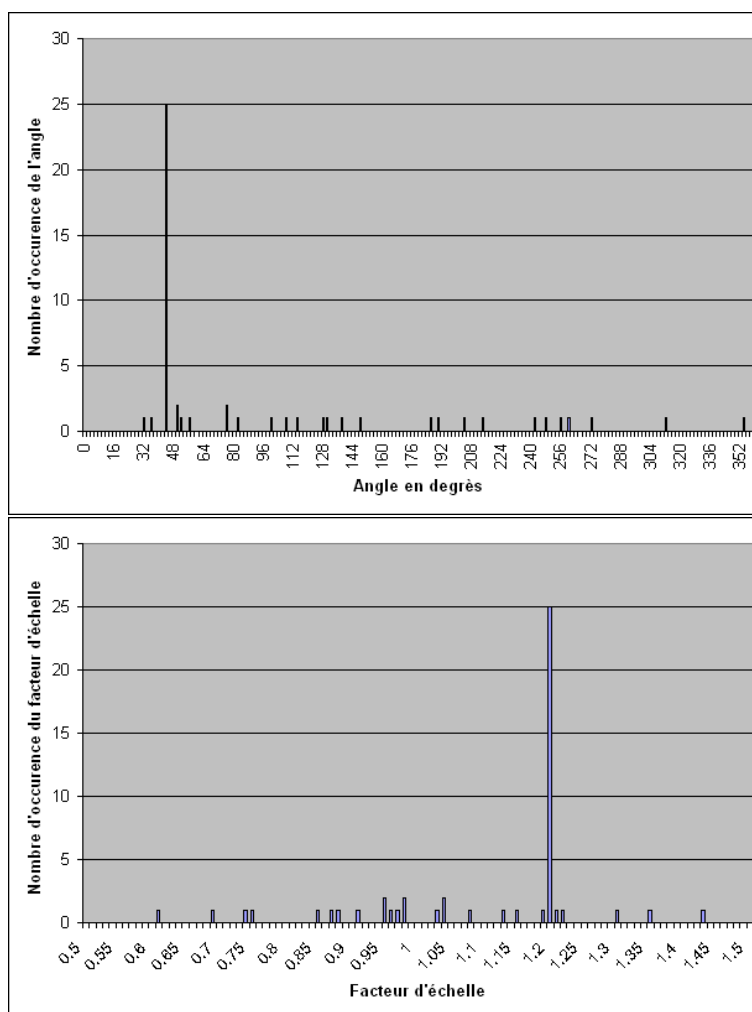


FIGURE 5.32 – Histogramme des angles d'attaques possibles et Histogramme des facteurs d'échelle possibles

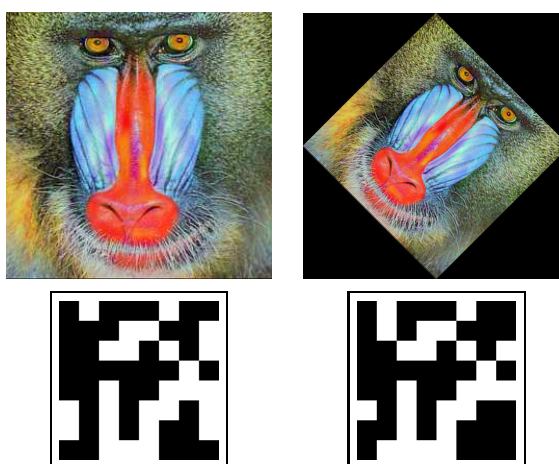


FIGURE 5.33 – Illustration de la méthode de resynchronisation face à des attaques géométriques dans le cadre du tatouage

Cependant, comme nous l'avons vu, la décision d'acceptation ou non de la signature reste basée sur des hypothèses fragiles. C'est pourquoi nous avons développé une réflexion sur la protection de la signature. L'opportunité est venue de la création d'un groupe de travail sur la sécurité dans le cadre d'une action du plan pluri-formation (PPF2008-2011) intitulé " Géométries-Images-Communications " incluant les universités de Poitiers, La Rochelle, Limoges, Tours, Orléans, ainsi que l'Ensma, action que je co-animais avec Philippe Gaborit du département DMI de XLIM. P. Gaborit a notamment comme domaine de recherches les codes correcteurs et c'est tout naturellement que nous avons débuté une collaboration afin d'intégrer certains développements menés autour des codes dans le schéma de tatouage afin d'augmenter la robustesse de la signature, mais aussi de modéliser plus finement la procédure d'acceptation de la marque. Nous proposons de décrire ce développement dans la section suivante.

5.4 Augmentation de la robustesse par utilisation des codes correcteurs

Différents travaux ont proposé d'augmenter la robustesse d'une stratégie de tatouage par l'utilisation des codes correcteurs [170, 57, 22, 238]... Le problème du tatouage est alors vu comme un problème de transmission d'un signal dans un canal très bruité. Le signal à transmettre est alors la signature, l'image support est le canal et le bruit est lié aux attaques.

Dans cette section, nous proposons de revoir les solutions développées notamment dans le cadre de la thèse de A. Wadood [1] à travers la collaboration menée avec P. Gaborit. L'objectif de ces travaux était de mettre en lumière la résistance de différents codes correcteurs dans notre schéma de tatouage par ondes face à différentes attaques, notamment, élément nouveau par rapport à la littérature, les attaques que nous nommerons couleur car elles modifient les caractéristiques colorimétriques de l'image. Ces travaux se poursuivent actuellement avec l'utilisation de nouvelles stratégies de protection mais aussi à travers la stéga-analyse comme je le décrirai à la fin de ce chapitre.

Dans ce cadre, nous avons étudié plus spécifiquement quatre familles de codes correcteurs, chacune des familles offrant des réponses différentes face aux attaques variées auxquelles elles ont été soumises. Nous avons proposé de mettre en concurrence ces différentes familles de codes car les attaques modifient différemment la marque insérée et les codes correcteurs ont des capacités de réponse variées face à ces différentes dégradations (par exemple un effacement par bloc ou un bruit additif aléatoire sur toute la signature). Ces quatre familles sont :

- les codes par duplication simple des bits (ce que nous avons introduit dès la première solution) ;
- les codes de Hamming [100] ;
- les codes Bose Chaudhuri Hocquenghem (BCH) [27] ;
- et les codes Reed-Solomon [193].

Ces différents codes ont déjà été testés dans la littérature pour une problématique de tatouage. Par exemple, dès 98, Wang et al proposent d'utiliser les codes de Hamming [249], Perreira et al. vont tester les codes BCH dans le contexte du tatouage [178], il existe des version hybrides en intégrant deux familles de codes comme les codes BCH et une répétition de la marque [22], et enfin certains travaux proposent d'utiliser les codes convolutionnels [244, 207]. On peut trouver aussi des comparaisons entre les différentes familles Reed Solomon, BCH et répétition [238]. L'originalité de notre approche a été de mettre en évidence l'adéquation de certains codes face à des attaques précises, et aussi d'intégrer ces codes dans un schéma de tatouage complet (jusqu'à son déploiement logiciel). Dans ce paragraphe, nous allons présenter la stratégie

mise en place notamment durant la thèse de A. Wadood et nous montrerons, à travers les résultats obtenus durant cette thèse, l'apport de ces différentes familles de codes et donc la potentialité de cet axe de recherche.

Bien sûr, comme nous le verrons, les caractéristiques de l'image (et donc du canal), comme la couleur ou l'aspect texturé, déterminent les effets des attaques sur l'image marquée. De même, l'algorithme de tatouage utilisé et le type de code correcteur vont jouer un rôle dans le compromis de l'invisibilité et de la robustesse. Dans l'étude faite, nous avons analysé les attaques suivantes : compression JPEG, bruit gaussien additif, filtrage passe-bas, modification de la teinte, saturation et luminosité.

Nous constatons que nous n'avons pas repris la partie s'intéressant aux attaques géométriques, puisque l'influence des codes est moindre face à ces attaques. En effet, l'élément important face aux attaques géométriques est la stratégie de resynchronisation comme nous l'avons décrit dans la section précédente.

En ce qui concerne l'algorithme de tatouage, nous utilisons celui introduit en début de chapitre qui insère la marque dans le domaine des ondelettes. Nous utilisons la version standard (à savoir basée sur une décomposition décimée) puisque nous souhaitons mettre en évidence uniquement l'apport des codes.

Avant d'analyser les résultats, nous allons revoir brièvement quelques notions sur les codes correcteurs qui nous permettront d'analyser les résultats en terme de robustesse, ainsi que la procédure proposée permettant d'introduire cette dimension "code" dans notre algorithme de tatouage.

5.4.1 Construction de la marque à partir de la signature : introduction des codes

Principe. Par la suite les codes correcteurs se résumeront sous la forme (n, k, d) avec n la longueur du code, donc le nombre total de bits dans le block encodé, k la dimension donc le nombre de bits de données que l'on peut encoder et d la distance de Hamming entre chaque paire du dictionnaire. Cette distance de Hamming définit la capacité à corriger du code à corriger les erreurs : un code (n, k, d) est capable de corriger t erreurs avec $t < \frac{d}{2}$

Le canal associé au tatouage est très bruité du fait des différentes attaques. L'objectif des codes correcteurs va être d'introduire de la redondance afin de protéger le signal (la signature) contre les dégradations dues au canal. Reprenant le schéma original (à savoir duplication de chaque bit puis répétition du motif), la robustesse du schéma de tatouage par code peut être améliorée par une redondance du motif issu du bloc du code correcteur.

Nous avons décidé de maintenir cette étape de duplication du motif car nous avons observé que les codes correcteurs ne pouvaient pas atteindre leur potentiel lorsque le taux d'erreurs dues au canal dépassait une certaine valeur critique. Plus précisément, avec un canal très destructeur, un code par exemple de type BCH n'apportait aucun avantage alors qu'une simple duplication de la signature apportait toujours une modeste protection. En revanche, l'utilisation du code BCH suivi d'une duplication du motif augmente les performances de la chaîne de décodage même lorsque le canal est très perturbé [238]. Nous profitons donc de la complémentarité des deux systèmes. La figure 5.34 résume le processus utilisé qui ressemble au schéma d'origine, avec simplement le remplacement de la répétition par un processus plus général associé aux codes appelé "Redondance".

Nous utilisons donc un modèle concaténé connu depuis bien longtemps [91]. La concaténation de codes permet d'augmenter la puissance des systèmes de codage au prix d'une augmentation de la complexité globale. La concaténation peut se faire de trois façons : parallèle, série ou hybride (parallèle et série) et sur deux ou plusieurs niveaux. Dans le cas d'une structure série à deux codes, comme nous le proposons, l'information est codée deux fois. Une première fois par le premier code appelé code externe, puis une seconde fois

par le deuxième, dit code interne. Les deux codes utilisés sont en général complémentaires. Dans notre cas, le code interne est la duplication de la marque qui a pour objectif de réduire le niveau d'erreur du canal et le code externe est code correcteur choisi (répétition, Hamming ou BCH) qui va permettre de réduire les erreurs sur la signature et ainsi faciliter la prise de décision finale.

Notons que les codes Reed-Solomon sont utilisés seuls car ils ont pour objectif de corriger des erreurs "burst" à savoir les erreurs par plages. Le décodage des Reed-Solomon sera fait avec l'utilisation d'un algorithme de décodage en liste [233, 98]. Le décodage en liste va permettre d'obtenir de meilleures performances que les algorithmes classiques ("bounded distance algorithms") quand le taux d'erreur est important. Ce dernier point illustre pleinement la richesse de la collaboration avec P. Gaborit puisque cela permet d'intégrer des approches plus fines dans l'utilisation des codes correcteurs.

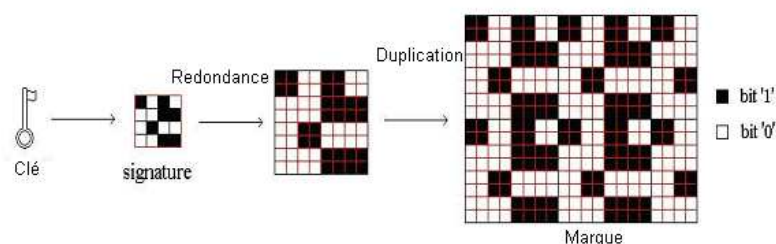


FIGURE 5.34 – Construction de la marque

Nous proposons de revoir très brièvement les éléments de bases des codes correcteur.

Codes correcteur par répétition de la marque. La robustesse du schéma de tatouage peut être augmentée avec une répétition de la marque comme codage interne suivie d'une duplication, d'un code Hamming ou d'un code BCH codes comme codage externe. Ce codage interne est adaptatif et dépend bien sûr de la taille de l'image. A la réception, la procédure inverse est appliquée, à savoir que l'on reconstruit tout d'abord la marque à partir des différentes répétitions puis la signature à partir du décodage répétition, Hamming ou BCH. Comme indiqué, cette duplication de la marque permet de lutter contre des taux d'erreur très importants.

Code par duplication. Le code par duplication construit la marque à partir de la signature, par une simple duplication de chaque bit de la signature. Si nous reprenons la schématisation de notre code sous la forme (n, k, d) , dans le cadre de la duplication nous avons toujours $(n, 1, n)$ avec chaque mot du code répété n fois. Le décodage est effectué à travers une étape de calcul de moyenne calculé sur le code reçu. Le code par duplication est utilisé comme codage externe et la répétition est utilisée comme codage interne. Cette chaîne est le schéma proposé à l'origine de cet axe mais indépendamment de la notion de codes.

Code de Hamming. Les codes de Hamming sont des codes en block linéaire, c'est-à-dire que l'on fait correspondre à chaque mot d'information un mot de code par une fonction linéaire. La représentation des codes binaires de Hamming sont de la forme $(2^m - 1, 2^m - 1 - m, m)$. Pour $m = 3$, nous avons $(7, 4, 3)$, à savoir que le code de Hamming encode 4 bits de données dans un block de 7 bits. Ces trois bits supplémentaires sont les bits de parité, et chacun de ces bits de parité code la parité de 3 des 4 bits de données, à chaque fois différent. De ce fait, le code de Hamming $(7, 4, 3)$ peut corriger 1 erreur par mot code.

Lors de la détection d'erreur, nous allons utiliser ce que l'on appelle la matrice de parité. La matrice de parité pour un code linéaire (n, k, d) est une matrice $(n - k) \times n$ H pour laquelle $Hx = 0$ pour tout x élément du code.

Si nous reprenons notre code de Hamming paramétré de la manière suivante $(7, 4, 3)$, quand nous multiplions le code reçu avec la matrice de parité, nous obtenons une parité d'une valeur comprise entre 000 et 111. Ces trois bits indiquent la localisation de l'erreur. 000 indique aucune erreur, et une valeur entre 001 et 111 indique la localisation de l'erreur dans les sept bits du mot du code reçu. Il suffit alors de faire pivoter la valeur du bit ainsi détectée comme erronée. C'est pourquoi tout mot avec au maximum une erreur sera décodé parfaitement et la taille de la redondance sera la plus faible possible.

Bose Chaudhuri Hocquenghem (BCH) codes Les codes BCH sont des codes cycliques c'est-à-dire ils ont la propriété d'être stables par permutation circulaire des mots : soit C un code block cyclique sur un champ fini F de blocks de longueur n . C est appelé code cyclique si pour tout mot du code $c = (c_1, \dots, c_n)$ de C , le mot $(c_n, c_1, \dots, c_{n-1})$ in F^n obtenu par décalage circulaire à droite de chaque composant est lui aussi un mot du code C .

L'objectif associé aux codes BCH est d'obtenir des codes de capacité de correction donnée (on se donne un nombre d'erreurs maximal et on en déduit le code). Les codes BCH sont tels que pour tous entiers positifs $m \geq 3$ et t avec $t \leq 2^{m-1} - 1$, il existe un code BCH de longueur $n = 2^m - 1$ qui est capable de corriger t erreurs et a comme dimension $k = n - m * t$. Dans le cadre de notre application de tatouage nous avons testé BCH(15, 7, 5) et BCH(63, 16, 23). Notons que le code BCH (63, 16, 23) est directement en lien avec notre algorithme puisque nous avons testé des signatures de taille 8×8 .

Codes de Reed-Solomon Les codes de Reed-Solomon [193, 217] sont des codes cycliques qui ont la particularité d'être non-binaires, c'est-à-dire qu'ils s'expriment à travers des symboles codés sur une séquence de m bits, où m est un entier de valeur supérieure à 2. Le code Reed-Solomon (n, k, d) consiste en des symboles de m bits pour tout n, k et d vérifiant $0 < k < n < 2^m + 2$ avec k le nombre de symboles de données que l'on peut encoder et n le nombre total de symboles dans le block encodé. Si nous reprenons la modélisation précédemment introduite

$$(n, k, d) = (2^m - 1, 2^m - 1 - 2t, 2t + 1)$$

nous constatons que les codes de Reed-Solomon sont capables de corriger toute combinaison maximale de t erreurs avec

$$t = \lfloor (d - 1)/2 \rfloor \quad (5.23)$$

La procédure de décodage utilise les algorithmes classiques "Bounded distance decoding algorithm" [116].

L'intérêt fondamental des codes Reed-Solomon est dans leur capacité à corriger simultanément plusieurs bits erronés en corrigeant un symbole d'un message. D'après ce qu'il est mentionné ci-dessus, un code RS corrige tous les bits erronés situés dans au plus t symboles du message. De ce fait, les codes RS sont particulièrement performants pour les erreurs par paquets comme nous l'illustrons sur la figure 5.35.

Dans cet exemple nous considérons un code RS $(n, k, d) = (40, 11, 30)$, chaque symbole est codé sur $m = 6$ bits. Puisque $d = 30$, ce code peut corriger $t = 14$ erreurs de symbole sur un bloc de 40. Considérons une dégradation par paquet qui touche 60 bits et donc dégrade 10 symboles comme illustré sur la figure

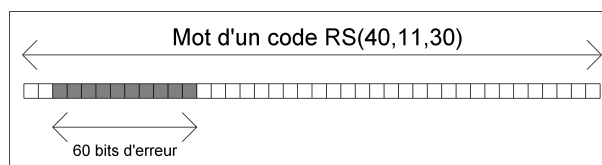


FIGURE 5.35 – Performance d'un code RS face aux erreurs par paquets

5.35. Le code Reed-Solomon (40, 11, 30) peut alors corriger l'ensemble des 14 symboles. Le code corrige par bloc de 6 bits et remplace l'ensemble du symbole corrompu par le symbole correct, indépendamment du nombre de bits dégradés dans le symbole. Ceci explique pourquoi le code RS est bien adapté pour les bruits de ce type par rapport au code binaire. Dans notre exemple, si les 60 bits sont perturbés aléatoirement sur la totalité plutôt que par paquets, cela peut affecter plus que 14 symboles et alors dépasser la capacité de correction du code. Il est bien connu que ce codage est très mauvais en cas de bruit impulsif de longue durée, ou de bruit aléatoire régulier.

Cette brève présentation des codes testés explique pourquoi nous avons introduit plusieurs codes dans notre schéma de tatouage car les attaques vont se traduire de différentes manières, soit aléatoirement sur la totalité des bits soit plutôt par paquets.

Comme nous l'avons dit pour le code RS, les algorithmes classiques de type "bounded distance decoding" corrigent jusqu'à $t = \lfloor (n - k)/2 \rfloor$ erreurs de symboles. L'évolution que nous avons introduite dans notre schéma de tatouage grâce à la collaboration avec P. Gaborit consiste en l'utilisation des décodages en listes. Les travaux de Sudan [233] et plus tard de Guruswami-Sudan [98] montrent que la capacité de décodage des Reed Solomon peut être augmentée jusqu'à $t_S = n - \sqrt{2kn}$ et $t_{GS} = n - \sqrt{nk}$ respectivement. Nous proposons de discuter de cette évolution dans le paragraphe suivant.

Décodage en listes L'idée des décodages en listes est que pour $t > (d - 1)/2$ l'algorithme de décodage va proposer une liste de mots plutôt qu'un unique mot. Le décodage en liste a été introduit à l'origine par Elias en 1957 [78], mais le premier algorithme utilisable pour une famille de codes, les Reed Solomon a été proposé par Sudan [233], et ensuite amélioré par Guruswami et Sudan en 1999 [98].

En contre partie, les méthodes de décodage en listes bien que très performantes souffrent d'une limite importante : elles sont très lentes et très complexes par rapport aux algorithmes classiques de décodage. Dans des contextes de communication, cette lenteur du décodage est prohibitive. En revanche, dans un contexte comme le nôtre ou dans des schémas de suivi[216], le décodage est en bout de chaîne et donc une complexité accrue peut être un défaut acceptable.

Nous ne détaillerons pas le fonctionnement des deux algorithmes testés de décodage en liste car nous nous éloignons du cadre de ce mémoire. Toutefois, nous illustrons les performances des différents algorithmes de décodage en fonction du rendement $R = k/n$ sur la figure 5.36. Nous pouvons observer que les performances de l'algorithme de Guruswami-Sudan sont meilleures que celles de Sudan lors que le rendement $R = k/n$ est élevé. A l'inverse lorsque ce rendement est faible, les deux algorithmes ont des performances similaires. Enfin les performances de ces deux algorithmes sont nettement supérieures aux algorithmes classiques (Bounded Distance) lorsque le rendement est faible.

Dans le cadre de la thèse de W. Abdul nous avons sélectionné l'algorithme de Sudan, car étant donné le paramétrage de l'algorithme de tatouage, le gain apporté par la méthode de Guruswami-Sudan n'est pas significatif, mais est associé à un accroissement de la complexité de l'algorithme très important. Les

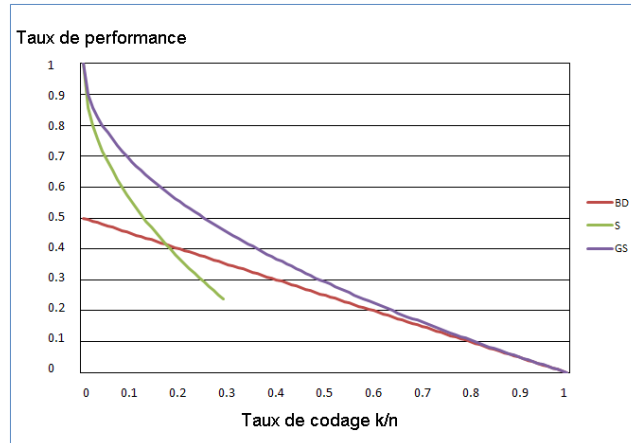


FIGURE 5.36 – Comparaison entre les performances de décodage pour un algorithme de type bounded distance (BD), décodage en liste de Sudan (S) et celui de Guruswami-Sudan (GS)

codes RS utilisés dans le cadre de notre tatouage sont paramétrés de la manière suivante : $RS(40, 11, 30)$, $RS(127, 9, 119)$ et $RS(448, 8, 441)$ qui correspondent à des taux de codage 0.275, 0.071 et 0.018 respectivement. Pour notamment les deux derniers nous sommes bien dans la zone avec un faible taux de codage et dans ce cas les performances des deux algorithmes de décodages par listes sont similaires (figure 5.36).

Nous proposons maintenant d'étudier les résultats les plus représentatifs obtenus avec ces codes correcteurs. Les mesures ont été faites durant la thèse de W. Abdul [1].

5.4.2 Mesure de la robustesse du tatouage avec les codes correcteurs

Nous avons conservé le protocole de mesures utilisé depuis le début du projet tatouage, à savoir que chaque image est marquée par 5 signatures différentes, la taille de la signature étant de 64 bits. Nous avons utilisé, comme base de mesures, les 15 images de différents types présentées à la fin de chapitre. Notons que nous n'avons pas repris les images FT utilisées au début de ce chapitre car elles présentent un contenu trop spécifique. Nous avons repris les attaques déjà exposées précédemment (hors attaques géométriques) à savoir compression JPEG, ajout de bruit, filtrage passe-bas, et modification de la teinte, de la saturation et de la clarté. Pour chaque attaque, nous présenterons le résultat sous la forme d'un graphique exprimant le taux d'erreur binaire (BER) en fonction du paramétrage de l'attaque :

$$BER = \frac{B_E}{B_T}$$

avec B_E le nombre de bits dégradés reçus et B_T le nombre total de bits de la signature.

Sur les graphes, nous ne représentons que la valeur moyenne du BER calculée pour les 15 images marquées et attaquées. Nous utilisons des forces de marquage légèrement plus faibles que dans la première partie du chapitre afin de bien mettre en évidence l'influence des codes.

5.4.2.1 Compression JPEG

Comme nous l'avons déjà introduit, l'attaque par compression JPEG touche majoritairement les hautes fréquences, donc pour des taux raisonnables de compression (et donc d'indice de qualité assez élevé) la

marque, insérée en bande de fréquence moyenne, va être peu touchée. Pour des taux plus élevés, les dégradations sont plus importantes et vont avoir des implications sur la marque. C'est pourquoi nous indiquons aussi le résultat du décodage obtenu sans protection pour bien différencier ce qui est lié à l'algorithme de tatouage et ce qui est lié au code.

Nous présentons sur la figure 5.37 les résultats des différents modes de protection sur les images de taille 1024 (nous sélectionnons cette taille suffisamment grande pour permettre de comparer les différents paramétrages des codes.) Nous retrouvons le résultat constaté durant la première partie de ce chapitre, à savoir qu'au delà d'un certain seuil (autour de 50) notre tatouage est robuste à la compression JPEG. De plus, on observe que la solution par duplication obtient des résultats tout à fait satisfaisants et comparables aux codes binaires. D'autre part les expérimentations menées avec les codes RS indiquent qu'ils sont inadaptés à ce type d'attaque. En effet, l'erreur due aux dégradations de la compression va se répandre sur toute la marque d'une manière aléatoire.

Il faut cependant affiner ces remarques générales car, comme nous l'avons indiqué dans l'introduction, les résultats sont aussi fortement dépendants du contenu de l'image. Pour illustrer cela, nous proposons de nous focaliser sur la mesure de robustesse du code RS pour différentes images (figure 5.38). Nous constatons des différences très importantes de robustesse comme par exemple entre l'image "Ours" et l'image "Plantes" (figure 7.2). Ces différences s'expliquent très probablement par le contenu fréquentiel des images, à savoir que dans l'image de l'ours, une très grande quantité de l'information est de la texture haute-fréquence, fortement touchée par la compression, alors que le domaine fréquentiel de la marque, à savoir les moyennes fréquences, contient probablement peu d'information, ce qui entraîne une grande fragilité de l'enfouissement. Le résultat est très différent pour l'image de plantes qui contient de nombreuses ruptures franches. Etant donné le mode de correction des codes RS, ils ne peuvent pas contrer cette attaque sur un tel contenu.

Pour renforcer ces constations, nous considérons la même étude mais pour le simple code par duplication (figure 5.39). Nous confirmons tout d'abord que la duplication est plus efficace que les codes RS face à la compression étant donné le type d'erreur associé. Ceci étant, on constate les mêmes différences entre les images en termes de robustesse, ce qui confirme bien également la sensibilité des résultats au contenu de l'image.

Pour conclure, nous pouvons dire que dans le cadre des attaques par compression JPEG, la protection par code binaire est bien plus performante, avec une légère préférence allant vers les codes BCH.

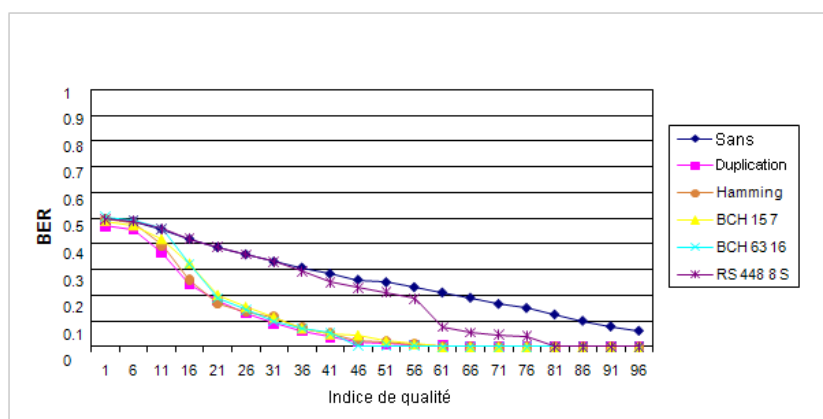


FIGURE 5.37 – Comparaison de la robustesse des différentes protections face à une attaque JPEG (image de taille 1024×1024)

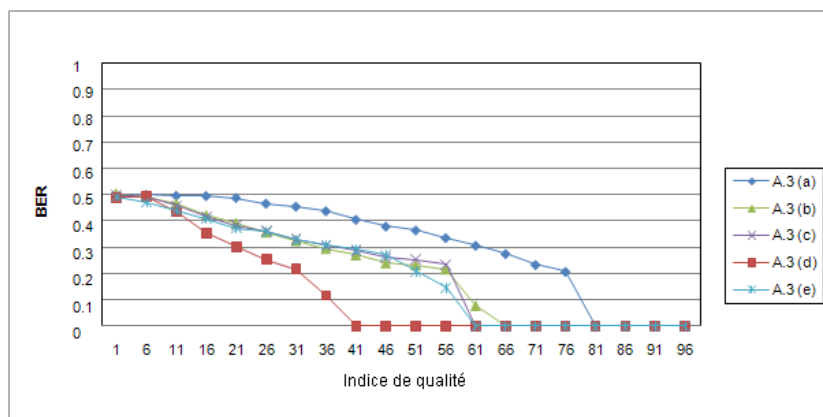


FIGURE 5.38 – Comparaison de différentes images face à l’attaque par compression pour le code RS.

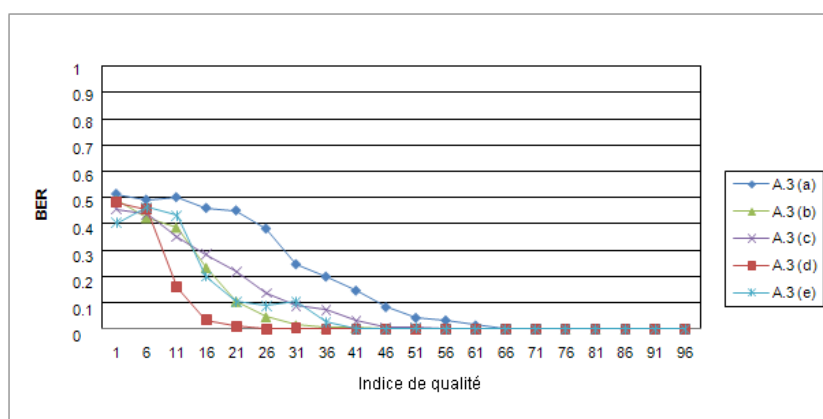


FIGURE 5.39 – Comparaison de différentes images face à l’attaque par compression pour le code par duplication.

5.4.2.2 Bruit blanc Gaussien additif

Nous reprenons l’attaque classique simulant l’ajout d’un bruit blanc Gaussien additif. Le paramètre dans ce cas va être le niveau de bruit. Dans le domaine des ondelettes, le bruit blanc gaussien va rester une perturbation gaussienne et va donc venir s’ajouter à chaque coefficient d’ondelettes. Nous sommes bien en présence d’une dégradation aléatoire sur chaque bit.

Nous indiquons sur la figure 5.40 le résultat de la protection des codes face à l’attaque par bruit. L’axe des abscisses correspond à la valeur du ratio signal sur bruit (en dB). Nous constatons sans surprise que les codes binaires permettent d’améliorer les résultats en terme de robustesse, et que le code RS qui comme nous le savons est efficace pour les erreurs par paquets, obtient de mauvais résultats.

5.4.2.3 Attaque par filtrage passe-bas

Nous analysons rapidement l’attaque par filtrage passe-bas, car celle-ci ne dégrade que très peu la marque, ou alors il faut utiliser un gabarit fréquentiel qui modifie la bande de fréquence associée à l’insertion et dans ce cas l’image attaquée est très fortement dégradée visuellement. Bien sûr, l’utilisation de codes ne modifie pas ce fait que nous avons déjà observé, comme nous pouvons le voir sur la figure 5.41.

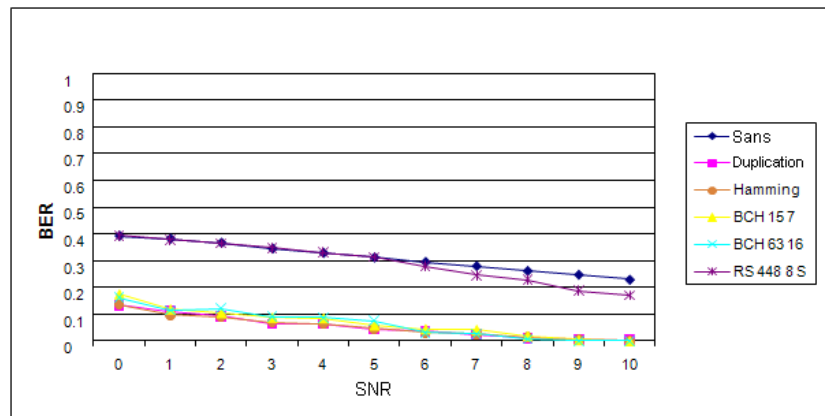


FIGURE 5.40 – Comparaison de la robustesse des différentes protections face à une attaque par bruit (image de taille 1024×1024)

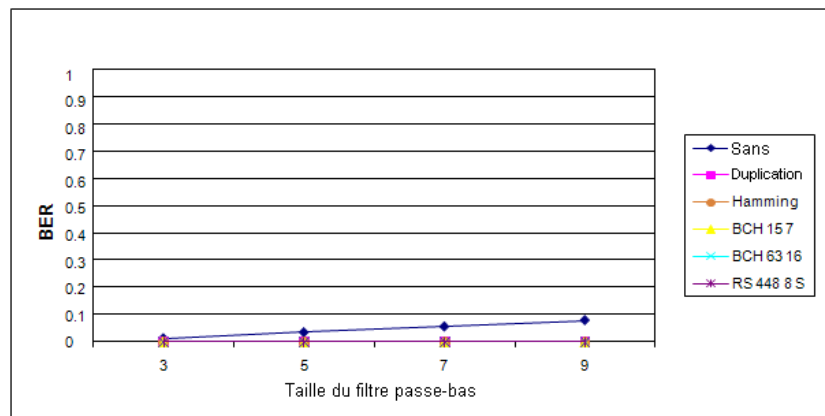


FIGURE 5.41 – Comparaison de la robustesse des différentes protections face à une attaque par filtrage passe-bas (image de taille 1024×1024)

Nous proposons maintenant d'étudier la robustesse des codes face à des attaques couleur, qui était l'un des objectifs du travail mené dans le cadre de la thèse de W. Abdul.

5.4.2.4 Attaque par modification de la teinte

La modification de la teinte va changer les couleurs présentes dans l'image par un mécanisme de rotation. Nous avons illustré en début de chapitre l'effet de cette attaque. La stratégie de tatouage que nous avons mise en place est très fragile face à cette attaque, puisque la modification des couleurs va perturber l'affectation des vecteurs couleur telle qu'elle a été faite lors de l'insertion.

En général, les différents codes donnent des résultats similaires, à savoir qu'au delà d'un certain angle de modification de la teinte, nous dépassons la capacité de correction et les résultats deviennent très mauvais. Les performances des codes RS sont légèrement meilleures car ils permettent une meilleure résistance pour des attaques de paramètres entre 10° et 30° comme nous pouvons le voir sur la figure 5.42. Cette résistance légèrement supérieure des codes RS s'explique par le fait qu'une modification des teintes de l'image va provoquer des changements par zone de l'image (zone de même couleur) et donc provoquer des erreurs par blocs. Comme nous l'avons introduit, ce type d'erreur est a priori mieux corrigé par les codes RS d'où les

résultats observés.

Bien sûr cette résistance a une limite puisque lorsque la modification de la teinte devient trop importante, l'ordre des vecteurs couleur change et dans ce cas, c'est l'algorithme de tatouage qui est entièrement mis en défaut. Dans ce cas, nous devons avoir un référentiel colorimétrique invariant.

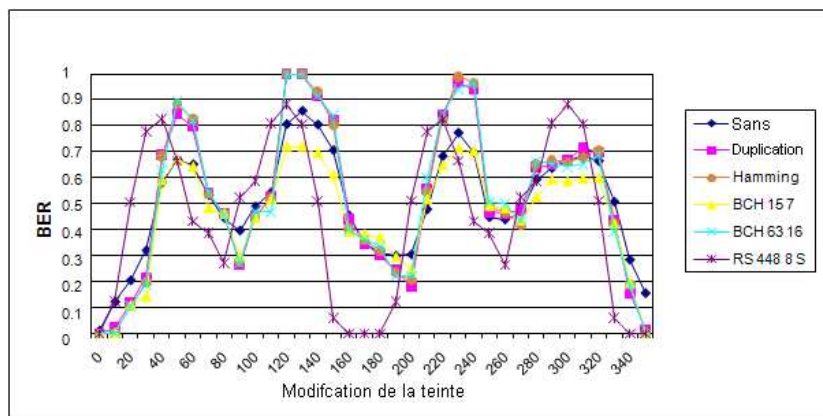


FIGURE 5.42 – Comparaison de la robustesse des différentes protections face à une attaque par modification de la teinte (image de taille 1024×1024)

5.4.2.5 Attaque par modification de la Saturation

Comme nous l'avons rappelé, la saturation mesure l'aspect coloré du pixel et correspond à la distance entre le pixel et l'axe des gris (figure 5.12). Nous rappelons que le paramètre de l'attaque de saturation va réduire ou augmenter cette distance en conservant toutefois la teinte ainsi que son intensité.

Lorsque le paramètre est négatif, l'image perd de son aspect coloré, et lorsqu'il est positif l'aspect coloré est renforcé, voir saturé.

Nous présentons sur l'image 5.43 la résistance des différents codes face à cette attaque. Comme nous l'avons vu en début de chapitre, l'attaque suivant un paramètre négatif ne pose aucune difficulté car nous sommes simplement en présence d'une contraction du nuage de points des pixels autour de l'axe des gris, mais aucune modification n'est faite. Seul le paramètre -100 est entièrement destructif puisque il rend grise l'image et donc c'est de nouveau l'algorithme de marquage qui est mis en défaut.

L'accroissement de l'aspect coloré est lui plus destructif et rapidement, il va mettre en défaut la grande majorité des codes. Ceci s'explique par le fait que la procédure de dilatation de la saturation va très rapidement provoquer des phénomènes de dépassement de capacité de représentation (nous arrivons à limite du cube RGB) et donc des erreurs. Toutefois, nous constatons de nouveau que les codes RS permettent d'améliorer très sensiblement les résultats. En effet, pour des paramétrages supérieurs à 30, les codes RS permettent de maintenir le taux $BER < 0.1$. Ceci s'explique de nouveau par le fait que les erreurs vont se produire par blocs (zone de l'image de même couleur). Bien sûr lorsque cette modification de la saturation devient trop importante le code ne peut plus corriger.

Ainsi que nous l'avons souligner précédemment, l'influence de cette attaque est dépendante du contenu initial de l'image. Nous proposons sur la figure 5.45 de détailler les résultats sur les images de taille 1024. Nous constatons que certaines images sont parfaitement corrigées par les codes RS alors que d'autres sont plus dégradées à partir d'un certain seuil.

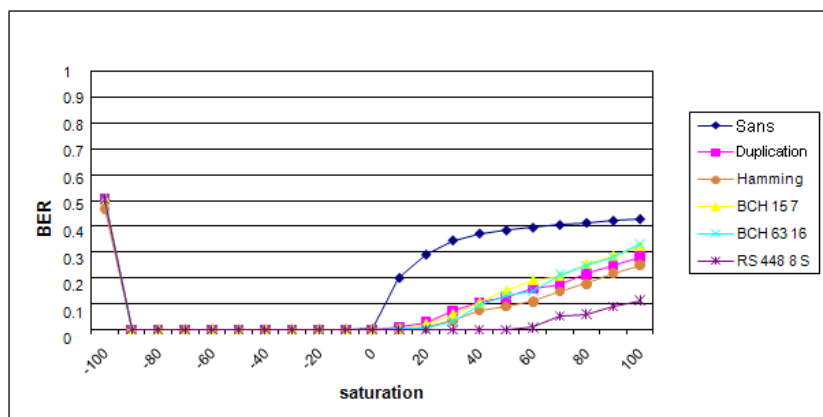


FIGURE 5.43 – Comparaison de la robustesse des différentes protections face à une attaque par modification de la saturation (image de taille 1024×1024)

Nous nous focalisons plus spécifiquement sur l'image "Feuilles" et l'image "Ours". Dans le cas de l'image "Feuilles" les codes RS sont mis en défaut car l'image d'origine est très fortement colorée avec une couleur dominante vers le rouge. Afin de confirmer l'impression visuelle, nous proposons une visualisation du nuage de points associé à l'image couleur (figure 5.44). Nous observons que ce nuage est entièrement décalé par rapport à l'axe des gris. Lorsque la saturation va être modifiée, le phénomène de dépassement va apparaître sur une très grande majorité de pixels rendant impossible la correction. A l'inverse, sur l'image Ours, nous observons que le nuage est plus proche de l'axe des gris, et dans ce cas, la modification de la saturation provoque moins de phénomènes de dépassement, et plus localisés dans certaines zones de l'image, provoquant un phénomène d'erreur par paquets. Les codes RS sont parfaitement adaptés pour corriger cette attaque.

5.4.2.6 Attaque par modification de l'intensité

Pour rappel, l'intensité correspond à la valeur de la projection du pixel sur l'axe des gris, et mesure ainsi une notion d'intensité, le paramétrage intervenant d'une manière multiplicative suivant l'axe formé par le pixel à modifier et l'origine. Lorsque l'attaque utilise un paramètre de -100 les pixels se confondent avec l'origine et l'image est entièrement noire, et un paramètre positif provoque un éloignement de l'origine et donc à partir d'une certaine valeur de nouveau un phénomène de dépassement du codage.

Cependant par rapport aux attaques précédentes, il est plus difficile de caractériser les dégradations. Etant donné "l'éclatement" des pixels, lors de l'augmentation de la clarté, les erreurs vont à la fois se disperser sur toute l'image, puisque se sont les points extrêmes du nuage qui arrivent à saturation, mais à partir d'un certain niveau, les erreurs vont apparaître aussi par zone.

Sur les images de taille 1024, nous constatons que les codes binaires obtiennent de meilleurs résultats que les codes RS (figure 5.46). De nouveau les résultats dépendent fortement du contenu comme nous l'illustrons sur la figure 5.49. Si nous étudions le nuage de points associés, par exemple, à l'image "Plant" avant et après attaque (figure 5.48), nous constatons le phénomène décrit précédemment, à savoir la dilatation du nuage de point, provoquant sur certains pixels un phénomène de dépassement.

Pourtant dans le cadre de cette attaque, et c'est la première fois, si nous étudions les résultats obtenus sur des tailles d'images différentes, par exemple des images de taille 512 comme sur la figure 5.47, le compor-

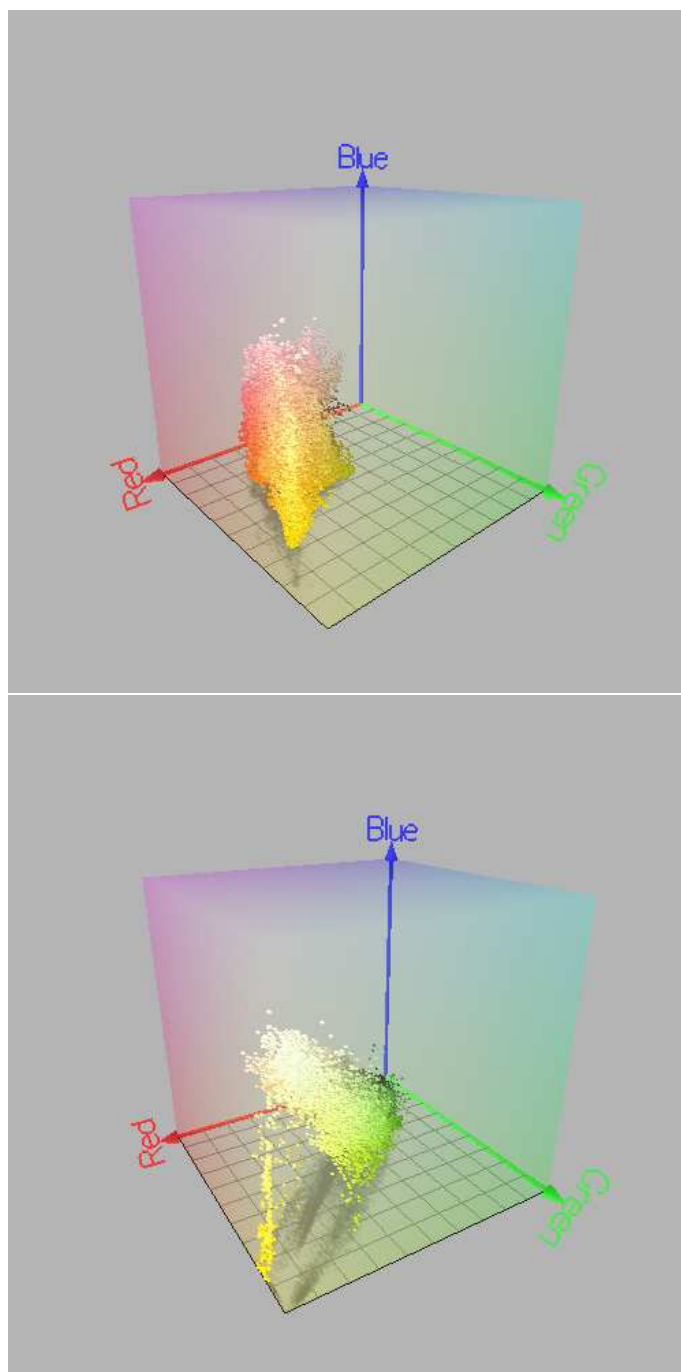


FIGURE 5.44 – Visualisation du nuage de points associé à l'image Feuilles et Ours respectivement.

tement des différents codes n'est pas aussi tranché, à savoir que toutes les familles de codes suivent un peu le même comportement jusqu'à un certain niveau. Ce qui différencie une image de taille 1024 d'une image de taille 512, c'est le paramétrage du code interne puisque forcément la duplication est plus réduite pour une image de plus petite taille. En fait, la réussite ou non est probablement plus liée au code interne à savoir la duplication de la signature. Pour mémoire, le code RS n'est pas associé à une duplication. Donc l'analyse de l'attaque par modification de l'intensité de l'image met plutôt en lumière la réussite d'un système de protection par concaténation.

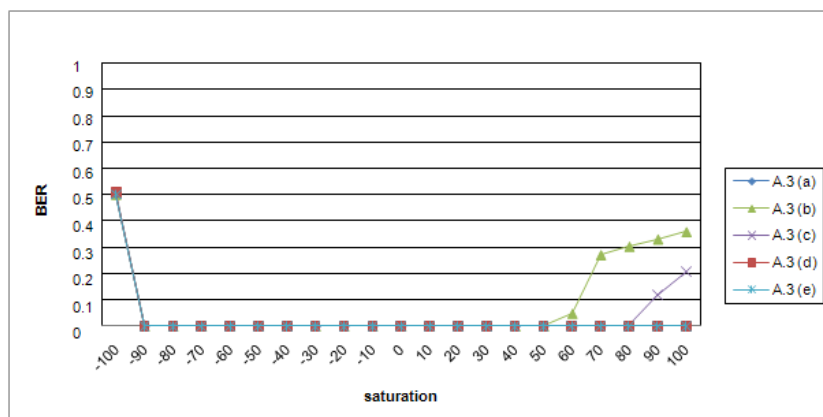


FIGURE 5.45 – Comparaison de la robustesse pour différentes images face à l’attaque par modification de la saturation pour le code RS.

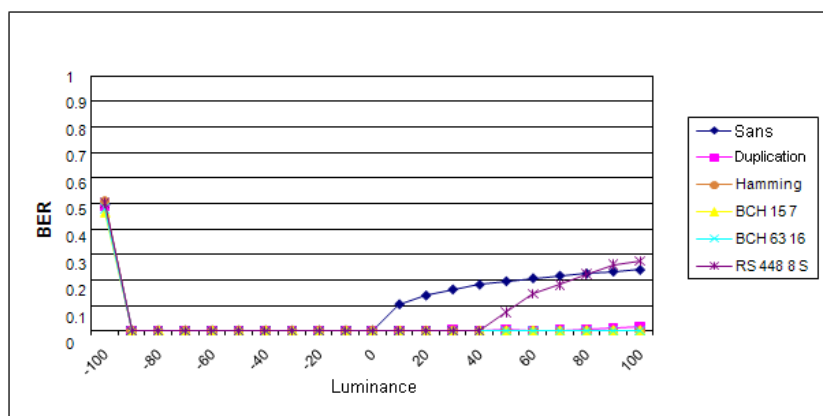


FIGURE 5.46 – Comparaison de la robustesse des différentes protections face à une attaque par modification de la Luminance (image de taille 1024×1024)

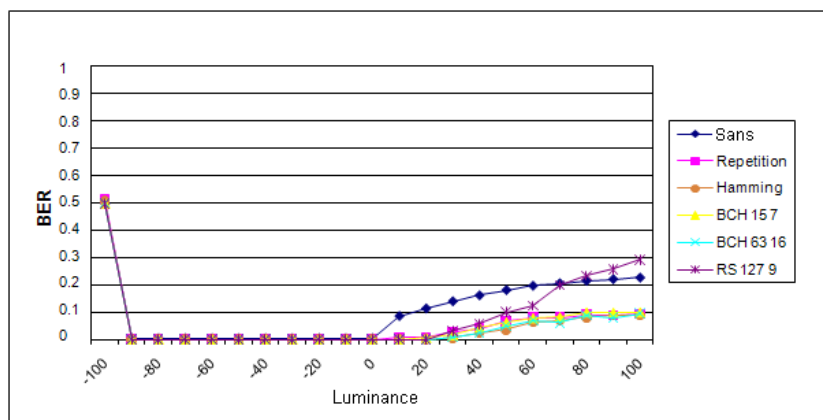


FIGURE 5.47 – Comparaison de la robustesse des différentes protections face à une attaque par modification de la Luminance (image de taille 512×512)

Toutes ces expérimentations ont montré l’importance d’utiliser différentes familles de codes, puisque chacune des stratégies va permettre de répondre à une certaine famille d’attaques. Ceci étant, nous ne pouvons pas émettre des conclusions et donc des spécifications très précises, car nous avons pu aussi observer

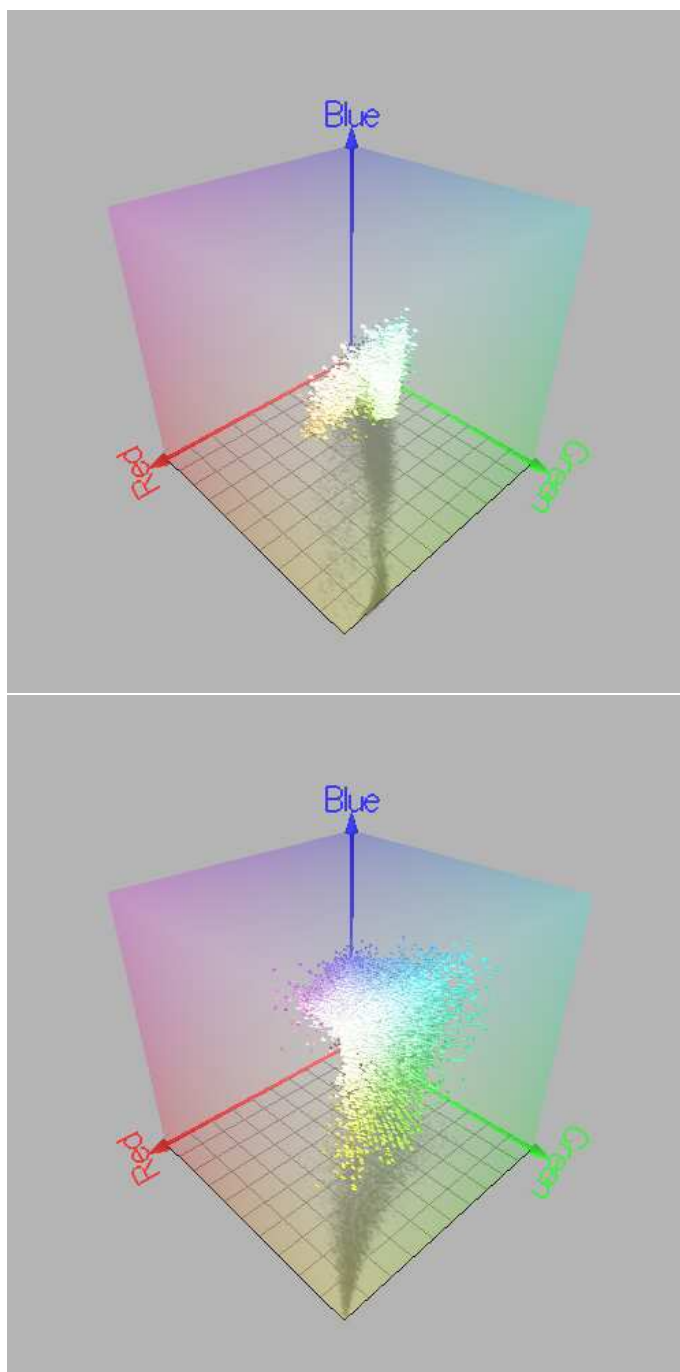


FIGURE 5.48 – Visualisation du nuage de points associé à l'image Plante avant et après la modification de luminosité suivant un paramétrage de 80.

que la qualité du décodage dépend également du contenu de l'image. Ce travail a posé les bases d'une solide collaboration scientifique avec le département DMI de XLIM que nous poursuivons actuellement à travers différents thèmes, certains sont dans la prolongation de ce travail à travers l'intégration de procédure de protection plus élaborée, d'autres concernent des domaines connexes comme le "Fingerprinting" ou la stéga-analyse.

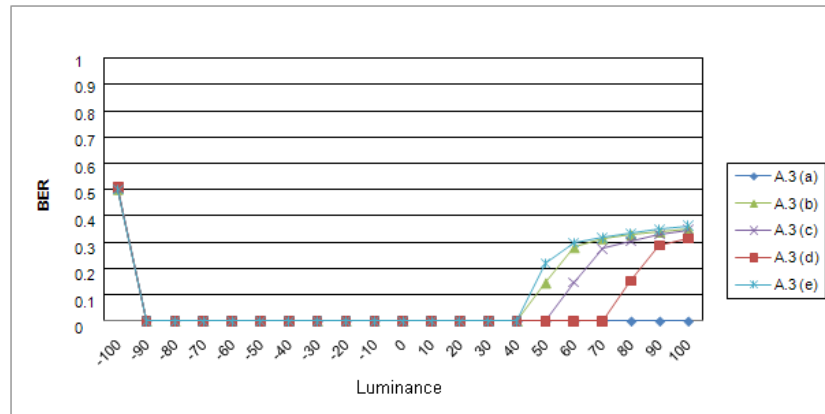


FIGURE 5.49 – Comparaison de la robustesse pour différentes images face à l'attaque par modification de la Luminance pour le code RS.

5.5 Conclusion

Les développements présentés dans ce chapitre ont permis de mettre en place une solution complète de tatouage. Les choix scientifiques se sont fait en trois étapes :

- Mise en place de l'algorithme basé ondelettes permettant d'intégrer la dimension couleur à travers une paramétrisation de la modification des vecteurs couleur. Cet algorithme résiste à de nombreuses attaques non désynchronisantes. Cette étape a été réalisée à travers le partenariat avec FT R& D et dans le cadre de la thèse de CIFRE A. Parisi [171].
- Intégration d'une transformée en ondelettes non-décimée afin de résister aux attaques provoquant une désynchronisation par translation et ajout d'un détecteur de points d'intérêt afin de contrer les attaques géométriques globales de type changement d'échelle et rotation. De nouveau cette étape s'est faite en collaboration et a donné lieu au déploiement d'un démonstrateur logiciel.
- Intégration lors de la création de la marque d'un code correcteur afin de renforcer la résistance du schéma de tatouage. Ce travail s'est fait en co-encadrement avec P. Gaborit du département DMI-XLIM dans le cadre de la thèse de A. Wadood.

A travers ces différents jalons, le schéma de tatouage est arrivé à un certain stade de maturité qu'il me semble difficile de dépasser significativement. Afin de faire évoluer la méthode, il faut maintenant revenir au schéma original et modifier entre autre l'espace d'insertion, à l'aide notamment des travaux présentés précédemment sur les nouvelles transformées. Dans le cadre de la thèse de A. Wadood, nous avons proposé un nouveau schéma utilisant la transformée Contourlet afin d'être au plus proche de la séparation fréquentielle faite par le système visuel humain, afin d'intégrer au mieux la dimension psychovisuelle. Nous allons poursuivre dans cette direction, en intégrant l'apport des codes correcteurs.

Cet axe portant sur la protection des données multimédia s'est concrétisé à travers le co-encadrement de deux thèses [171, 1], s'est construit à travers la collaboration avec P. Gaborit de XLIM-SIC, H. Laurent de OrangesLabs. Il a été valorisé dans la littérature à travers principalement [172, 2, 4, 5, 3], et dans différents brevets.

Les perspectives de cet axe portent aussi sur une nouvelle dimension applicative que nous décrivons maintenant.

5.6 Perspective : la stéga-analyse par métriques de qualité

Nous proposons de terminer ce chapitre à travers une brève description du contenu scientifique que je développe actuellement dans la continuité des travaux sur la sécurité et ayant comme sujet la stéga-analyse, notamment en collaboration avec une collègue de XLIM-SIC AS Capelle-Laizé.

Principe L'objectif d'une méthode de stéga-analyse est de tenter de détecter la présence d'un message caché dans un document primaire. L'un des concepts élémentaires de sécurité indique que plus la distribution de probabilité du message primaire est ressemblante à la distribution du message caché et plus il sera difficile de détecter ce message caché. Les méthodes de stéga-analyse vont chercher à détecter la rupture du modèle. Actuellement il n'existe aucune méthode pratique de stéganographie qui réussisse cette insertion parfaite. C'est pourquoi différentes stratégies sont proposées ayant pour but de détecter la présence du message.

Le domaine de la stéga-analyse est très vaste, avec de nombreuses méthodes proposant chacune des variantes par rapport à d'autres algorithmes. Afin de structurer la description de ce nouvel axe, nous allons tout d'abord lister les paramètres différenciant globalement les algorithmes entre eux avant de citer plus spécifiquement certaines familles sur lesquelles nous souhaitons nous pencher.

Tout d'abord il existe différentes types de stéga-analyses :

- Le pirate passif intercepte le message primaire et teste la présence d'un message caché.
- Le pirate actif peut modifier le média et ainsi perturber le message
- Le pirate malicieux cherche à modifier le message.

Dans notre travail, nous nous plaçons principalement dans le premier cas. Donc l'élément fondamental de notre algorithme de stéga-analyse est cette étape de détection. Finalement, dans de nombreux algorithmes l'approche de stéga-analyse consiste en un problème de classification à deux classes : à savoir que l'algorithme reçoit des données et classifie ce message comme une donnée perturbée ou non.

La variante principale qui peut exister au niveau du classifieur est l'information accessible à la méthode. Dans un premier cas, il n'existe aucune information sur les données que l'on reçoit mais simplement une suspicion de camouflage. Alors, doit être développée une méthode de stéga-analyse qui doit être capable de détecter toutes les formes de stéganographie. Bien sûr, ce cas appelé "blind steganalysis" est le cas le plus difficile. A l'opposé, si nous connaissons l'algorithme d'insertion qui a été potentiellement utilisé, nous développons une méthode de stéga-analyse dédiée à cette algorithme, ce sont les méthodes dites de stéga-analyse ciblées. Nous nous focalisons sur les méthodes "blind stega-analyse".

Maintenant, décrivons plus précisément les éléments que nous devons étudier dans le cadre des méthodes de stéga-analyse. Un schéma est actuellement assez figé, à savoir que la grande majorité des méthodes vont extraire tout d'abord des caractéristiques de l'image, ensuite elles définissent un discriminateur. Puis elles paramètrent ce discriminateur par une étape d'apprentissage à partir d'un ensemble d'images [10, 135]. Dans certains cas, des opérations de traitement d'image sont appliquées avant l'extraction des mesures [84, 101, 51]. Enfin, quelquefois un traitement est appliqué sur l'ensemble des mesures afin d'augmenter le caractère discriminatoire de ces mesures [111, 214].

A noter, pour conclure sur cette définition générale de notre objet d'étude qu'actuellement il n'existe pas une méthode universelle de stéga-analyse.

Pour résumer, les méthodes que nous proposons de développer, nous présenterons l'enchaînement d'opérations suivant :

1. prétraitement de l'image : ensemble d'opérations appliquées à l'image avant l'extraction des mesures. Cela peut être des opérations de traitement d'images classiques, mais aussi comme nous le verrons un changement d'espace de représentation (par exemple vers le domaine des ondelettes) avec pour objectif une amélioration de la détection ;
2. extraction des mesures : l'objectif est de sélectionner des caractéristiques sensible aux changements, toutefois la dimension du vecteur de mesures doit rester raisonnable afin de limiter le coût de calcul ;
3. définition et paramétrisation du discriminateur : on doit définir un discriminateur et ensuite dans une grande majorité des cas effectuer une procédure d'apprentissage à partir d'une large base de données d'images ;
4. classification : c'est l'étape centrale qui à partir du discriminateur permet de décider si une image est corrompue ou d'origine.

Nous proposons maintenant de citer les caractéristiques que nous abordons déjà ou que nous aborderons dans ce nouvel axe. Bien sûr cette liste ne peut pas être exhaustive.

Métrique de Qualité pour la stéga-analyse. Etant donné les compétences développées au sein du département XLIM-SIC sur les métriques de qualité (dans le cadre de la compression mais aussi du tatouage), l'une des premières approches que nous abordons est la détection de marques à partir de mesures de qualité d'images. En effet, il existe un tout petit nombre de travaux dans la littérature qui proposent de détourner les métriques de qualité d'image, à l'origine définies pour évaluer le résultat perceptuel d'un traitement pour la stéganalyse. On peut citer Avcibas et al. [11] qui font une analyse exhaustive des métriques avec références afin de déterminer lesquelles sont particulièrement adaptées pour une attaque mais aussi proposent d'en combiner un certain nombre puisque il existe différentes familles de métriques abordant l'image d'une manière complémentaire (différence de pixels mesures de corrélation, détection de contours, analyse spectrale ...). Toutes ces approches ont pour vocations de modéliser le comportement du système visuel humain. Il faut noter que la prise en compte de la couleur est un élément encore ouvert et pourtant fondamental il nous semble pour une possible détection. C'est l'un de nos points de fixation. Ces premières études ont montré que face à des algorithmes classiques tels que Digimarc, Coxs et PGS steganography, Steganos, S-tools et Jsteg, les métriques de qualité obtiennent des résultats très satisfaisants. Le problème reste bien sûr le choix des métriques qui sont actuellement en très grand nombre dans la littérature (secteur en plein essor mais essentiellement dans des problématiques de compression). Il me semble tout à fait logique d'introduire cette stratégie dans un cadre de stéga-analyse, car lorsque l'on étudie les algorithmes actuels, un grand nombre d'entre eux, au delà de la modélisation statistique faite en amont, se traduit numériquement par des opérations de différences entre pixels voisins, de mesures de corrélation, ... Opérations que l'on retrouve dans une métrique de qualité. C'est pourquoi dans cet axe de recherche, je collabore avec ma collègue C. Fernandez-Maloigne sur la pertinence des différentes métriques.

Moments d'ordre supérieur dans un espace transformée. Ensuite, le second point qui nous semble fondamental à analyser concerne les moments statistiques des coefficients dans un domaine transformée de type ondelettes ou DCT. Farid [84] est le premier à proposer l'utilisation des ondelettes pour la stéga-analyse. Cette approche se retrouve ensuite dans de nombreuses méthodes proposées dans la littérature. Si l'on veut résumer cette classe d'algorithme, pour chaque bande d'ondelettes des paramètres statistiques comme la moyenne, la variance le skewness et le kurtosis sont calculés et constitueront le vecteur de mesures.

Les évolutions que l'on va rencontrer vont être essentiellement dans la mesure qui est faite à partir de ces coefficients, mais aussi dans le choix des fonctions analysantes utilisées. On peut remarquer que, de nouveau, l'intégration de la couleur reste un problème particulièrement absent. Tous les travaux développés au département XLIM-SIC, autour des transformées, ont pour objectif de capter au mieux l'information de l'image (Curvelet, ondelettes complexes), travaux que j'ai présentés dans ce mémoire. Il me semble pertinent d'adapter ou d'analyser ces nouvelles décompositions plus performantes dans un cadre de stéga-analyse.

Notons enfin que la difficulté liée à ces méthodes est la sensibilité à la base d'images utilisée lors de l'apprentissage. En effet, aucun *a priori* n'est fait et l'hypothèse sous-jacente est que la distribution des coefficients des images d'origine peut se caractériser à partir d'un nombre restreint de paramètres.

Analyse des liens spatiaux. Une autre famille de méthodes se propose d'analyser les relations spatiales existantes entre les pixels. Ceci peut se faire de nombreuses manières, l'objectif étant de détecter une rupture dans la modélisation attendue. Par exemple, on peut calculer les matrices de transition suivant une approche markovienne [234], ou d'autres méthodes utilisent la notion de matrice de co-occurrence comme [51]. Cette notion de transition peut aussi être exprimée dans le domaine DCT. La notion de lien entre pixels peut aussi être analysée à travers des approches de type filtrage comme le calcul d'un gradient exprimant la différence entre deux pixels successifs [101]. Cette classe de solution nous semble de nouveau fondamentale à incorporer dans notre étude, d'autant plus que de nouveau on peut remarquer que la dimension couleur est absente.

Problème du discriminateur. Le discriminateur utilisé n'est pas propre au problème de la stéga-analyse et les solutions choisies sont issues des recherches menées dans la problématique plus générale de la reconnaissance de formes. Dans le cadre de la stéga-analyse on va trouver des classifications par régressions multiples, par réseaux de neurones, par support vecteur-machine, classifieur de Bayes,... L'un des éléments clés est l'évaluation de l'impact du choix du discriminateur dans un problème de stéga-analyse. Ce point là fait l'objet d'une attention particulière, à savoir pour un jeu de mesures déterminées, évaluation de plusieurs discriminateurs.

La compétence scientifique complémentaire sur ce sujet est apportée par ma collègue A.S. Capelle-Laizé, notamment à travers une réflexion sur des discriminateurs à base de logique floue.

Premiers résultats. Cet axe de recherches étant très récent, il ne serait pas raisonnable de conclure déjà sur la pertinence des pistes citées ci-dessus. D'autant plus qu'il me semble que ce travail doit être mené sur une longue période et à une échelle plus large (par exemple dans le cadre d'un groupe de travail du GDR ISIS) pour être confronté aux approches plus classiques.

Toutefois, de nos premiers travaux, nous avons pu constater que certaines mesures de dégradations avec référence (donc qui nécessitent l'image d'origine) permettent de détecter des images marquées. Plus précisément dans une population composée :

- d'images non modifiées ;
- d'images filtrées passe-bas ;
- d'images bruitées ;
- d'images compressées ;
- d'images stéganographiées par Stools ;
- d'images stéganographiées par StegoMagic ;

- d’images stéganographiées par The Third Eye ;
- et enfin d’images stéganographiées par wbStego.

Un discriminateur de type k plus proches voisins détecte les images stéganographiées, ceci à l’aide de mesures de dégradation avec références. Nous retrouvons en cela les résultats de [10]. En revanche, les premiers développements autour des métriques sans référence (l’image avant transformation n’est plus disponible) n’ont pas encore obtenu des résultats satisfaisants.

Ce travail qui nous semble novateur au niveau de la communauté va donc se poursuivre avec cet aspect fédérateur de compétences transverses.

Chapitre 6

Conclusion et perspectives

Ce qui constitue l'axe centrale de mes recherches et directions de recherche et qui a été décrit dans ce mémoire est la construction d'une représentation numérique multiéchelles permettant la description la plus fine possible des données définies sur une grille discrète à valeurs vectorielles. Différents outillages mathématiques et algorithmes ont été proposés et/ou testés, toujours dans le souci de donner un sens à l'espace de représentation. Le volet applicatif est fortement présent notamment à travers la réflexion autour de solutions pour la sécurité des données multimédia. Ces travaux ont été menés à travers l'encadrement de huit thèses, la gestion de nombreux projets contractuels et différents échanges avec des collègues de ma communauté et d'autres communautés scientifiques. Des choix ont bien sûr été faits en ce qui concerne les apports et axes présentés dans ce mémoire. Ils se sont structurés en quatre parties.

Le premier chapitre a posé les bases et les problématiques inhérentes à mon travail de recherche. Dans la recherche d'une représentation multiéchelle qui prend en compte l'aspect géométrique et la dimension vectorielle des données, nous avons dû intégrer des connaissances et des compétences d'autres domaines, notamment de l'informatique graphique. Dans un premier temps, à travers une collaboration avec E. Andrès de l'équipe IG de SIC, nous avons construit une nouvelle stratégie discrète pour la transformée Ridgelet/Curvelet. L'un des apports a été l'utilisation du formalisme de la géométrie discrète permettant ainsi le déploiement d'une technique rapide, parfaitement restructurable et qui a logiquement obtenu des résultats performants pour le débruitage de données. La définition apportée par la géométrie discrète nous a permis d'étendre ce travail à toute dimension en collaboration avec un jeune collègue D. Helbert[106]. Notre second apport a été l'introduction d'une transformée en ondelettes géométrique adaptative qui s'appuie sur un schéma lifting et une structure de graphe permettant l'estimation de la distance géodésique. Cela nous a permis d'intégrer la modélisation de la géométrie de l'image dans le calcul de la transformée. Les résultats, notamment dans le cadre de la restauration de données, montrent un comportement de la transformée prometteur qui illustre la manière fine avec laquelle est extraite l'information 2-D. Cet axe, complémentaire au premier s'est développé dans le cadre d'une collaboration industrielle avec France Telecom R&D et notamment à travers la thèse de G. Lebrun [132]. Ce travail s'est fait en co-encadrement avec S. Pateux de France Telecom.

En complément de ces travaux, la thèse de A. Maalouf [141] proposait une autre manière de modéliser la géométrie de l'image à travers les représentations Bandelettes et leurs évolutions. Ce travail a permis de mettre en commun les outils de type "Equations aux Dérivés Partielles" développés par mon collègue B. Augereau et les dernières générations d'ondelettes et a obtenu de très bons résultats applicatifs.

Cet axe a été valorisé dans la littérature à travers principalement ces publications [147, 142, 106, 41, 145,

149, 148, 146, 144, 143, 40, 43, 133].

Le second chapitre avait pour objectif de présenter les travaux menés dans le cadre de l'utilisation de formalismes algébriques pour analyser et traiter les images numériques couleur. J'ai créé cet axe de recherche dès mon recrutement au sein du département XLIM-SIC. Il a trouvé son implication scientifique dans de nombreux encadrements dont deux thèses (P. Denis et R. Soulard [62, 222]) et nous avons utilisé avec succès ces techniques dans la thèse de F. Petit [179], suite aux discussions menées avec ma collègue A.S. Capelle-Laizé.

Ce travail sur les algèbres géométriques a été un élément fédérateur dans la construction d'un thème scientifique commun au sein de la fédération PRIDES (Programme Régional de recherche en Images, Données et Systèmes). J'ai été co-animateur depuis sa création d'une action "Géométries et images", dans le cadre de cette fédération PRIDES, regroupant les laboratoires d'informatiques et de traitement des images des universités de Poitiers, de la Rochelle et de l'Ensm. L'objectif de cette fédération était la mise en commun des savoirs et l'impulsion de projets collaboratifs en traitement d'images au niveau de la région Poitou-Charentes. Cette action a notamment permis de développer de solides collaborations sur la problématique de l'utilisation des algèbres de Clifford pour l'analyse d'images. Le chapitre de livre rédigé en collaboration avec M. Berthier [37], directeur du Laboratoire Mathématiques, Image et Applications de LA Rochelle, en est une illustration, ou encore le dépôt d'une ANR dans le cadre de l'appel ASTRID. Comme nous l'avons vu, les quaternions, contrairement aux nombres complexes, peuvent décrire complètement une information couleur, par exemple sur leurs parties imaginaires. On peut alors manipuler un vecteur couleur par de simples opérations comme le produit. Nous avons alors proposé différentes ré-écritures d'opérations couleur à travers des transformations géométriques sur des vecteurs couleur au moyen de ce formalisme. Parmi les travaux menés au département XLIM-SIC, nous avons proposé une évolution du détecteur de ruptures couleur de Sangwine. Ensuite, dans ce chapitre, nous avons vu que les algèbres géométriques $\mathcal{G}_3$ permettaient de manipuler des entités géométriques telles que scalaires, vecteurs, bivecteurs. De nouveau, nous avons proposé différentes variantes d'opérations couleur en utilisant ce formalisme et construit une nouvelle approche spatiale de détection de ruptures couleur, en généralisant celle définie avec les quaternions. Enfin, dans la dernière partie, nous avons discuté de la question scientifique qui constituait le point de départ de mon travail sur les quaternions : la définition d'une transformée de Fourier pour les images couleur. Cependant, malgré les différentes évolutions proposées ou étudiées, nous n'avons pas pu prouver un apport numérique majeur par rapport à l'analyse chromatique par transformée Complexe. Les techniques apportent une manière élégante de décrire certains problèmes. Cependant, le bilan nous semble pour l'instant mitigé et, même à travers les évolutions les plus récentes, nous n'avons pas pu dépasser certaines limites dans l'analyse de l'image couleur. Cet axe a été valorisé dans la littérature à travers principalement ces publications [64, 37, 45, 63, 180, 181, 197].

La réflexion sur les outils mathématiques pour l'analyse d'images à valeurs vectorielles se traduit aussi par une collaboration avec la DGA. L'objectif est d'étudier les algèbres géométriques dans le domaine de l'analyse d'images hyperspectrales pour la détection d'anomalies et d'objets et pour la fusion d'images. Cette étude est menée dans la thèse de S. Rousseau [196] en co-encadrement avec D. Helbert et en lien avec J. Blanc-Talon de la DGA.

Au niveau national, j'ai assuré la co-animation avec D. Tschumperlé du Greyc une action du GDR ISIS ayant pour sujet les outils mathématiques pour l'image couleur, dans laquelle les algèbres de Clifford avait pleinement leur place.

Le troisième chapitre a décrit nos travaux les plus récents concernant la définition de décomposition en ondelettes adaptée à l'image. Cette approche est très séduisante et nous semble prometteuse car elle pose la réflexion à partir de fondements issus du signal et de la physique afin de fournir des outils pour caractériser et comprendre le signal 2-D. A travers nos travaux concernant la transformée QWT, nous avons montré que ces outils peuvent fournir des réponses applicatives grâce à la qualité de la représentation. Nous avons proposé des pistes à la fois dans des problématiques de caractérisation de textures et aussi de codage d'images. Ensuite, nous avons décrit une approche ambitieuse à travers la transformée en ondelettes monogéniques, annonçant les travaux actuels sur la définition d'une représentation couleur 2-D au plus proche du concept physique du signal analytique 1-D. Ce développement pose des questions fondamentales et permet, comme nous l'avons vu de faire une connexion inattendue avec les travaux présentés au tout début du mémoire sur les transformées Ridgelet. Les deux stratégies semblent à l'origine plutôt disjointes, l'une cherche une analyse directionnelle, l'autre reformule une notion fondamentale du signal, et cependant ces deux approches se rejoignent montrant finalement la complémentarité des axes de recherche développés au sein du département XLIM-SIC. Ce chapitre pose clairement les fondements des travaux à venir. Cet axe s'illustre à travers la thèse de R. Soulard [222] et il a été valorisé dans la littérature à travers principalement ces publications [228, 226, 227, 224, 225].

Comme indiqué, ces travaux ont par ailleurs trouvé une pleine application dans l'ANR CAIMAN (dont les partenaires sont Thales Communications, l' IETR (UMR 6164) de l'INSA de Rennes, , le laboratoire des Equipes Traitement des Images et du Signal (UMR 8051) de Cergy, qui se propose de travailler sur de nouveaux codeurs d'images. La transformée QWT obtient des résultats très encourageants en qualité de codage et a fait l'objet de tests menés par A.S. Baquet, post-doctorante au département SIC durant l'année 2011, co-encadrée par mes collègues C. Larabi, C. Perrine et moi-même.

Enfin le dernier chapitre a permis de mettre en lumière la création d'une chaîne applicative complète pour une solution de tatouage. Cette chaîne a nécessité l'intégration de différents blocs novateurs. Nous avons tout d'abord développé un algorithme "basé ondelettes" intégrant la dimension couleur à travers une paramétrisation de la modification des vecteurs couleur. Cette étape a été réalisée à travers le partenariat avec FT R& D et dans le cadre de la thèse de A. Parisi [171]. Ensuite, nous avons abordé le problème des attaques désynchronisantes par l'intégration d'une transformée en ondelettes non-décimée contre les attaques par translation, et l'ajout d'un détecteur de points d'intérêt afin de contrer les attaques géométriques globales de type changement d'échelle et rotation. De nouveau, cette étape a donné lieu au déploiement d'un démonstrateur logiciel. Enfin, nous avons introduit le concept de code correcteur lors de la création de la marque afin de renforcer la résistance du schéma de tatouage. Nous avons construit un lien entre les attaques possibles et les codes adaptés, ce qui nous semble une approche originale par rapport à l'existant.

Cet axe portant sur la protection des données multimédia s'est concrétisé à travers le co-encadrement de deux thèses [171, 1], et s'est construit à travers l'échange avec différents chercheurs P. Gaborit de XLIM-SIC, H. Laurent de OrangesLabs. Il a été valorisé dans la littérature à travers principalement ces publications [172, 2, 4, 5, 3], et dans différents brevets.

Ce travail se développe au sein du groupe de travail " Sécurisation logicielle de la transmission d'images", dont je suis co-animateur avec P. Gaborit (XLIM-DMI), dans le cadre de la mise en place du plan pluri-formation (PPF2008-2011) intitulé " Géométries-Images-Communications " incluant les universités de Poitiers, La Rochelle, Limoges, Tours, Orléans et l'Ensm. L'objectif de cette action est le développement de projets transverses faisant appel à différentes compétences complémentaires traitant de la notion de protec-

tion de données multimédia. J'ai enfin été co-organisateur de la journée "Nouvelles avancées en tatouage d'images" avec PUECH William et LE GUELVOUIT Gaëtan le 12 mars 2009 à Paris.

Je propose de présenter maintenant les différentes perspectives et axes annoncés dans le mémoire, axes que je souhaite développer dans les années à venir, toujours dans le même cadre général scientifique.

Intégration de la géométrie et de la couleur par des structures de graphes. La première perspective que je souhaite mettre en avant est la poursuite de la réflexion concernant la transformée en ondelettes géométrique adaptative s'appuyant sur une structure de graphe, ou sur une structure plus complexe. Cet axe reste d'actualité car il nous reste à approfondir l'intégration de la dimension couleur dans la transformée. Ce qui nous semble novateur, c'est que nous ne proposons pas une transformée en niveaux de gris que nous cherchons à étendre à des données vectorielles, mais nous définissons directement une transformée adaptée à des données vectorielles. Cependant, il nous faut approfondir la modélisation de la couleur et des disparités colorées. De plus, à travers les applications de débruitage, on constate que ce type de transformée peut aussi permettre de remettre en question la stratégie de régularisation, à savoir qu'au lieu d'un seuillage sur une notion d'amplitude de coefficients, il nous semble maintenant également judicieux d'utiliser la modélisation de la topologie traduite par notre graphe. Ce nouveau travail qui s'amorce requière lui aussi des compétences scientifiques complémentaires, apportées par les collègues de l'équipe IG de XLIM-SIC qui proposent des représentations topologiques et géométriques plus avancées.

Utilisation du concept du signal analytique pour la définition d'une transformée en ondelettes 2-D couleur Le sujet décrit dans le chapitre sur les algèbres de Clifford est un sujet d'étude très récent et de ce fait il ouvre de nombreuses questions. Les techniques apportent une manière élégante de décrire certains problèmes. Cependant, le bilan nous semble pour l'instant mitigé même à travers les évolutions les plus récentes. C'est pourquoi, nous sommes revenus aux sources du traitement de l'image, à savoir l'analyse du signal pour définir une représentation à la fois adaptée à un support bidimensionnel mais aussi à la couleur. C'est suivant cette direction que nous souhaitons poursuivre, notamment en continuant l'étude du codeur QWT à travers une modélisation probabiliste des coefficients (notamment de l'information de phase) et de l'intégration de la notion de points d'intérêts. De plus, cette application a un rôle structurant au niveau du laboratoire puisqu'il permet la réunion de nombreuses compétences, notamment avec la dimension "modélisation du canal de transmission". Cette approche intégrant la dimension qualité et canal dans le codeur nous semble originale et doit être poursuivie.

Ensuite, en reprenant les concepts du signal analytique, nous avons vu que la définition d'une phase couleur était complexe et non résolue et ceci dès la construction du signal monogène couleur. C'est pourquoi, nous proposons de poursuivre, notamment à travers la thèse de R. Soulard, un travail original proposant une extension couleur des travaux de Unser autour des ondelettes monogéniques avec une intégration des concepts concernant les tenseurs de structure étendus dans le cadre de la couleur. En parallèle, nous devons mener une étude ambitieuse théorique et numérique à partir du lien existant entre Radon et Riesz. Ceci doit permettre tout d'abord d'explicitier les connexions entre la transformée Ridgelet et l'analyse par ondelettes monogéniques. Ensuite à partir de notre schéma numérique de Radon discrète inversible, nous définirons un nouvel algorithme de calcul de la transformée de Riesz et par conséquent un schéma original de calcul de la transformée en ondelettes monogéniques.

Sécurité des données multimédia Enfin en ce qui concerne la stratégie de sécurisation des données multimédia, nous proposons de revenir au schéma original et de modifier l'espace d'insertion, à travers les nouvelles transformées décrites précédemment. Nous étudions actuellement un nouveau schéma utilisant la transformée Contourlet afin d'être au plus proche de la séparation fréquentielle faite par le système visuel humain pour intégrer au mieux la dimension psychovisuelle. Cette réflexion se poursuit en collaboration avec nos collègues étudiant la modélisation du système visuel humain (C. Fernandez-Maloigne, H. Saadane). En parallèle, nous allons développer un schéma utilisant l'information de phase de la transformée monogénique, qui apporte une description de la géométrie de l'image.

Le travail sur les codes correcteurs en collaboration avec P. Gaborit se poursuit avec l'intégration de nouvelles familles de codes mais aussi l'incorporation de nouvelles attaques, comme la numérisation, qui pose de nouvelles contraintes pour la protection. Notons que ce travail s'intègre dans le LABEX SIGMA-Lim obtenu par le laboratoire XLIM présentant un volet sécurité important.

Enfin, j'ai débuté cette année un nouvel axe ayant comme sujet la stéga-analyse suivant une approche qui propose de détourner les métriques de qualité d'image, à l'origine définies pour évaluer le résultat perceptuel d'un traitement. Je co-encadre un post-doctorat avec une collègue A.S. Capelle-Laizé sur ce sujet. D'une manière générale, la communauté "stéga-analyse" développe des techniques principalement issues du monde des statistiques, et nous souhaitons évaluer la pertinence d'une approche "image". Cela pose aussi la question du meilleur domaine de représentation pour identifier les ruptures du modèle.

Ces travaux doivent être menés en collaboration avec d'autres chercheurs, et dans le cadre de différents projets. Nous souhaitons pour cela poursuivre les collaborations existantes, et créer des nouveaux liens avec des spécialistes de domaines connexes, comme l'Informatique Graphique pour la modélisation des données, ou les mathématiques appliqués pour poursuivre l'intégration de certains outils pour la manipulation de données à valeurs vectorielles. Nous avons l'ambition pour certains thèmes (stéga-analyse par approche image, analyse du signala anlytique 2-D ..) de tenter de fédérer une communauté pour l'instant trop réduite ou dispersée.

D'une manière complémentaire, nous poursuivrons notre démarche de dépôt de différents projets académiques (ANR, Européen, ...) mais aussi contractuels avec différents industriels comme nous le faisons actuellement (SAGEMCOM, Technicolor ...).

Chapitre 7

Annexes

7.1 annexe : bancs de filtres quaternioniques

7.1.1 Préambule

Comme nous l'avons vu dans le premier chapitre, la transformée de Fourier a d'importantes limitations. La plus importante est la perte de notion chronologique dans le domaine fréquentiel. L'un des buts de la transformée en ondelettes discrète et du banc de filtres associé est justement d'étudier le comportement fréquentiel de différentes bandes de fréquences à travers une série de filtrage complémentaire. C'est pourquoi les études sur Fourier ne sont pas suffisantes. Nous avons voulu définir une représentation qui permet d'extraire les différents composants fréquentiels du signal. Cette représentation ne doit pas être redondante, si le signal est stocké sur N points, la représentation doit contenir N coefficients. De plus cette représentation doit être parfaitement inversible. Enfin nous désirons un algorithme simple et rapide.

Comme nous savons, un banc de filtres est un ensemble de filtres liés par des opérateurs d'échantillonnage. Le sous-échantillonnage est un opérateur de décimation, le sur-échantillonnage est un opérateur de dilatation. Dans un banc de filtres à deux canaux, les filtres d'analyse/synthèse sont de types passe-bas et passe-haut. Nous rappelons les notations et le concept du banc de filtres sur la figure 7.1.

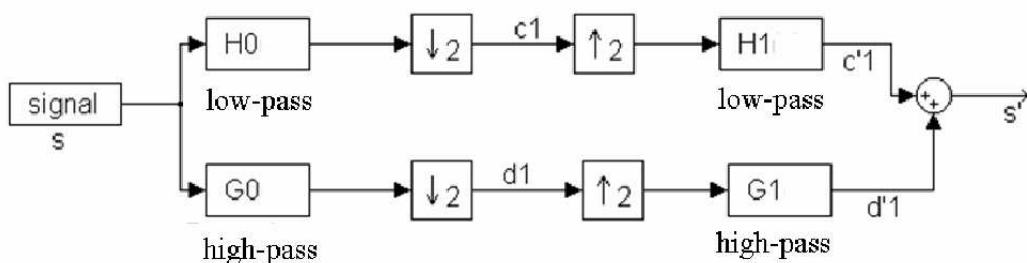


FIGURE 7.1 – Banc de filtres à 2 canaux

L'objectif de ce paragraphe est de construire les conditions sur les différents filtres afin d'obtenir dans le domaine quaternionique une reconstruction parfaite : $s = s'$ avec $s \in \mathcal{L}^2([0..N-1] \times [0..N-1], \mathbb{H})$.

7.1.2 L'existant sur les ondelettes quaternioniques

Il n'existe pas à notre connaissance actuellement de définition de banc de filtres quaternioniques. En revanche, nous pouvons trouver trois propositions de décompositions en ondelettes quaternioniques discrètes :

- Bayro [23] a proposé une transformée en ondelettes quaternioniques pour les images en niveaux de gris ($s \in \mathcal{L}^2([0..N-1]^2, \mathbb{R})$) avec le principe des filtres de Gabor modulés en quadrature. Les filtres utilisés sont définis par :

$$h[n_1, n_2] = g_\sigma[n_1, n_2] e^{2i\pi\xi_1 n_1} e^{2j\pi\xi_2 n_2}$$

avec σ l'écart-type de la fenêtre Gaussienne et ξ_1, ξ_2 les coordonnées fréquentielles. Cette définition correspond à une généralisation du travail de Bülw. Cette expression montre que la décomposition de Bayro étudie d'une manière disjointe les structures locales de parité différente selon n_1 et n_2 . A partir des coefficients de Gabor Quaternionique, un concept de trois phases est introduit (similaire à celui de Bülw [35]). Dans le papier cité, l'auteur utilise ce concept de phase pour estimer le flot géométrique de l'image analysée. Notons que cette décomposition en ondelettes est définie pour l'analyse et que l'aspect reconstruction n'est pas discutée.

- Chan et al [48, 49] proposent une transformée en ondelettes quaternioniques pour les images niveaux de gris suivant une extension directe du concept de la transformée en ondelettes complexe développée par Kingsbury. En 1-D la transformée en ondelettes complexe utilise des bases d'ondelettes complexes qui forment une paire de Hilbert (le spectre de l'ondelette complexe n'a pas d'énergie dans sa partie négative). La transformée en ondelette complexe est redondante mais le module des coefficients de la transformée est invariant à la translation. Chan et al proposent d'étendre ce concept au 2-D. Ceci est fait en définissant quatre éléments de bases à partir de la transformée de Hilbert 1-D calculée sur les fonctions d'ondelettes 2-D obtenues par produit tensoriel. La transformée de Hilbert 1-D est appliquée suivant ligne et colonne. Les quatre éléments sont associés aux quatre réels constituant un quaternion fabriquant ainsi l'ondelette quaternionique. Selon les auteurs cette nouvelle décomposition est approximativement une transformée Fourier quaternionique telle qu'elle est définie par Bülw mais localisée dans le temps. Comme pour le premier travail, les auteurs utilisent la notion de phase pour estimer le flot géométrique.
- Olhede et al [167] proposent la transformée en ondelettes hyperanalytique pour les images niveaux de gris. Ce travail est proche de celui de Chan : les quatre éléments de la base sont définis à partir de la transformée de Hilbert 1-D calculée sur les fonctions d'ondelettes 2-D obtenues par produit tensoriel et la transformée de Hilbert 1-D est appliquée suivant ligne et colonne. Notons qu'une seconde famille est décrite à partir des travaux de Felsberg et al [85], cette famille est appelée "Monogenic wavelets". Le concept d'ondelettes monogéniques a été discuté dans le chapitre 3.

Comme nous le constatons, tous ces travaux sont liés à ceux de Bülw et sont approximativement une Transformée de Fourier Quaternionique de Bülw localisée dans le temps. Toutes ces transformées permettent d'étudier séparément les différentes symétries selon les deux directions pour les images niveaux de gris.

Comme pour la transformée de Fourier, dans notre travail nous avons voulu définir un outil pour les images couleur et donc à valeur dans le domaine des quaternions. Ce problème n'est pas étudié actuellement dans la littérature. De plus, notre principal but n'est pas la transformée en ondelettes mais bien l'algorithme de transformée à savoir le banc de filtres afin de définir une transformée numérique à reconstruction parfaite. Nous devons pour cela étudier les trois éléments d'un banc de filtres : l'opération de filtrage, de sous-

échantillonnage et de sur-échantillonnage, dans le contexte des quaternions.

7.1.3 Opération de filtrage : produit de convolution

Nous avons déjà présenté globalement le produit de convolution quaternionique. Etant donné, l'objectif de construction d'un banc de filtres, nous reprenons en détail. Comme pour le cas complexe, l'opération de filtrage dans le domaine quaternionique va être perçu comme un concept de boîte noire avec une entrée e et une sortie s . Le filtrage se traduit alors par un produit de convolution entre l'entrée e et la réponse impulsionnelle h du filtre. Comme nous l'avons vu dans les précédentes sections,

Définition 5 la fonction définie par $\forall t \in \mathbb{R}$, and $e, h \in \mathcal{L}^2(\mathbb{R}, \mathbb{H})$ $s(t) = e \circledast h(t) = \int e(\tau).h(t - \tau)d\tau$ est le produit de convolution entre e et h .

Dans le cadre d'un système discret la réponse impulsionnelle est un signal discret $\{h[k]\}_{k \in \mathbb{N}}$. Le lien entre l'entrée et la sortie est alors un produit de convolution discret :

Définition 6 (DQCV) Soit $N \in \mathbb{N}$, $e, h \in \mathcal{L}^2([0..N-1], \mathbb{H})$: la fonction définie par $\forall n \in [0..N-1]$

$$s[n] = e \circledast h[n] = \sum_{\tau=0}^{N-1} e[|n - \tau|_N]h[|\tau|_N]$$

est le produit de convolution discret à droite entre e et h .

La notation $|p|_N$ indique que les indices sont calculés modulo N .

Nous savons que dans le domaine complexe le produit de convolution entre deux signaux dans le domaine temporel correspond à un produit terme à terme dans le domaine fréquentiel :

$$g[n] = s \circledast h[n] \iff G[k] = S[k].H[k]$$

Comme le produit quaternionique n'est pas commutatif, cette propriété n'est pas toujours vraie pour les quaternions. Si nous prenons la transformée de Fourier quaternionique telle qu'elle est définie par Sangwine, nous rappelons que nous avons la relation suivante :

Soit $(e, \vec{\mu}, \vec{\nu}, \vec{\mu}\vec{\nu})$ une base orthonormale de $\mathbb{H}$, on peut décomposer h suivant sa partie parallèle et perpendiculaire à μ :

$$h[n] = h_a[n] + h_b[n].\nu$$

où $h_a[n] = h_e[n] + h_\mu[n].\mu$ et $h_b[n] = h_\nu[n] + h_{\mu\nu}[n].\mu$ avec h_e, h_μ, h_ν et $h_{\mu\nu}$ la projection de h sur $(e, \mu, \nu, \mu\nu)$.

Soit $h, e \in \mathcal{L}^2(0..N-1, \mathbb{H})$ et s définis par $s = h \circledast e$ tel que $s \in \mathcal{L}^2(0..N-1, \mathbb{H})$. alors pour $\forall k \in [0..N-1]$,

$$\mathcal{F}_\mu(s)[k] = \mathcal{F}_\mu(h_a)[k]\mathcal{F}_\mu(e)[k] + \mathcal{F}_\mu(h_b)[k]\nu\mathcal{F}_\mu(e)[-k]$$

Sous certaines conditions, la relation précédente peut être simplifiée. C'est le cas lorsque s est pair :

$$e[n] = e[-n] \iff E[k] = E[-k]$$

alors la convolution dans le domaine temporel correspond à un simple produit dans le domaine fréquentiel.

Afin de simplifier le problème de la définition des filtres, nous inversons le problème et nous nous interrogeons sur la relation existante dans le domaine spatial lorsque que nous calculons un produit terme à terme dans le domaine fréquentiel. Le résultat est similaire à la relation précédente, à savoir

Soit $h, e \in \mathcal{L}^2(0..N-1, \mathbb{H})$ et s définis par $S[k] = H[k].E[k]$ alors $\forall n \in [0..N-1]$,

$$s[n] = h_a \otimes e[n] + (h_b.\nu) \otimes e[-n]$$

Tous ces problèmes sont discutés en détails sous une autre forme par Pei et al [175].

7.1.4 Définition du sous-échantillonnage

Nous voulons modéliser l'influence dans le domaine fréquentiel Quaternionique du sous-échantillonnage dans le domaine spatial. Une opération de sous-échantillonnage peut être décrite par : nous convertions uniquement les échantillons d'indice pair. L'opération se note $s \rightarrow \downarrow 2 \rightarrow s'$.

Proposition 2 Soit $N \in \mathbb{N}$, $s \in \mathcal{L}^2([0..N-1], \mathbb{H})$ $\mu \in \mathbb{P} \cap \mathbb{S}$ et $S[k] = F_\mu^l(s)[k]$ alors comme pour le cas complexe, l'opération de sous-échantillonnage se traduit dans le domaine fréquentiel Quaternionique par :

$$s' = [\downarrow 2] s \implies S'[k] = \frac{1}{2} \left[S[k] + S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots \frac{N}{2} - 1$$

avec S' une séquence $\frac{N}{2}$ -périodique.

Pour analyser le sous-échantillonnage, nous le découpons en deux processus élémentaires :

Soit u la séquence s avec tous ces éléments impaires mis à 0 :

$$u[n] = \begin{cases} s[n] & \text{si } n \text{ paire} \\ 0 & \text{sinon} \end{cases}$$

$u[n]$ peut être défini par :

$$u[n] = \frac{1}{2} (s[n] + (-1)^n s[n])$$

Le second terme inclut $(-1)^n$ afin que l'addition supprime les éléments impairs n . Puisque $\mu \in \mathbb{P} \cap \mathbb{S}$ nous pouvons écrire $(e^{-\mu\pi})^n = (-1)^n$ et donc

$$u[n] = \frac{1}{2} (s[n] + (e^{-\mu\pi})^n s[n])$$

Dans le domaine fréquentiel, nous avons :

$$\begin{aligned} U[k] &= \sum_{n=0}^{N-1} e^{-2\mu\pi \frac{kn}{N}} . u[n] = \sum_{n=0}^{N-1} \frac{1}{2} e^{-2\mu\pi \frac{kn}{N}} (s[n] + (e^{-\mu\pi})^n s[n]) \\ &= \frac{1}{2} \left[\sum_{n=0}^{N-1} e^{-2\mu\pi \frac{kn}{N}} . s[n] + \sum_{n=0}^{N-1} e^{-2\mu\pi \frac{kn}{N} - \mu\pi n} . s[n] \right] \text{ pour } k = 0 \dots N-1 \end{aligned}$$

alors

$$U[k] = \frac{1}{2} \left[S[k] + S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots N-1$$

La seconde opération est telle que :

$$s'[n] = u[2n]$$

car le résultat ne conserve que les éléments d'indice pair.

Dans le domaine fréquentiel, nous avons

$$S'[k] = \sum_{n=0}^{\frac{N}{2}-1} e^{-2\mu\pi \frac{kn}{N/2}} \cdot u[2n] \text{ pour } k = 0 \dots \frac{N}{2} - 1$$

avec le changement de variable $n' = 2n$

$$= \sum_{n'=0}^{N-1} e^{-2\mu\pi n' \frac{2}{N}(\frac{k}{2})} \cdot u[n'] = U[k] \text{ pour } k = 0 \dots \frac{N}{2} - 1$$

Il en découle que le sous-échantillonnage du signal peut être défini dans le domaine fréquentiel quaternionique :

$$s' = [\downarrow 2] s \implies S'[k] = \frac{1}{2} \left[S[k] + S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots \frac{N}{2} - 1$$

Puisque S est N -périodique, nous observons que S' est $\frac{N}{2}$ -périodique.

7.1.5 Définition du sur-échantillonnage

Durant le processus de reconstruction, la première étape consiste à reconstruire la taille d'origine de la séquence. Pour cela, nous appliquons une opération de sur-échantillonnage, à savoir qu'un zéro est inséré un échantillon sur deux.

Comme précédemment nous voulons caractériser cette opération de ré-échantillonnage dans le domaine fréquentiel. Le signal ré-échantillonné a des zéros pour tous ces éléments d'indices impairs :

$$s' = [\uparrow 2] s \implies \begin{cases} s'[2n] = s[n] \\ s'[2n+1] = 0 \end{cases} \text{ avec } \{s[n]\}_{n=0 \dots \frac{N}{2}-1}$$

Proposition 3 Soit $N \in \mathbb{N}$, $s \in \mathcal{L}^2([0, \frac{N}{2} - 1], \mathbb{H})$ $\mu \in \mathbb{P} \cap \mathbb{S}$ et $S[k] = F_\mu^l(s)[k]$ alors

$$s' = [\uparrow 2] s \implies S'[k] = S[k] \text{ pour } k = 0 \dots N - 1$$

avec S une séquence $\frac{N}{2}$ -periodique.

La transformée de Fourier quaternionique de s' est

$$S'[k] = \sum_{n=0}^{N-1} e^{-2\mu\pi \frac{kn}{N}} \cdot s'[n] \text{ pour } k = 0 \dots N - 1$$

Si n est paire alors $s'[n] = s[n/2]$ sinon $s'[n] = 0$. La transformée de Fourier quaternionique peut alors se ré-écrire :

$$\begin{aligned} S'[k] &= \sum_{n=0 \text{ } n \text{ paire}}^{N-1} e^{-2\mu\pi \frac{kn}{N}} \cdot s[n/2] \text{ nous posons } b = n/2 \\ &= \sum_{b=0}^{N/2-1} e^{-2\mu\pi \frac{(2k)b}{N}} \cdot s[b] = \sum_{b=0}^{N/2-1} e^{-2\mu\pi \frac{kb}{N/2}} \cdot s[b] \\ &= S[k] \text{ pour } k = 0 \dots N - 1 \end{aligned}$$

7.1.6 Banc de filtres quaternionique à reconstruction parfaite

Comme nous l'avons déjà mentionné, le but est de définir les conditions permettant d'obtenir une reconstruction parfaite : $s = s'$ avec $s \in \mathcal{L}^2([0..N-1], \mathbb{H})$.

Dans cette étude nous considérons que l'opération de filtrage se traduit par un produit dans le domaine fréquentiel :

Définition 7 Soit $N \in \mathbb{N}$, $e \in \mathcal{L}^2([0..N-1], \mathbb{H})$, $E = F_{-\mu}^l(e)$ la transformée DQFT gauche du signal original e de direction μ et $H \in \mathcal{L}^2([0..N-1], \mathbb{H})$ $H = F_{-\mu}^l(h)$ la réponse fréquentielle du filtre et $\mu \in (\mathbb{P} \cap \mathbb{S})$.

La fonction définie par $\forall k \in [0..N-1] S[k] = H[k].E[k]$ est l'opération de filtrage de e par h .

Puisque l'entrée et la sortie sont liées par une simple relation, le produit fréquentiel est plus aisé pour une construction des filtres du banc.

Les conditions pour une reconstruction parfaite sont exprimées directement dans le domaine fréquentiel quaternionique : nous considérons $H \in \mathcal{L}^2([0..N-1], \mathbb{H})$ une séquence N -periodique correspondant à la réponse fréquentielle du filtre passe-bas (avec un index 0 pour l'analyse et un index 1 pour la reconstruction), $G \in \mathcal{L}^2([0..N-1], \mathbb{H})$ une séquence N -periodique correspondant à la réponse fréquentielle du filtre passe-haut (avec un index 0 pour l'analyse et un index 1 pour la reconstruction) et $S = F_{-\mu}^l(s)$ une séquence N -periodique correspondant à la transformée DQFT gauche du signal original s de direction μ avec $s \in \mathcal{L}^2([0..N-1], \mathbb{H})$ et $\mu \in (\mathbb{P} \cap \mathbb{S})$.

Si nous étudions la chaîne passe-bas sur la figure 7.1 :

$$\begin{aligned} c_1 &= [\downarrow 2] (\text{filtering}(h_0, s)) \\ C_1[k] &= \frac{1}{2} \left[H_0[k]S[k] + H_0[k + \frac{N}{2}]S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots \frac{N}{2} - 1 \end{aligned}$$

Puis sont appliqués un filtrage et un sur-échantillonnage :

$$C'_1[k] = \frac{1}{2} H_1[k] \left[H_0[k]S[k] + H_0[k + \frac{N}{2}]S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots N-1$$

Nous obtenons la même relation pour la chaîne passe-haut avec les filtres H changés en filtres G :

$$D'_1[k] = \frac{1}{2} G_1[k] \left[G_0[k]S[k] + G_0[k + \frac{N}{2}]S[k + \frac{N}{2}] \right] \text{ pour } k = 0 \dots N-1$$

Le banc de filtres combine les deux canaux pour obtenir s'

$$S'[k] = \frac{1}{2} [G_1[k]G_0[k] + H_1[k]H_0[k]] . S[k] + \frac{1}{2} \left[G_1[k]G_0[k + \frac{1}{2}] + H_1[k]H_0[k + \frac{1}{2}] \right] S[k + \frac{1}{2}]$$

Nous en déduisons que le banc de filtres permet une reconstruction parfaite si :

– Sans distortion :

$$G_1[k]G_0[k] + H_1[k]H_0[k] = 2 \text{ pour } k = 0 \dots N-1 \quad (7.1)$$

– Suppression du repliement

$$G_1[k]G_0[k + \frac{N}{2}] + H_1[k]H_0[k + \frac{N}{2}] = 0 \text{ pour } k = 0 \dots N-1 \quad (7.2)$$

Nous observons que nous obtenons les mêmes conditions qu'un banc de filtres classique mais avec la différence que l'ordre des éléments est primordial. Puisque le produit quaternionique n'est pas commutatif, en utilisant une autre définition pour l'opération de filtrage, nous obtenons d'autres conditions.

Afin d'illustrer la transformée en ondelettes quaternioniques pour l'image couleur, nous proposons maintenant une simple construction : le banc de filtres couleur de Shannon.

7.1.7 Un exemple : le banc de filtres quaternioniques de Shannon

Les ondelettes de Shannon sont construites à partir de la multirésolution de Shannon qui propose d'approximer les fonctions par leurs restrictions à des intervalles basse-fréquences. Cette multirésolution est associée au filtre passe-bas suivant :

$$H(\xi) = \sqrt{2} \cdot 1_{[-\frac{1}{4}, \frac{1}{4}]} \text{ pour } \xi \in [-\frac{1}{2}, \frac{1}{2}] \text{ avec } H(\xi) = \sum_n h[n] e^{-2j\pi\xi n}$$

Nous proposons d'adapter le banc de filtres de Shannon au domaine quaternionique..

Pour construire le banc de filtres quaternioniques, nous pouvons ajouter différentes contraintes sur les filtres. Par exemple,

- Les coefficients d'ondelettes doivent être des quaternions purs (pour être des coefficients couleur).

Dans ce cas, les signaux c_1 et d_1 de la figure 7.1 doivent être tels que $c_1, d_1 \in \mathcal{L}^2([0..N-1], \mathbb{P})$.

La transformée de Fourier quaternionique de c_1 et d_1 doit alors vérifier une propriété de symétrie antihermitienne.

Dans notre cas, nous proposons de suivre le concept du banc de filtres à deux canaux avec des filtres miroirs en quadrature : dans ce cas, les filtres de reconstruction sont la version inversée des filtres d'analyse : $h_1[n] = h_0[-n]$ et $g_1[n] = g_0[-n]$. Nous savons que si h_1, g_1 sont réels, alors $H_1[k] = \bar{H}_0[k]$ et $G_1[k] = \bar{G}_0[k]$.

Dans ce cas, nous proposons que les deux filtres passe-bas soient définis de la façon suivante :

$$H_0[k] = H_1[k] = \begin{cases} \sqrt{2} & \text{pour } k \in [0, \frac{N}{4}-1] \cup [\frac{3N}{4}+1, N-1] \\ 0 & \text{pour } k \in [\frac{N}{4}+1, \frac{3N}{4}-1] \end{cases}$$

Dans ce cas G_0 et G_1 doivent satisfaire :

$$\begin{cases} G_0[k]G_1[k] = 0 \\ G_0[k + \frac{N}{2}]G_1[k] = 0 \end{cases} \text{ pour } k \in [0, \frac{N}{4}-1] \cup [\frac{3N}{4}+1, N-1]$$

$$\begin{cases} G_0[k]G_1[k] = 2 \\ G_0[k + \frac{N}{2}]G_1[k] = 0 \end{cases} \text{ pour } k \in [\frac{N}{4}+1, \frac{3N}{4}-1]$$

Nous proposons la solution simple suivante :

$$G_0[k] = G_1[k] = \begin{cases} 0 & \text{pour } k \in [0, \frac{N}{4}-1] \cup [\frac{3N}{4}+1, N-1] \\ \sqrt{2} & \text{pour } k \in [\frac{N}{4}+1, \frac{3N}{4}-1] \end{cases}$$

Il reste les points d'intersection aux coordonnées $k = \frac{N}{4}$ et $k = \frac{3N}{4}$ car à ces points précis $H_0[k + \frac{N}{2}]H_1[k] \neq 0$ si $H_0[\frac{N}{4}] = H_0[\frac{3N}{4}] = \sqrt{2}$.

Nous proposons les relations suivantes pour les points d'intersection :

$$\begin{aligned} H_0[N/4] &= H_0[3N/4] = 1 ; H_1 = \bar{H}_0 \\ G_0[N/4] &= -1.G_0[3N/4] = \mu ; G_1 = \bar{G}_0 \end{aligned}$$

avec $\mu \in \mathbb{P} \cap \mathbb{S}$. Avec cette construction H_0 , G_0 , H_1 , and G_1 satisfont les conditions de reconstruction parfaite. Notons que H_0 et G_0 présentent une symétrie hermitienne.

Nous avons défini quatre filtres pour la décomposition/reconstruction en ondelettes d'un signal 1-D. Pour son extension à des données 2-D comme les images, nous proposons une simple approche par produit tensoriel. Pour cela, nous appliquons l'algorithme 1-D sur chaque ligne, puis sur la matrice résultante nous appliquons l'algorithme 1-D sur les colonnes.

Pour conclure sur cette construction, nous pouvons faire différentes remarques sur notre définition du filtre passe-haut. En effet, nous avons proposé une construction empirique de G_0 mais dans l'approche classique complexe, il existe une relation systématique permettant de construire G_0 à partir du filtre passe-bas [245, 152] :

$$G_0(\xi) = e^{-2j\pi\xi}.\bar{H}_0(\xi + \frac{1}{2})$$

Nous proposons d'expliciter cette relation dans le cadre quaternionique et de vérifier si elle permet de satisfaire les conditions de reconstructions parfaites (7.1) et (7.2) (avec $H_1[k] = \bar{H}_0[k]$ et $G_1[k] = \bar{G}_0[k]$) :

$$G[k]\bar{G}[k] + H[k]\bar{H}[k] = 2 \text{ pour } k = 0 \dots N-1 \quad (7.3)$$

$$\bar{G}[k]G[k + \frac{N}{2}] + \bar{H}[k]H[k + \frac{N}{2}] = 0 \text{ pour } k = 0 \dots N-1 \quad (7.4)$$

Nous proposons deux généralisations de la construction :

$$G[k] = e^{-2\mu\pi \frac{k}{N}} \bar{H}[k + \frac{N}{2}] \quad (7.5)$$

$$\text{ou} \quad (7.6)$$

$$G[k] = \bar{H}[k + \frac{N}{2}] e^{-2\mu\pi \frac{k}{N}} \quad (7.7)$$

Ces deux relations ne sont pas équivalentes puisque le produit n'est pas commutatif.

– Pour la généralisation (7.5) :

$$G[k]\bar{G}[k] = e^{-2\mu\pi \frac{k}{N}} \bar{H}[k + \frac{N}{2}].H[k + \frac{N}{2}].e^{2\mu\pi \frac{k}{N}} = e^{-2\mu\pi \frac{k}{N}} |H[k + \frac{N}{2}]|^2 . e^{2\mu\pi \frac{k}{N}} = \left| H[k + \frac{N}{2}] \right|^2$$

La première condition (7.3) est satisfaite si H est tel que (condition classique associée à une base orthonormale) :

$$\left| H[k + \frac{N}{2}] \right|^2 + |H[k]|^2 = 2 \quad (7.8)$$

$$\bar{G}[k]G[k + \frac{N}{2}] = H[k + \frac{N}{2}]e^{2\mu\pi \frac{k}{N}} . e^{-2\mu\pi \frac{k+0.5N}{N}} \bar{H}[k] = -H[k + \frac{N}{2}]\bar{H}[k]$$

La seconde condition (7.4) est vérifiée si H est telle que

$$H[k + \frac{N}{2}]\bar{H}[k] = \bar{H}[k]H[k + \frac{N}{2}]$$

. C'est toujours vrai si H est réel.

– Pour la généralisation (7.7) :

$$G[k]\bar{G}[k] = \bar{H}[k + \frac{N}{2}]e^{-2\mu\pi\frac{k}{N}}.e^{2\mu\pi\frac{k}{N}}H[k + \frac{N}{2}] = \left|H[k + \frac{N}{2}]\right|^2$$

Nous obtenons la même relation que (7.5).

$$\bar{G}[k]G[k + \frac{N}{2}] = e^{2\mu\pi\frac{k}{N}}.H[k + \frac{N}{2}].\bar{H}[k]e^{-2\mu\pi\frac{k+0.5N}{N}}$$

Nous ne pouvons pas simplifier cette expression puisque le produit n'est pas commutatif. Toutefois, la seconde condition (7.4) est toujours vérifiée si H est réel.

Nous pouvons conclure que si H est réel et qu'il satisfait la relation (7.8), nous pouvons alors utiliser l'expression (7.5) ou (7.7) pour définir le filtre passe-haut. C'est le cas pour le filtre passe-bas associé à la décomposition de Shannon, et donc G peut être défini par $G[k] = e^{-2\mu\pi\frac{k}{N}}\bar{H}[k + \frac{N}{2}]$ pour $k = 0 \dots N - 1$. C'est l'extension directe de la multirésolution de Shannon au domaine quaternionique.

Tout d'abord nous proposons d'illustrer la construction précédente à travers l'illustration de trois fonctions de la base associées à notre banc de filtres (figure 7.2). Comme pour la description du domaine fréquentiel, nous initialisons le domaine des ondelettes avec un Dirac et nous étudions la réponse spatiale après transformée inverse. Nous observons que ces atomes correspondent à une version colorée des ondelettes "classiques" ($\mu = \mu_{grey}$).

Pour illustrer cette décomposition, nous créons une image test à partir de trois objets colorés (figure 7.3.a). Après décomposition, nous observons que la représentation contient tous les éléments attendus : la partie passe-bas correspond à une approximation de l'image couleur ; la partie passe-haut tend à présenter peu de coefficients significatifs : un coefficient est large uniquement si un contour est présent dans le support de l'ondelettes. De plus la couleur des coefficients donne des informations sur la couleur des discontinuités.

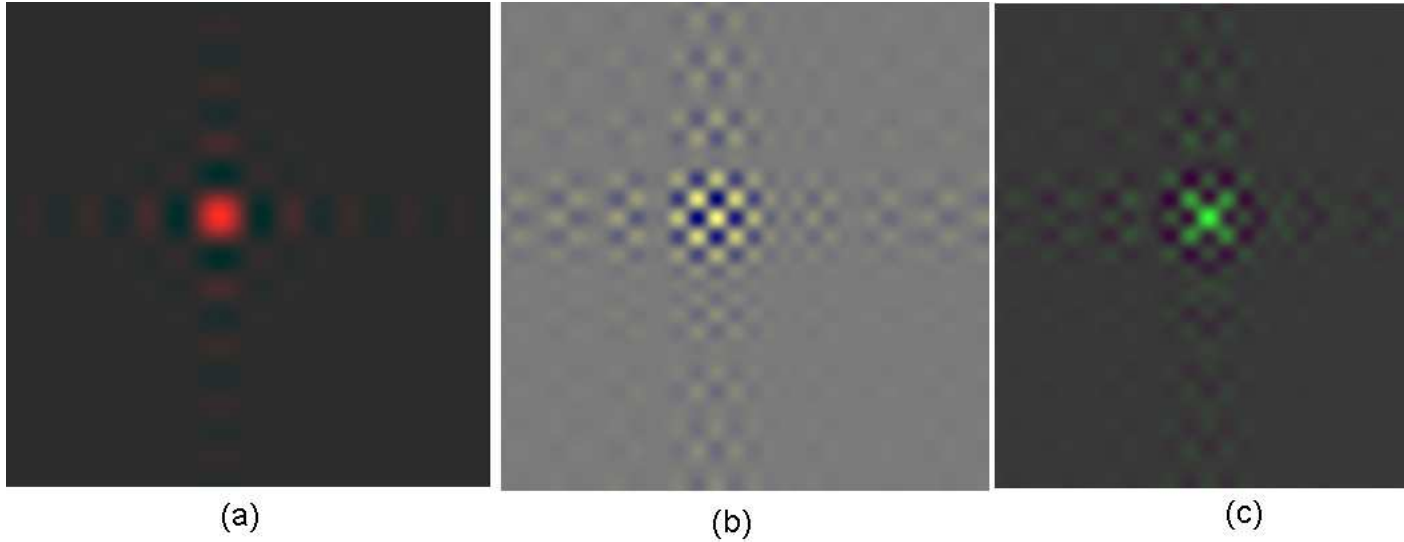


FIGURE 7.2 – Fonctions associées au banc de filtres quaternioniques

7.2 Annexe : Image de Test pour les techniques de tatouage

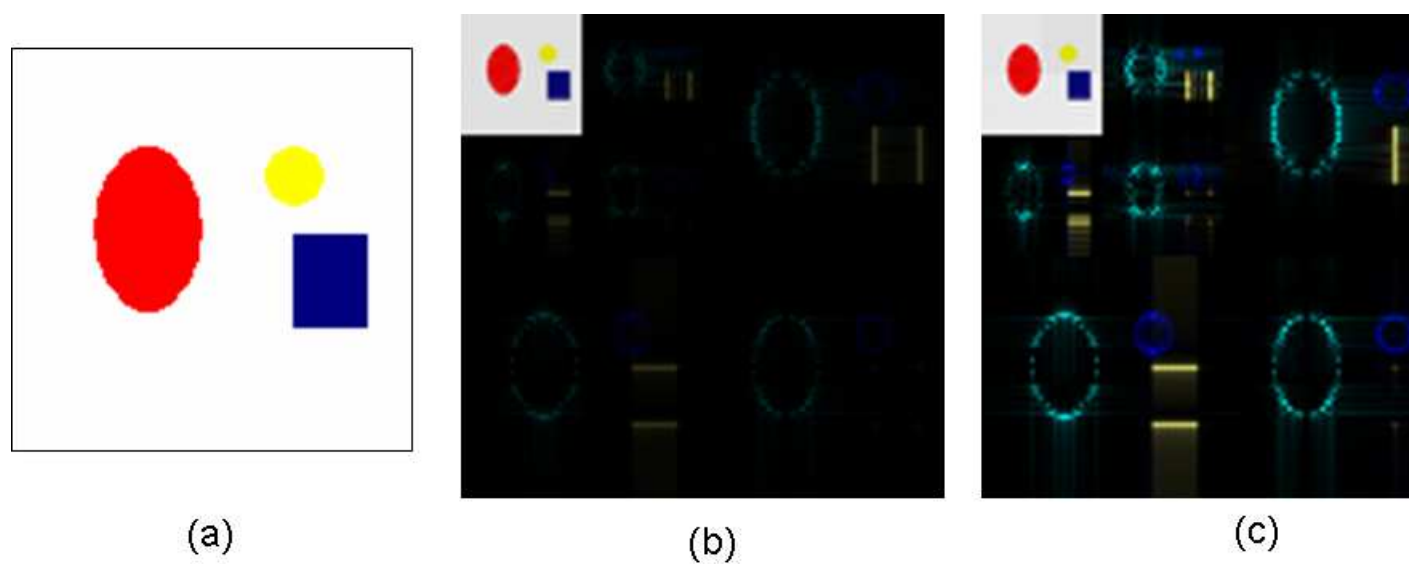


FIGURE 7.3 – Décomposition en ondelettes d’une image de test : (a) Image originale (b) Les coefficients d’ondelettes (c) La décomposition normalisée pour chaque plan

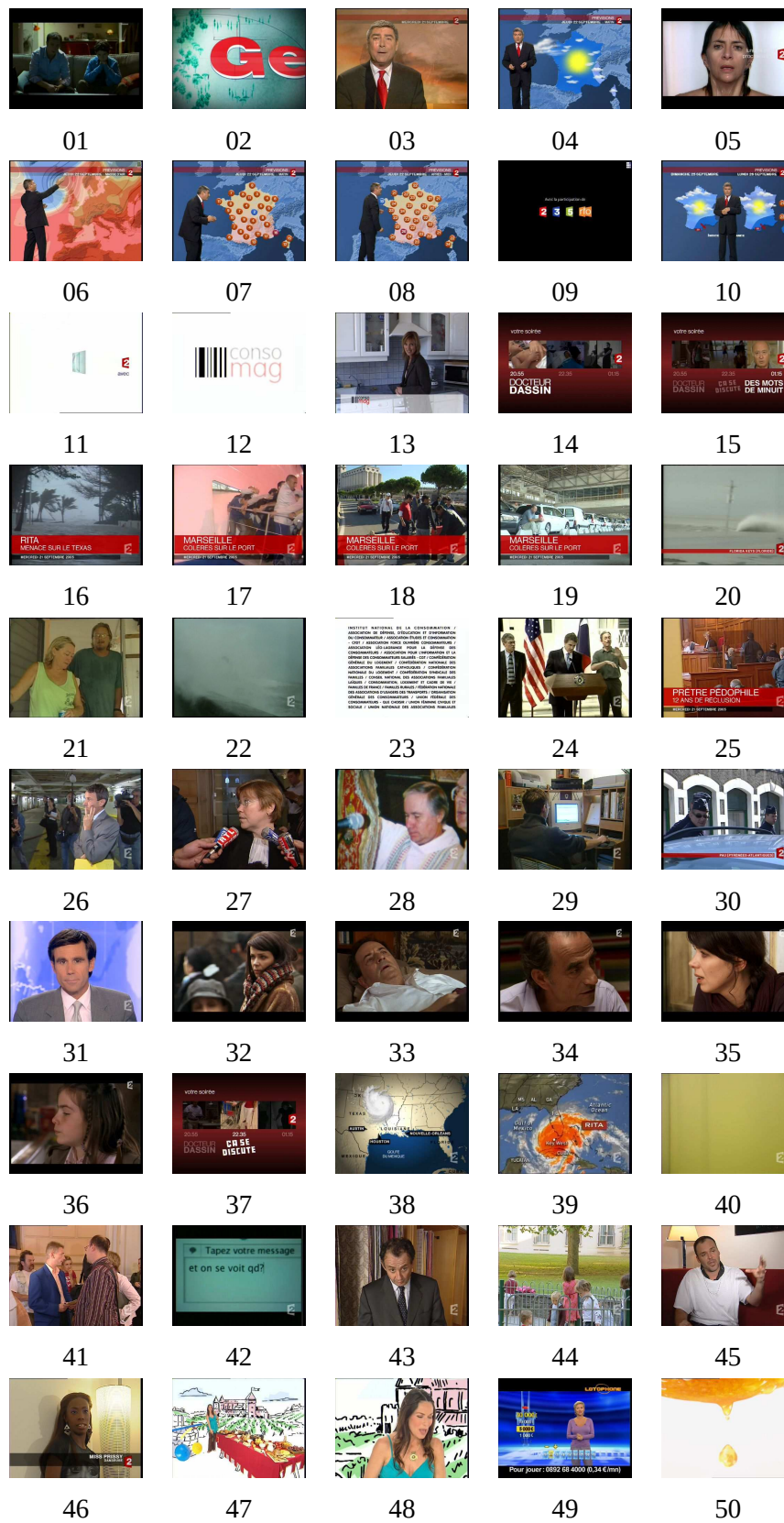


FIGURE 7.4 – Images de test France Telecom



FIGURE 7.5 – Exemple de l’effet d’une compression de qualité 1% (à gauche) et 96% (à droite) sur l’image "lenna.bmp"

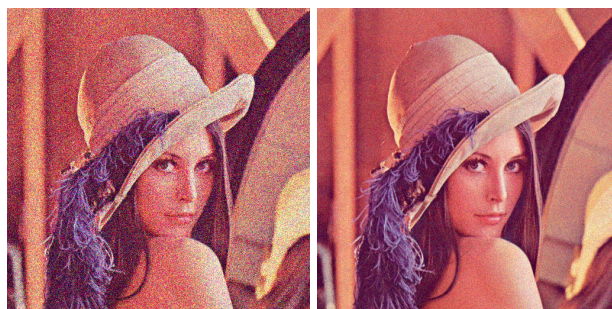


FIGURE 7.6 – Exemple de l’effet de l’ajout de bruit 1 (à gauche) et de 10 (à droite) sur l’image "lenna.bmp"



FIGURE 7.7 – Exemple de la modification du gamma de 0.1 (à gauche) et de 5 (à droite) sur l’image "lenna.bmp"



FIGURE 7.8 – Illustration de la modification de la saturation pour l’image Lenna avec comme paramètres -100 et 100 respectivement.



FIGURE 7.9 – Exemple de modification du contraste de -5 (à gauche) et de $+5$ (à droite) sur l'image "lenna.bmp"



FIGURE 7.10 – Illustration de la modification de la teinte pour l'image Lenna avec comme paramètre -100° et 100° respectivement.



FIGURE 7.11 – Exemple de filtrage passe-bas de type moyenneur pour un filtre de taille 3 (à gauche) et 11 (à droite) sur l'image "lenna.bmp"

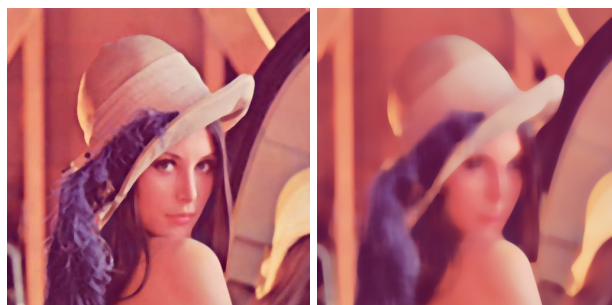


FIGURE 7.12 – Exemple de filtrage de type médian pour un filtre de taille 3 (à gauche) et 11 (à droite) sur l'image "lenna.bmp"



FIGURE 7.13 – Exemple de quantification en 2 couleurs (à gauche) et en 252 couleurs (à droite) sur l'image "lenna.bmp"

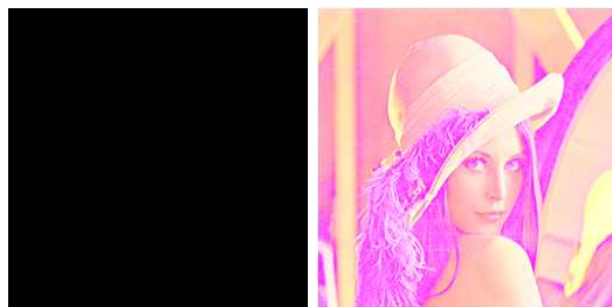


FIGURE 7.14 – Illustration de la modification de la Luminance pour l'image Lenna avec comme paramètres -100 et 100 respectivement.



Image "Ours"



Image "Feuille"



Image "Singe"



Image "Plante"

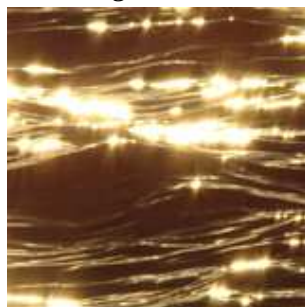


Image Eau

FIGURE 7.15 – Images de test Code Correcteur

Bibliographie

- [1] Wadood Abdul. *Codes Correcteurs et transformée psychovisuelle pour le tatouage d'images*. Thèse de doctorat, Université de Poitiers, 2011.
- [2] Wadood Abdul, P. Carré, and Philippe Gaborit. Human visual system-based color image steganography using the contourlet transform. In *SPIE 2010*, page 10.1117/12.838899, États-Unis, January 2010.
- [3] Wadood Abdul, Philippe Carré, and Philippe Gaborit. List decoding of Reed Solomon codes for wavelet based colour image watermarking scheme. In *IEEE International Conference on Image Processing - ICIP 2009*, Cairo, Égypte, November 2009.
- [4] Wadood Abdul, Philippe Carré, Hakim Saadane, and Philippe Gaborit. Watermarking using multiple visual channels for perceptual color spaces. In *IEEE International Conference on Image Processing*, pages 2597–2600, Hong-Kong, Chine, 2010.
- [5] Wadood Abdul, Philippe Gaborit, and P. Carré. Private anonymous fingerprinting for color images in the wavelet domain. In *SPIE 2010*, page 10.1117/12.838996, États-Unis, January 2010.
- [6] B.K. Alpert. *Wavelets and Other Bases for Fast Numerical Linear Algebra*. Academic Press, 1993.
- [7] E. Andres. Modélisation analytique discrète d'objets géométriques. 2000.
- [8] E. Andres, R. Acharya, and C. Sibata. Discrete analytical hyperplanes. *Graphical Models and Image Processing*, 59(5) :302–309, 1997.
- [9] Dawit Assefa, Lalu Mansinha, Kristy F. Tiampo, Henning Rasmussen, and Kenzu Abdella. Local quaternion fourier transform and color image texture analysis. *Signal Processing*, 90(6) :1825 – 1835, 2010.
- [10] I. Avcibas, N. Memon, and B. Sankur. Steganalysis using image quality metrics. *IEEE Transaction on Image Processing*, 12(2) :221–229, 2003.
- [11] I. Avcibas, B. Sankur, and K. Sayood. Statistical evaluation of image quality measures. *Journal of Electronic Imaging*, 11(2) :206–223, 2002.
- [12] A. Averbuch, R. Coifman, D. Donoho, M. Israeli, and J. Walden. Fast slant stack : A notion of radon transform for data in a cartesian grid which is rapidly computable, algebraically exact, geometrically faithful and invertible. *SIAM Scientific Computing*, 72890, 2001.
- [13] A. Averbuch and Y. Shkolnisky. 3d discrete x-ray transform. *Applied and Computational Harmonic Analysis*, 17(3) :259–276, 2004.
- [14] C M Bachmann, T L Ainsworth, and R A Fusina. Exploiting manifold geometry in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3) :441–454, 2005.

- [15] R.H. Bamberger and M.J.T. Smith. A filter bank for the directional decomposition of images : Theory and design. *IEEE Trans. Signal Processing*, 40(4) :882–893, April 1992.
- [16] M. Barni, F. Bartolini, V. Cappelini, and A. Piva. A dct-domain system for robust image watermarking. *Signal Processing*, 66(3) :357–372, May 1998.
- [17] P. Bas, J.-M. Chassery, and B. Macq. Image watermarking : an evolution to content based approaches. In *Pattern Recognition*, volume 35, pages 545–565, March 2002.
- [18] Patrick Bas, Nicolas Le Bihan, and Jean-Marc Chassery. Color watermarking using quaternion fourier tranform. In *IEEE International conference on Acoustics Speech and Signal Processing (ICASSP)*, volume 3, pages 521–524, 2003.
- [19] Patrick Bas, Nicolas Le Bihan, and Jean-Marc Chassery. Utilisation de la transformée de fourier quaternionique en tatouage d’images couleur. In *Colloque GRETSI*, Paris, France, Septembre 2003.
- [20] Thomas Batard, Michel Berthier, and Christophe Saint-Jean. Clifford Fourier Transform for Color Image Processing. In Gerik Bayro-Corrochano, Eduardo ; Scheuermann, editor, *Geometric Algebra Computing*, Engineering and Computer Science, pages 135–162. Springer Verlag, 2010.
- [21] S. Battiato, D. Catalano, G. Gallo, and R. Gennaro. A color opponency watermarking scheme for digital images. In *Security and watermarking of multimedia contents II, proceedings of SPIE*, volume 3971, pages 510–515, 2000.
- [22] S. Baudry, J. Delaigle, B. Sankur, B. Macq, and Henri Maitre. Analyses of error correction strategies for typical communication channels in watermarking. *Signal Processing*, 81 :1239–1250, June 2001.
- [23] E. Bayro-Corrochano. Multiresolution image analysis using the quaternion wavelet transform. *Numerical Algorithms*, 39 :35–55, 2005.
- [24] Eduardo Bayro-Corrochano. The theory and use of the quaternion wavelet transform. *J. Math. Imaging Vis.*, 24(1) :19–35, 2006.
- [25] Gwendoline Blanchet, Lionel Moisan, and Bernard Rougé. Measuring the global phase coherence of an image. In *IEEE International Conference on Image Processing*, 2008.
- [26] H. Blum. Biological shape and visual science. *Theoretical Biology*, 38 :205–287, 1973.
- [27] R.C. Bose and D.K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3(1) :68 – 79, 1960.
- [28] Fred Brackx, Nele De Schepper, and Frank Sommen. The clifford-fourier transform. *The Journal of Fourier Analysis and Applications*, 11,6 :669–681, December 2005.
- [29] Fred Brackx, Nele De Schepper, and Frank Sommen. The two-dimensional clifford-fourier transform. *Journal of Mathematical Imaging and Vision*, 26,1-2 :5–18, 2006.
- [30] M. Brady. A fast discrete approximation algorithm for the radon transform. *SIAM J. Computer*, 27(1) :107–119, February 1998.
- [31] J. Bresenham. Algorithms for computer control of a digital plotter. *IBM Sys. Journal*, 4(1) :25–30, January 1965.
- [32] P. Brodatz. *Textures : A Photographic Album for Artists and Designers*. Dover publications, New York, 1966.

- [33] M. Bronstein, A. Bronstein, and M. Zibulevsky. Applications of non-uniform fft for acoustic diffraction tomography. Technical report, <http://visl.technion.ac.il/bron/works>, march 2002.
- [34] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A non-local algorithm for image denoising. In *CVPR (2)'05*, pages 60–65, 2005.
- [35] Thomas Bülow. *Hypercomplex Spectral Signal Representations for the processing and Analysis of Images*. PhD thesis, Kiel University, Germany, August 1999.
- [36] P. J. Burt and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, COM-31,4 :532–540, 1983.
- [37] L. Macaire C. Fernandez, editor. *Imagerie couleur numerique : Avancees et perspectives*. Hermes, 2012.
- [38] E. Candès. *Ridgelets : Theory and Applications*. PhD thesis, Dpt of Statistics, Stanford University, <http://www.acm.caltech.edu/emmanuel/>, August 1998.
- [39] P. Carré and E. Andrès. Discrete ridgelet transform : application to the denoising of color images. In *CGIP'2002*, April 2002. Poitiers, France.
- [40] P. Carré and E. Andrès. Ridgelet transform based on reveillés discrete lines. In *DGCI'02*, volume 2301 of *Lecture Notes in Computer Science*, pages 417–427. Springer Verlag, April 2002. Bordeaux, France.
- [41] P. Carré and E. Andres. Discrete analytical ridgelet transform. *Signal Processing*, 84(11) :2165–2173, November 2004.
- [42] P. Carré and L. Fuchs. Geometric algebras : multicomponent images analysis and geometrical modelisation. In *Workshop Clifford Algebras and Lie Groups for multidimensional image processing*, La Rochelle, 2009.
- [43] P. Carré, D. Helbert, and E. Andrès. 3-d fast ridgelet transform. In *ICIP 03*. IEEE, September 2003. Barcelona, Spain.
- [44] Philippe Carré and Eric Andres. Ridgelet transform based on Réveillès discrete lines. *Springer Verlag*, 2301 of *Lecture notes in Computer Science* :417–427, April 2002.
- [45] Philippe Carré and Patrice Denis. Quaternionic wavelet transform for colour images. In *Wavelet Applications in Industrial Processing IV*, volume 6383, Boston, Massachusetts, USA, October 2006.
- [46] Turgay Celik and Tardi Tjahjadi. Multiscale texture classification using dual-tree complex wavelet transform. *Pattern Recogn. Lett.*, 30(3) :331–339, 2009.
- [47] Turgay Celik and Tardi Tjahjadi. Bayesian texture classification and retrieval based on multiscale feature vector. *Pattern Recognition Letters*, 32(2) :159 – 167, 2011.
- [48] W.L. Chan, H. Choi, and R. Baraniuk. Directional hypercomplex wavelets for multidimensionak signal analysis andprocessing. In *Intl Conf. on Image Processing*. IEEE, May 2004.
- [49] W.L. Chan, H.H. Choi, and R.G. Baraniuk. Coherent multiscale image processing using dual-tree quaternion wavelets. *IEEE Transactions on Image Processing*, 17(7) :1069–1082, July 2008.
- [50] V. Chappelier and C. Guillemot. Oriented wavelet transform on a quincunx pyramid for image compression. In *International Conference on Image Processing*, pages I : 81–84, 2005.

- [51] X.C. Chen, Y.H. Wang, T.N. Tan, and L. Guo. Blind image steganalysis based on statistical analysis of empirical matrix. In *Proceedings of 18th International Conference on Pattern Recognition*, volume 3, 2006.
- [52] V. M. Chernov. Discrete orthogonal transforms with data representation in composition algebras. In *Proceedings Scandinavian Conference on Image Analysis*, pages 357–364, 1995.
- [53] H. Choi and R. Baraniuk. Multiscale manifold representation and modeling. *ICASSP '05*, 4(iv) :569–572, March 2005.
- [54] L.D. Cohen. Chemins minimaux et modeles deformables en analyse d’images. *Traitement du Signal*, 20 :225–241, Décembre 2003.
- [55] R. Coifman and S. Donoho. Translation-invariant de-noising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, volume 103 of *Lecture Notes in Statistics*, pages 125–150, New York, 1995. Springer-Verlag.
- [56] Ming Cui, Jiuxiang Hu, Anshuman Razdan, and Peter Wonka. Color-to-gray conversion using iso-map. *The Visual Computer*, 26(11) :1349–1360, 2009.
- [57] J. Darbon, B. Sankur, and H. Maitre. Error correcting code performance for watermark protection. *Security and Watermarking of Multimedia Contents 3*, SPIE Volume 4314 :663–672, August 2001.
- [58] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *Journal of Fourier Analysis and Applications*, 4(3) :247–269, may 1998.
- [59] Franck Davoine. Triangular meshes : A solution to resist to geometric distorsions based watermark-removal softwares. In *10th European Signal Processing Conference EUSIPCO'2000, Tampere, Finland*, pages 1985–1988, September 2000.
- [60] D. Delanay and B. Macq. Generalized 2-d cyclic patterns for secret watermarking generation. In *Proc. ICIP*, volume 2, pages 77–80, September 2000.
- [61] G. Demarcq. *Le signal monogene couleur : théorie et applications*. PhD thesis, Université de La Rochelle, 2010.
- [62] P. Denis. *Quaternions et Algèbres Géométriques, de nouveaux outils pour les images numériques couleur*. PhD thesis, University of Poitiers, 2007.
- [63] Patrice Denis and Philippe Carré. Colour gradient using geometric algebra. In *EUSIPCO2007, 15th European Signal Processing Conference, Poznań, Poland*, September 2007.
- [64] Patrice Denis, Philippe Carré, and Christine Fernandez-Maloigne. Spatial and spectral quaternionic approaches for colour images. *Computer Vision and Image Understanding*, 107 :74–87, July 2007.
- [65] R. Deriche and D. Tschumperlé. Vector-valued image regularization with pdes : A common framework for different applications. In *Pattern Analysis and Machine Intelligence*, volume 27, April 2005.
- [66] E.W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1 :269–271, 1959.
- [67] W. Ding, F. Wu, and S. Li. Lifting-based wavelet transform with directionnally spatial prediction. In *Picture Coding Symposium*, December 2004.

- [68] M. Do. *Directional Multiresolution Image Representations*. PhD thesis, Department of Communication Systems, Swiss Federal Institute of Technology Lausanne, November 2001.
- [69] M.N. Do and M. Vetterli. The contourlet transform : An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12) :2091–2106, December 2005.
- [70] D. Donoho. Wavelet shrinkage and w.v.d. : A 10-minute tour. Technical report, Stanford University, 1992.
- [71] D. Donoho. De-noising by soft-thresholding. *IEEE Trans. on Information Theory*, 41(3) :613–627, 1995.
- [72] D. Donoho and A. Flesia. Digital ridgelet transform based on true ridge functions. In J. Stoeckler and G. V. Welland, editors, *Beyond Wavelets*. Academic Press, 2002.
- [73] D. L. Donoho and X. Huo. Beamlets and multiscale image analysis. *Analysis*, 1 :149–196, 2002.
- [74] D.L. Donoho. Wedgelets : Nearly minimax estimations of edges. *Annals of Statistics*, 27(3) :859–897, 1999.
- [75] D.L. Donoho. Orthonormal ridgelet and linear singularities. *Mathematical Analysis*, 31 :1062–1099, 2000.
- [76] Julia Ebling and Gerik Scheuermann. Clifford convolution and pattern matching on vector fields. In *IEEE Visualization*, pages 193–200, 2003.
- [77] Julia Ebling and Gerik Scheuermann. Clifford fourier transform on vector fields. *IEEE Transactions on Visualization and Computer Graphics*, 11,4 :469–479, July/August 2005.
- [78] P. Elias. List decoding for noisy channels. Technical report, Technical Report 335, Research Laboratory of Electronics, MIT, 1957.
- [79] Todd A. Ell. *Hypercomplex spectral transformation*. PhD thesis, University of Minnesota, 1992.
- [80] Todd A. Ell and Stephen J. Sangwine. Quaternion involutions and anti-involutions. *Computers & Mathematics with Applications*, 53(1) :137 – 143, 2007.
- [81] Todd A. Ell and Stephen John Sangwine. Decomposition of 2d hypercomplex fourier transforms into pairs of fourier transforms. In *Proc. EUSIPCO*, pages 151–154, 2000.
- [82] Todd A. Ell and Stephen John Sangwine. Hypercomplex fourier transform of color images. *IEEE Transactions on Signal Processing*, 16,1 :22–35, January 2007.
- [83] Todd Anthony Ell. *Hypercomplex spectral transformations*. PhD thesis, University of Minnesota, 1992.
- [84] H. Farid. Detecting hidden messages using higher-order statistical models. In *Proceedings of IEEE International Conference on Image processing*, volume 2, 2002.
- [85] Michael Felsberg. *Low-Level Image Processing with the Structure Multivector*. PhD thesis, Kiel University, Germany, March 2002.
- [86] Michael Felsberg and Gerald Sommer. Optimized fast algorithms for the quaternionic fourier transform. *8th Int. Conference on Computer Analysis of Images and Patterns*, 1689 :209–216, 1999.
- [87] Michael Felsberg and Gerald Sommer. The monogenic signal. *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001.

- [88] J. Fessler and B. Sutton. A min-max approach to the multidimensional nonuniform fft : Application totomographic image reconstruction. To be Published in IEEE Trans. on Signal Processing, 2001.
- [89] D. J. Fleet and D. J. Hegger. Embedding invisible information in color image. In *IEEE International Conference on Image Processing, Santa Barbara*, volume 1, pages 532–535, october 1997.
- [90] L. Florack, B. ter Haar Romeny, J. Koenderink, and M. Viergever. General intensity transformations and second order invariants.
- [91] G. D. Forneyi. *Concatenated Codes*. The MIT Press, 1966.
- [92] K. Fourmont. Non-equispaced fast fourier transforms with applications to tomography. Technical report, Math Department at the University of Muenster, [http ://www.math.uni-muenster.de](http://www.math.uni-muenster.de), May 2002.
- [93] Wolfgang Funk. Image watermarking in the fourier domain based on global features of concentric ring areas. In *Security and Watermarking of Multimedia Contents IV, Proceedings of SPIE*, volume 4675, pages 596–599, 2002.
- [94] Jean Gallier. *Geometric methods and applications : for computer science and engineering*. Springer-Verlag, London, UK, 2000.
- [95] Philippe-Henri Gosselin. Quaternions et images couleur. Master’s thesis, Université de Poitiers, 2002.
- [96] A. Gouze, M. Antonini, and M. Barlaud. Quincunx lifting scheme for lossy image compression. In *International Conference on Image Processing*, volume 1, pages 665–668, 2000.
- [97] S. Grassin. *Analyse TF des signaux discrets : application aux images*. PhD thesis, Université de Rennes, 1997.
- [98] V. Guruswami and M. Sudan. Improved decoding of reed-solomon codes and algebraic geometry codes. *IEEE Trans. Inf. Theory*, Volume 45 :1755–1764, 1999.
- [99] S. L. Hahn. Hilbert transforms in signal processing. *Artech House, Boston, London*, 1996.
- [100] R. W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2) :147–160, 1950.
- [101] J.J. Harmsen and W.A. Pearlman. Steganalysis of additive noise modelable information hiding. In *Proceedings of the SPIE, Security, Steganography, and Watermarking of Multimedia Contents V*, 2003.
- [102] C. Harris and M.J. Stephens. A combined corner and edge detector.
- [103] F. Hartung, J. K. Su, and B. Girod. Spread spectrum watermarking : Malicious attacks and counterattacks. In *IS&T/SPIE Conference on Security and Watermarking of Multimedia Contents, San Jose, California, USA*, volume 3657, pages 147–158, January 1999.
- [104] Serkan Hatipoglu, Sanjit K. Mitra, and Nick Kingsbury. Texture classification using dual-tree complex wavelet transform. volume 1, pages 344 –347 vol.1, jul. 1999.
- [105] David Helbert. *système de fusion partagé : d’un prétraitement des informations à une reconstruction dynamique*. Thèse de doctorat, Université de Poitiers, 2005.
- [106] David Helbert, Philippe Carré, and Éric Andrès. 3-D Discrete Analytical Ridgelet Transform. *IEEE Transactions on Image Processing*, 15(12) :pages 3701–3714, 2006.

- [107] Stefan Held, Martin Storath, Peter Massopust, and Brigitte Forster. Steerable wavelet frames based on the riesz transform. *IEEE Transactions on Image Processing*, 19(3) :653–667, 2010.
- [108] J. R. Hernandez and F. Perez-Gonzalez. The impact of channel coding on the performance of spatial watermarking for copyright protection. In *IEEE ICASSP, Seattle, WA*, volume 5, pages 2973–2976, May 1998.
- [109] Paul R. Hill, David R. Bull, and Cedric Nishan Canagarajah. Rotationally invariant texture features using the dual-tree complex wavelet transform. In *IEEE International Conference on Image Processing*, 2000.
- [110] Eckhard Hitzer. Clifford fourier and clifford wavelet transforms. *Technology*, (August), 2009.
- [111] T. Holotyak, J. Fridrich, and S. Voloshynovskiy. Blind statistical steganalysis of additive steganography using wavelet higher order statistics. In Springer, editor, *Proceedings of ninth IFIP TC-6 TC-11 Conference on Communications and Multimedia Security, Lecture Notes in Computer Science*, volume 3677, 2005.
- [112] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian. A real-time algorithm for signal analysis with the help of the wavelet transform. In J.M. Combes, A. Grossmann, and P. Tchamitchian, editors, *Wavelet, Time-Frequency Methods and Phase Space*, pages 289–297. Springer-Verlag, Berlin, 1989.
- [113] Ebroul Izquierdo. Using invariant image features for synchronization in spread spectrum image watermarking. In *EURASIP Journal on Applied Signal Processing*, volume 4, pages 410–417, 2002.
- [114] Laurent Jacques, Laurent Duval, Caroline Chaux, and Gabriel Peyré. A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity. *Computer Vision and Pattern Recognition*, 2011.
- [115] M. Jansen, H. Choi, S. Lavu, and R.G. Baraniuk. Multiscale image processing using normal triangulated meshes. In *International Conference on Image Processing*, pages II : 229–232, 2001.
- [116] J. Justesen and T. Høholdt. A course in error-correcting codes. *European Mathematical Society*, 2000.
- [117] N G Kingsbury. The dual-tree complex wavelet transform : a new efficient tool for image restoration and enhancement. In *Proc. European Signal Processing Conference, EUSIPCO 98*, 1998.
- [118] Nick Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis*, 10(3) :234 – 253, 2001.
- [119] L. Kitchen and A. Rosenfeld. Gray level corner detection. *Pattern Recognition Letters*, pages 95–102, 1982.
- [120] B. De Knock. *Multidimensional distributions and generalized Hilbert transforms in Clifford analysis*. PhD thesis, University Gent, 2008.
- [121] I. Koren and A. Laine. A discrete dyadic wavelet transform for multidimensional feature analysis. In M. Akay, editor, *Time-Frequency and Wavelet transforms in biomedical engineering*. IEEE Press, New York, 1997.
- [122] P. KOVESI. Image features from phase congruency. Technical Report 4, University of Western Australia, 1995.

- [123] M. Kutter. Watermarking resisting to translation, rotation, and scaling. In *SPIE Conference on Multimedia Systems and Applications, Boston, Massachusetts*, volume 3528, pages 423–431, november 1998.
- [124] M. Kutter, S.K. Bhattacharjee, and T. Ebrahimi. Towards second generation watermarking schemes. In *6th International Conference on Image Processing ICIP'99, Kobe, Japon*, volume 1, pages 320–323, october 1999.
- [125] M. Kutter, F. Jordan, and F. Bossen. Digital signature of color images using amplitude modulation. In *Processings of SPIE storage and retrieval for image and video databases*, volume 3022, pages 518–526, february 1997.
- [126] A. Laine and J. Fan. Texture classification by wavelet packet signatures. *IEEE Trans. on PAMI*, 15(11) :1186–1191, November 1993.
- [127] Kieran G. Larkin, Donald J. Bone, and Michael A. Oldfield. Natural demodulation of two-dimensional fringe patterns. ii. stationary phase analysis of the spiral phase quadrature transform. *J. Opt. Soc. Am. A*, 18 :1871–1881, 2001.
- [128] B. Le Guen, S. Pateux, and J. Weiss. Modèle énergétique pour la représentation d’images par ondelettes déformées. In *CORESA 2006 Proceedings*, 2006.
- [129] B. Le Guen, S. Pateux, and J. Weiss. Non-geometric energy formulation for adaptive image compression. In *International Conference on Image Processing*, pages IV : 161–164, 2007.
- [130] E. Le Pennec. *Bandelettes et représentation géométrique des images*. PhD thesis, École Polytechnique, Décembre 2002.
- [131] E. Le Pennec and S. Mallat. Image compression with geometrical wavelets. In *International Conference on Image Processing*, pages Vol I : 661–664, 2000.
- [132] G. Lebrun. *Ondelettes géométriques adaptatives : vers une utilisation de la distance géodésique*. PhD thesis, University of Poitiers, 2009.
- [133] G. Lebrun, P. Carré, and S. Pateux. Two-layers bandelets : Algorithmic interpretation. In *Wavelet Applications in Industrial Processing V. SPIE*, October 2007.
- [134] C.S. Lee, C.K. Lee, and K.Y. Yoo. New lifting based structure for undecimated wavelet transform. *Electronic Letters*, 36(22) :1894–1895, October 2000.
- [135] W.N. Lie and G.S. Lin. A feature-based classification technique for blind image steganalysis. *IEEE Transactions on Multimedia*, 7(6) :1007–1020, 2005.
- [136] C.-Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. Man Lui. Rotation, scale, and translation resilient watermarking for images. In *IEEE Transactions on Image Processing*, volume 10, pages 767–782, May 2001.
- [137] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Springer, 1994.
- [138] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60 :91–110, November 2004.
- [139] Y. Lu and M.N. Do. Crisp-contourlet : a critically sampled directional multiresolution image representation. In *Proc. SPIE Conf. on Wavelet Applications in Signal and Image Processing*, August 2003.

-
- [140] Y. M. Lu and M. N. Do. Multidimensional directional filter banks and surfacelets. *IEEE Transactions on Image Processing*, 04 :918–931, April 07.
- [141] A. Maalouf. *Coop  ration des ondelettes et des   quations aux d  riv  es partielles pour le traitement d’images multispectrales*. PhD thesis, University of Poitiers, 2008.
- [142] A. Maalouf, P. Carr  , B. Augereau, and C. Fernandez-Maloigne. Cooperation of the Partial Differential Equation Methods and the Wavelet Transform for the Segmentation of Multivalued Images. *Signal Processing - Image Communication*, Volume 23(Number 1) :pages 14–30, January 2008.
- [143] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Bandelet-Based Anisotropic Diffusion. In *IEEE International Conference on Image Processing*, pages 23–30, San Antonio,   tats-Unis, 2007.
- [144] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Clouds Removal from Remotely Sensed Images by Using a Bandelet–Based Reconstruction Technique. In *Proceedings of SPIE Optics East Conference on Wavelet Applications in Industrial Processing V*, pages 23–30, Boston,   tats-Unis, 2007.
- [145] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Foveal wavelet-based color active contour. In *IEEE International Conference on Image Processing*, pages 23–30, San Antonio,   tats-Unis, 2007.
- [146] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Regularization of Multivalued Images by Means of a Wavelet-Based Partial Differential Equation. In *EUSIPCO2007, 15th European Signal Processing Conference*, pages 23–30, Poznan, Pologne, 2007.
- [147] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. A Bandelet-Based Inpainting Technique for Clouds Removal From Remotely Sensed Images. *IEEE Transactions on Geosciences and Remote Sensing*, pages 77–85, December 2008.
- [148] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Bandelet-Based Video Inpainting. In *16th European Signal Processing Conference*, pages 23–30, Suisse, August 2008.
- [149] Aldo Maalouf, Philippe Carr  , B. Augereau, and Christine Fernandez-Maloigne. Inpainting using geometrical grouplets. In *EUSIPCO2008, 16th European Signal Processing Conference, Lausanne, Switzerland*, pages 15–20, Suisse, August 2008.
- [150] J F A Magarey and N G Kingsbury. Motion estimation using a complex-valued wavelet transform. *IEEE Transaction on Signal Processing*, 46(4) :1069–84, 1998.
- [151] S. Mallat. A theory for multiresolution signal decomposition : the wavelet transform. *IEEE Trans. on PAMI*, 11(7) :674–693, 1989.
- [152] S. Mallat. *A wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.
- [153] S. Mallat. Geometrical grouplets. *Appl. Comput. Harmon. Anal.*, March 2008.
- [154] Bahri Mawardi and Eckhard MS Hitzer. Clifford fourier transformation and uncertainty principle for the clifford geometric algebra $cl_{3,0}$. *Advances in Applied Clifford Algebras*, 16,1 :41–61, February 2006.
-

- [155] Andrew McCabe, Terry Caelli, Geoff West, and Adam Reeves. Encoding and processing spatio-chromatic image information using complex fourier transform methods. Technical Report 4, School of Computing, Curtin University of Technology, Australia, 1997.
- [156] J. McNamara. A fast nearest-neighbor algorithm based on a principal axis search tree. *Pattern Analysis and Machine Intelligence*, 23(9) :964–976, September 2001.
- [157] Y. Meyer. *Ondelettes et opérateurs*, volume I. Hermann, Paris, 1990.
- [158] M. K. Mhcah and R. Venkatesan. Blind image watermarking via derivation and quantization of robust semi-global statistics. In *Proc. IEEE ICASSP, Florida, FL*, June 2002.
- [159] M. K. Mhcah, R. Venkatesan, and M. Kesal. Cryptanalysis of discrete-sequence spread spectrum watermarks. In *Proceedings of 5th Information Hiding Workshop, Holland*, October 2002.
- [160] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. In *Proceedings of the IEEE Computer Society International Conference on Computer Vision*, pages 525–531, 2001.
- [161] M. Miller and N. Kingsbury. Statistical image modelling using interscale phase relationships of complex wavelet coefficients. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 789–792, 2006.
- [162] Francois Mourougaya. *Generic Texture description of retro-engineering on a printing chain*. Thèse de doctorat, Université de Poitiers, 2003.
- [163] C.E. Moxey, Stephen John Sangwine, and Todd A. Ell. Hypercomplex correlation techniques for vector images. *IEEE Transactions on Signal Processing*, 51 :1941–1953, July 2003.
- [164] F. Natterer. Numerical methods in tomography. *Acta Numerica*, 1999.
- [165] A. Nikolaidis and I. Pitas. A region-based technique for chaotic image watermarking. In *10th European Signal Processing Conference EUSIPCO'2000, Tampere, Finland*, pages 1681–1684, September 2000.
- [166] A. Nikolaidis and I. Pitas. Region-based image watermarking. In *IEEE Transactions on Image Processing*, volume 10, pages 1726–1740, November 2001.
- [167] S. Olhede and G. Metikas. The hyperanalytical wavelet transform. May 2006.
- [168] Sofia C. Olhede and Georgios Metikas. The monogenic wavelet transform. *IEEE Transactions on Signal Processing*, 57 :3426–3441, 2009.
- [169] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. In *European Conference on Computer Vision 2006*, pages IV : 568–580, 2006.
- [170] A. Parisis, P. Carré, C. Fernandez-Maloigne, and N. Laurent. Color image watermarking with adaptive strength of insertion. *IEEE ICASSP*, 3 :85–9, May 2004.
- [171] Alice Parisis. *Tatouage d'images couleur*. Thèse de doctorat, Université de Poitiers, 2004.
- [172] Alice Parisis, Philippe Carré, Christine Fernandez, and Nathalie Laurent. Color image watermarking with adaptive strength of insertion. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Montréal Québec*, volume 3, pages 85–88, May 2004.
- [173] Alice Parisis, Philippe Carré, and Nathalie Laurent. Procédé et dispositif de tatouage d'une marque dans une image. Technical report, Brevet INPI 0313171, 2003.

- [174] alice Parisis, philippe Carré, and Nathalie Laurent. Technique de recalage d'une image après rotation ou changement d'échelle.application particulière aux images tatouées. Technical report, Brevet INPI 03O4592, 2003.
- [175] S. Pei, J. Ding, and J. Chang. Efficient implementation of quaternion fourier transform, convolution, and correlation by 2-d complex fft. *IEEE Trans. Signal Processing*, 49(11) :2783–2797, november 2001.
- [176] Soo-Chang Pei, Ja-Han Chang, and Jian-Jiun Ding. Commutative reduced biquaternions and their fourier transform for signal and image processing applications. *IEEE Trans. Signal Processing*, 52,7 :2012–2031, November 2004.
- [177] Soo-Chang Pei, Jian-Jiun Ding, and Ja-Han Chang. Efficient implementation of quaternion fourier transform, convolution, and correlation by 2-d complex fft. *IEEE Trans. Signal Processing*, 49 :2783–2797, November 2001.
- [178] S. Pereira and T. Pun. Fast robust template matching for affine resistant image watermarks. In *Information Hiding*, pages 199–210, 2000.
- [179] F. Petit. *Traitement et analyse d'images couleur sous-marines : modèles physiques et représentation quaternionique*. PhD thesis, University of Poitiers, 2010.
- [180] Frédéric Petit, Anne-Sophie Capelle-Laizé, and Philippe Carré. Underwater image enhancement by attenuation inversion with quaternions. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, Tapei, Taïwan, Province De Chine, 2009.
- [181] Frédéric Petit, Anne-Sophie Capelle-Laizé, and Philippe Carré. Hue-Based Quaternionic Criterion for Focused-Color Extraction. In *Proceedings of International Conference Image Processing (ICIP)*, Hong-Kong, Hong-Kong, 2010.
- [182] G. Peyré. *Géométrie multi-échelle pour les images et les textures*. PhD thesis, École Polytechnique, Décembre 2005.
- [183] G. Peyré. Texture synthesis with grouplets. *Pattern Analysis and Machine Intelligence*, March 2009.
- [184] G. Peyré. A review of adaptive image representations. *IEEE Journal of Selected Topics in Signal Processing*, 5(5) :896–911, 2011.
- [185] Gabriel Peyré and Laurent D. Cohen. Geodesic remeshing using front propagation. *Int. J. Comput. Vision*, 69 :145–156, August 2006.
- [186] D. Potts and G. Steidl. A new linogram algorithm for computerized tomography. *IMA Journal of Numerical Analysis*, 21 :769–782, 2001.
- [187] D. Potts, G. Steidl, and M. Tasche. Fast fourier transforms for nonequispaced data : a tutorial. In *Modern Sampling Theory : Mathematics ans Applications*. 2001.
- [188] Honeyman P Provos, N. Hide and seek : An introduction to steganography. *IEEE Security and Privacy Magazine*, 1 :32–44, 2003.
- [189] W. Puech. Safe transfer of image based on color transformation for watermarking. In *4th COST 276 Workshop on Transmitting Processing and Watermarking Multimedia Contents, Bordeaux, France*, pages 1–6, march 2003.

- [190] Imtnan-Ul-Haque Qazi, Olivier Alata, Jean-Christophe Burie, Mohamed Abadi, Ahmed Moussa, and Christine Fernandez-Maloigne. Parametric models of linear prediction error distribution for color texture and satellite image segmentation. *Computer Vision and Image Understanding*, 115(8) :1245–1262, 2011.
- [191] Imtnan-Ul-Haque Qazi, Olivier Alata, Jean-Christophe Burie, Ahmed Moussa, and Christine Fernandez-Maloigne. Choice of a pertinent color space for color texture characterization using parametric spectral analysis. *Pattern Recognition*, 44(1) :16–31, 2011.
- [192] Johann Radon. Über die bestimmung von funktionen durch ihre integral durch ihre integralwerte langs gewisser mannigfaltigkeiten. *Berichte Sachsische Akademie der Wissenschaften*, 69 :262–279, April 1917.
- [193] I. S. Reed and G. Solomon. Polynomial codes over certain finite fields. *SIAM Journal of Applied Math.*, 8 :300–304, 1960.
- [194] Jean-Pierre Reveillès. Géométrie discrète, calcul en nombres entiers et algorithmique. Habilitation, Université Louis Pasteur de Strasbourg, 1991.
- [195] J. Roerdink and M. Westenberg. Data-parallel tomographic reconstruction : a comparison of filtered backprojection and direct fourier reconstruction. *Parallel Computing*, 24 :2129–2144, 1998.
- [196] S. Rousseau. *Algèbre géométrique pour l’imagerie multispectrale*. PhD thesis, University of Poitiers, 2011.
- [197] Sylvain Rousseau, David Helbert, Philippe Carré, and Jacques Blanc-Talon. Metric Tensor for Multicomponent Edge Detection. In *ICIP*, page 2, 2010.
- [198] J.J.K. Ruanaidh and T. Pun. Rotation, scale and translation invariant spread spectrum digital image watermarking. In *Signal Processing*, volume 66, pages 303–317, 1998.
- [199] S. Sangwine and T. Ell. Hypercomplex fourier transforms of colour images. In *Inter. Conference on Image Processing*, volume 1, pages 137–140, Thessaloniki, Greece, 2001. IEEE.
- [200] S.J. Sangwine. Colour in image processing. *Electronics and Communication Engineering Journal*, 12(5) :211–219, October 2000.
- [201] Stephen John Sangwine. Fourier transforms of colour images using quaternion, or hypercomplex, numbers. *Electronics Letters*, 32, 21 :1979–1980, October 1996.
- [202] Stephen John Sangwine. The discrete quaternion fourier transform. In *Proceedings of the 6th International Conference on Image processing and its applications*, volume 2, 14-17, pages 790–793, July 1997.
- [203] Stephen John Sangwine. Colour image edge detector based on quaternion convolution. *Electronics Letters*, 34, 10 :969–971, May 1998.
- [204] Stephen John Sangwine and Todd A. Ell. Hypercomplex auto- and cross-correlation of color images. In *IEEE International Conference on Image Processing (ICIP)*, volume 4, pages 319–322, 1999.
- [205] Stephen John Sangwine and Todd A. Ell. Hypercomplex fourier transforms of colour images. In *Proceedings ICIP*, pages 137–140, 2001.
- [206] J. Sankaranarayanan, H. Samet, and A. Varshney. A fast k-neighborhood algorithm large point clouds. In *Point-Based Graphics*, July 2006.

- [207] M. Schlauweg, D. Prufrock, and E. Muller. Soft feature-based watermark decoding with insertion/deletion correction. In *Proceedings of the 9th international conference on Information hiding*, pages 237–251, 2007.
- [208] C. Schmid. *Appariement d’images par invariants locaux de niveaux de gris*. PhD thesis, 1996.
- [209] P. SCHRÖDER and W. SWELDENS. Spherical wavelets : Efficiently representing functions on the sphere. *Computer Graphics Proceedings (SIGGRAPH 95)*, pages 161–172, 1995.
- [210] I. W. Selesnick. Hilbert transform pairs of wavelet bases. *IEEE Signal Processing Letters*, 8(1) :170–173, 2001.
- [211] I. W. Selesnick. The design of approximate hilbert transform pairs of wavelet bases. *IEEE Trans. on Signal Processing*, 50(5) :1144–1152, 2002.
- [212] Ivan W. Selesnick, Richard G. Baraniuk, and Nick G. Kingsbury. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine* [123] November, 2005.
- [213] J.A. Sethian. Fast marching methods. *SIAM Review*, 41(2) :199–235, 1999.
- [214] Y.Q. Shi, G.R. Xuan, and D.K. Zou. Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction- error image, and neural network. In *Proceedings of IEEE International Conference on Multimedia and Expo*, 2005.
- [215] Y. Shkolnisky, A. Averbuch, and D. Donoho. 2d fourier based discrete radon transform. Technical report, Dpt of Statistics, Stanford University, [http ://www.math.tau.ac.il/ amir1/publications.html](http://www.math.tau.ac.il/~amir1/publications.html).
- [216] A. Silverberg, J. Staddon, and J. Walker. Efficient traitor tracing algorithms using list decoding. *ASIACRYPT*, pages 175–192, 2001.
- [217] Bernard Sklar. Digital communications : Fundamentals and applications. *Prentice Hall*, 2001.
- [218] Nir A. Sochen, Ron Kimmel, and Ravi Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7(3) :310–318, 1998.
- [219] V. Solachidis and I. Pitas. Circularly symmetric watermark embedding in 2d dft domain. In *IEEE International Conference on Acoustics, Speech and Signal Processing, Phoenix, USA*, March 1999.
- [220] V. Solachidis and I. Pitas. Self-similar ring shaped watermark embedding in 2d-dft domain. In *10th European Signal Processing Conference, EUSIPCO’00, Tampere, Finlande*, pages 1977–1980, september 2000.
- [221] R. Soulard. Quaternions et algèbres géométriques pour le traitement d’images. Master’s thesis, University of Poitiers, 2009.
- [222] R. Soulard. *Algèbre de Clifford et transformée en ondelettes*. PhD thesis, University of Poitiers, 2011.
- [223] R. Soulard and P. Carré. Transformées en ondelettes analytiques pour les images couleur. Presentation GDR ISIS, journée Transformées multiresolution géométrique, Date : 2011-04-01.
- [224] Raphael Soulard and Philippe Carré. Quaternionic wavelets for image coding. In *2010 European Signal Processing Conference (EUSIPCO-2010)*, Aalborg, Denmark, 8 2010.
- [225] Raphaël Soulard and Philippe Carré. Quaternionic wavelets for texture classification. In *Proceedings of the 35th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, pages 4134–4137, Dallas (TX, USA), March 2010.

- [226] Raphael Souillard and Philippe Carré. Color Extension of monogenic Wavelets with geometric algebra : application to color image denoising. In *ICCA9 Proceedings*, page 179, Weimar, Allemagne, July 2011. 10 pages.
- [227] Raphaël Souillard and Philippe Carré. Color monogenic wavelets for image analysis. In *ICIP*, Brussels, september 2011.
- [228] Raphael Souillard and Philippe Carré. Quaternionic wavelets for texture classification. *Pattern Recognition Letters*, 32(13) :1669–1678, July 2011.
- [229] Adrian Spinei, Denis Pellerin, and Jeanny Hérault. Spatiotemporal energy-based method for velocity estimation. *Signal Processing*, 65(3) :347 – 362, 1998.
- [230] A. Spira, R. Kimmel, and Nir A. Sochen. A short- time beltrami kernel for smoothing images and manifolds. *IEEE Transactions on Image Processing*, 16 :1628–1636, 2007.
- [231] S. Stankovic, I. Djurovic, and I. Pitas. Watermarking in the spacespatial–frequency domain using two–dimensional radon–wigner distribution. In *IEEE Transactions on Image Processing*, volume 10, pages 650–658, April 2001.
- [232] J.-L. Starck, E. Candès, and D. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6) :670–684, November 2002.
- [233] M. Sudan. Decoding of reed-solomon codes beyond the error correction bound. *Journal of complexity*, vol 12 :pp. 180–193, 1997.
- [234] K. Sullivan, U. Madhow, S. Chandrasekaran, and B.S. Manjunath. Steganalysis of spread spectrum data hiding exploiting cover memory. In 5681, editor, *Proceedings of the SPIE, Security, Steganography, and Watermarking of Multimedia Contents VII*, 2005.
- [235] W. SWELDENS. Wavelets and the lifting sheme : A 5 minute tour. *Z. Angew. Math. Mech.*, 76 :44–41, 1996.
- [236] W. Sweldens and P. Schröder. Building your own wavelets at home. In *Wavelets in Computer Graphics*, pages 15–87, 1996.
- [237] J.B. Tenenbaum, V. de Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500) :2319–2323, December 2000.
- [238] N. Terzija, M. Repges, K. Luck, and W. Geisselhardt. Digital image watermarking using discrete wavelet transform : Performance comparison of error correcting codes. *Visualization, Imaging, and Image Processing*, September 2002.
- [239] P. Toft. *The Radon Transform : theory and implementation*. PhD thesis, Dpt of Mathematical Modeling, Technical University of Denmark, 1996.
- [240] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. of International Conference on Computer Vision*, pages 839–846, 1998.
- [241] Alain Trémeau and Damien Muselet. Most recent trends in color image watermarking. *Journal of Imaging Science and Technology*, 53(1), 2009.
- [242] Michael Unser, Daniel Sage, and Dimitri Van De Ville. Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform. *Trans. Img. Proc.*, 18 :2402–2418, November 2009.

- [243] D. Van De Ville and M. Unser. Complex wavelet bases, steerability, and the Marr-like pyramid. *IEEE Transactions on Image Processing*, 17(11) :2063–2080, November 2008.
- [244] B. Verma, S. Jain, DP Agarwal, and A. Phadikar. A new color image watermarking scheme. *Infocomp, Journal of computer science*, 5(2) :37–42, 2006.
- [245] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice Hall PTR, New Jersey, 1995.
- [246] An Vo, Soontorn. Oraintara, and Truong T. Nguyen. Using phase and magnitude information of the complex directional filter bank for texture image retrieval. pages IV : 61–64, 2007.
- [247] An Vo, Soontorn Oraintara, and Truong T. Nguyen. Statistical image modeling using distribution of relative phase in the complex wavelet domain. In *Eusipco 2008, 16th European Signal Processing Conference, August 25-29, 2008, Lausanne, Switzerland*, 08 2008.
- [248] M. B. Wakin, J. K. Romberg, H. Choi, and R. G. Baraniuk. Wavelet-domain approximation and compression of piecewise smooth images. *IEEE Transactions on Image Processing*, 15(5), May 2006.
- [249] J.Z. Wang and G. Wiederhold. Wavemark : Digital image watermarking using daubechies’ wavelets and error correcting coding. In *Proceedings of SPIE*, volume 3528, pages 432–439, 1998.
- [250] Wang Xiao-Hui, Zhou Yue, Wang Yong-Gang, and Zhu WeiWei. Color texture segmentation based on quaternion-gabor features. In Josef Kittler Martinez-Trinidad, Jesus Carrasco Ochoa, and Josef Kittler, editors, *Progress in Pattern Recognition, Image Analysis and Applications*, volume 4225 of *Lecture Notes in Computer Science*, pages 345–353. Springer Berlin / Heidelberg, 2006.
- [251] Lexing Ying, Laurent Demanet, and Emmanuel C. Fast discrete curvelet transforms. *SIAM Multiscale Model. Simul.*, (5) :861–899, 2005.
- [252] Ming Zhang and B.K. Gunturk. Multiresolution bilateral filtering for image denoising. *Image Processing, IEEE Transactions on*, 17(12) :2324 –2333, dec. 2008.
- [253] Jun Zhou, Yi Xu, and Xiaokang Yang. Quaternion wavelet phase based stereo matching for uncalibrated images. *Pattern Recogn. Lett.*, 28(12) :1509–1522, 2007.
- [254] Y. M. Zhu, F. Peyrin, and R. Goutte. Equivalence between the two-dimensional real and analytic signal wigner distributions. *IEEE Transaction on ASSP*, 37 :1631–1634, 1989.