



HAL
open science

Méthodes fréquentielles pour la reconnaissance d'images couleur : une approche par les algèbres de Clifford

José Mennesson

► **To cite this version:**

José Mennesson. Méthodes fréquentielles pour la reconnaissance d'images couleur : une approche par les algèbres de Clifford. Mathématiques générales [math.GM]. Université de La Rochelle, 2011. Français. NNT : 2011LAROS337 . tel-00703445

HAL Id: tel-00703445

<https://theses.hal.science/tel-00703445v1>

Submitted on 1 Jun 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.



UNIVERSITÉ DE LA ROCHELLE

*ÉCOLE DOCTORALE SCIENCES ET
INGÉNIERIE POUR L'INFORMATION*

Laboratoire Mathématiques, Image et Applications



THÈSE présentée par :

José MENNESSON

Le 18 Novembre 2011

pour obtenir le grade de : **Docteur de l'université de La Rochelle**
Mention : **Automatique, Image et Signal**

**Méthodes fréquentielles pour la reconnaissance
d'images couleur -
Une approche par les algèbres de Clifford**

JURY :

Johel MITÉРАН

Stephen J. SANGWINE

Jesús ANGULO

Ludovic MACAIRE

Laurent MASCARILLA

Christophe SAINT-JEAN

Professeur, Université de Bourgogne, Rapporteur

Professeur, Université d'Essex, Rapporteur

Chargé de Recherche, École des MINES ParisTech, Examineur

Professeur, Université de Lille, Examineur

Maître de conférences HDR, Université de La Rochelle,
Directeur de thèse

Maître de conférences, Université de La Rochelle, Co-
directeur de thèse

Remerciements

Tout d'abord, j'aimerais remercier Laurent Mascarilla d'avoir accepté d'être mon directeur de thèse et de m'avoir guidé durant ces trois dernières années. Ses conseils sages et éclairés m'ont permis de réaliser ma thèse dans de bonnes conditions. Je remercie également très vivement Christophe Saint-Jean qui m'a apporté son aide lorsque j'en avais le plus besoin. Tout cela, malgré son emploi du temps chargé. Il a co-dirigé cette thèse avec patience et a su me transmettre des connaissances aussi bien méthodologiques que pratiques.

Je tiens à remercier très sincèrement Johel Mitéran, professeur à l'Université de Bourgogne et Stephen J. Sangwine, Professeur à l'Université d'Essex, d'avoir accepté d'être les rapporteurs de ma thèse. Je les remercie grandement pour leur lecture approfondie de mon manuscrit, leurs diverses remarques qui m'ont permis de l'améliorer, ainsi que pour leur participation au jury.

Je veux également remercier Jesús Angulo, chargé de recherche à l'école des MINES ParisTech et Ludovic Macaire, Professeur à l'Université de Lille pour l'intérêt qu'ils ont porté à l'égard de mon travail et d'avoir accepté de participer à mon jury.

Je remercie aussi Michel Berthier, Directeur du laboratoire MIA, pour l'accueil au sein de son laboratoire ainsi que pour tous les conseils qu'il a pu me donner. Malgré la dure tâche que j'avais à accomplir, l'environnement agréable et chaleureux de ce laboratoire m'a permis de travailler dans les meilleures conditions possibles. Je tiens par ailleurs à remercier le personnel administratif de ce laboratoire pour leur réactivité et leur gentillesse.

Mes remerciements vont aussi à toutes celles et ceux qui m'ont soutenu et qui ont cru en moi durant ces trois années de labeur. Mes pensées vont plus particulièrement à Sloven qui n'a cessé de me donner des conseils et de me soutenir moralement sans se ménager. Pour cela, je lui exprime ma profonde gratitude. J'aimerais aussi remercier mes collègues de bureau qui m'ont supporté pendant toutes ces années : Charles pour les discussions alambiquées que nous avons pu avoir, Ghina pour sa bonne humeur inaltérable et Agathe pour ses conseils avisés. Je remercie aussi mes collègues enseignants qui m'ont fait confiance et ont cru en moi comme Renaud Péteri, Bertrand Vachon, Michel Ménard, et tous les autres qui se reconnaîtront.

Je ne peux oublier tous ceux qui ont contribué de près ou de loin à cette thèse comme Thomas, Guillaume, Romain, Nadeem, Salman, tous mes amis enseignants, doctorants et étudiants.

Merci aussi à Janick qui est resté depuis toutes ces années un de mes meilleurs amis et qui m'a toujours soutenu durant cette épreuve.

D'un point de vue plus personnel, je tiens à remercier toute ma famille qui m'a apporté tout le soutien dont j'avais besoin même si la distance qui nous sépare est grande. Plus particulièrement, je remercie ma mère qui m'a toujours encouragé et soutenu dans les choix que j'ai fait. Pour cela, je lui dédie cette thèse qui n'aurait jamais existé sans elle.

Table des matières

Remerciements	3
Notations	9
Introduction générale	11
1 Les transformées de Fourier : un état de l'art	15
1.1 Les transformées de Fourier pour les images en niveaux de gris	18
1.1.1 La transformée de Fourier complexe	19
1.1.1.1 Définition	19
1.1.1.2 Propriétés	20
1.1.1.3 Interprétation	22
1.1.2 La transformée de Fourier-Mellin classique et sa version analytique	25
1.1.2.1 Définitions	25
1.1.2.2 Propriétés	28
1.1.3 La transformée de Fourier abstraite	29
1.1.4 La M_2 -transformée	31
1.1.4.1 Définition	31
1.1.4.2 Propriétés	33
1.1.5 La transformée de Fourier quaternionique de Bülow [Bülow 99]	34
1.1.5.1 Définition	35
1.1.5.2 Propriétés	36
1.1.6 Discussion	37
1.2 Les transformées de Fourier pour les images couleur	38
1.2.1 Les généralisations de la transformée de Fourier	38
1.2.2 La transformée de Fourier quaternionique de Sangwine et Ell [Sangwine 00]	40
1.2.2.1 Définitions	40
1.2.2.2 Interprétation en termes de rotation	42
1.2.2.3 Propriétés de la transformée de Fourier quaternionique	43
1.2.2.4 Choix de μ_1 et construction de la base $\{\mu_1, \mu_2, \mu_3\}$	44
1.2.2.5 Algorithme de calcul	46
1.2.2.6 Visualisation et interprétation générale	47
1.2.3 La transformée de Fourier Clifford couleur de Batard <i>et al.</i> [Batard 10]	49
1.2.3.1 Définitions	49
1.2.3.2 Interprétation en termes de rotation	53
1.2.3.3 Propriétés de la transformée de Fourier Clifford couleur	54
1.2.3.4 Construction de B	55
1.2.3.5 Construction de la base $\{\mathbf{c}, \mathbf{e}_4, \mathbf{v}, \mathbf{w}\}$	56
1.2.3.6 Algorithme de calcul	57
1.2.3.7 Visualisation et interprétation générale	59

1.2.4	Discussion	61
1.3	Conclusion	63
2	Les descripteurs de Fourier	65
2.1	Propriétés générales des descripteurs	68
2.2	Les descripteurs de Fourier pour les images binaires	70
2.2.1	Les descripteurs de Fourier 1D usuels	71
2.2.1.1	Discussion	73
2.3	Les descripteurs de Fourier pour les images en niveaux de gris	74
2.3.1	Les descripteurs de Fourier génériques	74
2.3.1.1	Représentations polaires et transformée de Fourier	75
2.3.1.2	Obtention des invariances	76
2.3.1.3	Discussion	77
2.3.2	Les descripteurs de Fourier-Mellin	78
2.3.2.1	Les descripteurs de Fourier-Mellin classiques	78
2.3.2.2	Les descripteurs de Fourier-Mellin Analytiques	79
2.3.2.3	Discussion	80
2.3.3	Les descripteurs de Fourier généralisés	80
2.3.3.1	Définitions	80
2.3.3.2	Méthode de calcul	81
2.3.3.3	Discussion	82
2.3.4	Comparaison des descripteurs de Fourier	82
2.4	Les descripteurs de Fourier pour les images couleur	83
2.4.1	Méthode 1D utilisant des signatures dépendantes de la couleur	83
2.4.2	Méthode marginale 2D	84
2.4.3	Méthode par projection de l'image sur un seul plan couleur	85
2.4.4	Les moments de Fourier-Mellin quaternioniques	86
2.4.5	De nouveaux descripteurs de Fourier couleur : GCFD1 et GCFD2	87
2.4.6	Un nouvel ensemble de descripteurs obtenu à partir de la transformée de Fourier Clifford reconstruite : les <i>GCFD3</i>	88
2.4.7	Comparaison des descripteurs couleur	90
2.5	Tableau récapitulatif	91
2.6	Conclusion	93
3	Applications à la reconnaissance d'images couleur	95
3.1	Évaluation des descripteurs de Fourier couleur	98
3.1.1	Les bases d'images couleur	98
3.1.2	Extraction des descripteurs	100
3.1.3	Classification	101
3.1.4	Comparaison des <i>GFD</i> , <i>GCFD</i> et ϕ_{μ_1}	102
3.1.4.1	Sur la base COIL-100	102
3.1.4.2	Sur la base color FERET	104
3.1.4.3	Sur la base color FERET bruitée	105
3.1.5	Choix du bivecteur	105
3.1.5.1	Sur la base COIL-100	107

3.1.5.2	Sur la base color-FERET	107
3.1.5.3	Sur la base ALOI	108
3.1.6	Discussion	108
3.2	Une application de la transformée de Fourier : la corrélation de phase	110
3.2.1	La corrélation de phase pour les images en niveaux de gris	110
3.2.2	La corrélation de phase pour les images couleur	112
3.2.2.1	La corrélation de phase quaternionique couleur	112
3.2.2.2	Proposition d'une corrélation de phase Clifford couleur	114
3.2.3	Comparaison des corrélations de phase quaternionique et Clifford	117
3.2.3.1	Expérimentations sur des images synthétiques	117
3.2.3.2	Utilisation dans le cadre de la reconnaissance d'images	118
3.2.3.3	Expérimentations sur la base COIL-100	119
3.2.3.4	Expérimentations sur la base color FERET	120
3.2.4	Discussion	121
3.3	Conclusion	121
 Conclusion générale et perspectives		123
 Annexes		125
A Les quaternions		127
B Les algèbres de Clifford		133
C Algorithmes		141
 Publications		147
 Table des figures		148
 Bibliographie		151

Notations

Domaines de définition

- \mathbb{R} ensemble des nombres réels
- \mathbb{C} ensemble des nombres complexes
- \mathbb{Z} ensemble des nombres entiers
- $SimO^+(\mathbb{R}^2) = \mathbb{R}_+^* \times S^1$ ensemble des similitudes directes planes
- $M_2 = \mathbb{R}^2 \rtimes SO(2)$ ensemble des déplacements du plan
- $\mathbb{H}, \mathbb{H}_0, \mathbb{H}_1$ ensemble des quaternions, purs, unitaires
- S^1 ensemble des nombres complexes de module 1
- $SO(2)$ le groupe spécial orthogonal de dimension 2

Transformées de Fourier

- \hat{f} et \check{f} la transformée de Fourier usuelle et son inverse
- \hat{f} et \hat{f}_σ la transformée de Fourier Mellin et sa version analytique
- $\hat{f}(r)_n$ la M_2 -transformée
- \hat{f}^q la transformée quaternionique de Bülow [Bülow 99]
- $\hat{f}_{\mu_1}^L$ et $\hat{f}_{\mu_1}^R$ la transformée quaternionique couleur à gauche et à droite de Sangwine et Ell [Sangwine 00]
- \widehat{f}_B la transformée Clifford couleur de Batard *et al.* [Batard 09]

Quaternions

- $q = a + bi + cj + dk$ un quaternion
- $\mathcal{S}(q)$ la partie scalaire d'un quaternion
- $\mathcal{V}(q)$ la partie imaginaire (vectorielle) d'un quaternion
- $\mathcal{I}(q), \mathcal{J}(q), \mathcal{K}(q)$ les différentes parties imaginaires d'un quaternion
- \times le produit vectoriel

Algèbres de Clifford

- $\mathbb{R}_{3,0}$ l'algèbre de Clifford $\mathbb{R}_{3,0}$
- $\{e_1, \dots, e_n\}$ les vecteurs de base
- \sim la réversion
- A_r^{-1} l'inverse d'un r-blade
- A_r^* le dual d'un r-blade
- $\overline{A_r}$ le conjugué d'un r-blade A_r
- $|A|$ la magnitude d'un multivecteur
- \cdot le produit intérieur
- \wedge le produit extérieur
- $*$ le produit scalaire

Introduction générale

Contexte

Ce mémoire est une contribution dans le domaine de la reconnaissance d'objets dans les images couleur. L'approche usuelle, que nous adoptons ici, repose sur l'utilisation de méthodes de classification [Duda 01], supervisées ou non, prenant en entrée des vecteurs d'attributs issus des images considérées et donnant en sortie la classe associée. Ces vecteurs d'attributs sont constitués de valeurs numériques encodant les caractéristiques discriminantes des objets d'intérêt [Gonzalez 01]. En traitement d'images, ces attributs sont classiquement qualifiés de “descripteurs” et sont le plus souvent calculés sur des images en niveaux de gris, voire binaires dans le cas de documents numérisés. Leur extension, relativement récente, à la couleur est, le plus souvent, obtenue en considérant indépendamment chaque plan (RVB , Lab , etc.) comme une image en niveaux de gris. Les valeurs obtenues sont ensuite simplement concaténées pour former un unique vecteur d'attributs. Cette façon de procéder, dite marginale, même si elle conduit à des résultats parfois concluants tout en permettant la réutilisation de méthodes éprouvées, est intrinsèquement biaisée puisqu'elle suppose l'indépendance des plans couleur et ignore leurs corrélations.

Pour notre part, nous avons cherché à utiliser un maximum d'information colorimétrique en inscrivant notre démarche dans celle de Sangwine et Ell [Sangwine 00] qui propose une représentation de l'espace RVB par l'ensemble des quaternions purs. Toutefois, le cadre mathématique retenu ici est celui des algèbres de Clifford [Hestenes 84] dont l'algèbre des quaternions est un cas particulier [Hamilton 44]. Les éléments d'une telle algèbre sont appelés multivecteurs et représentent des “portions” de sous-espaces vectoriels de dimension arbitraire avec une norme et une orientation. Les descripteurs peuvent être calculés directement sur l'image, dans le domaine spatial, ou après transformation. Nous avons choisi de travailler dans le domaine fréquentiel obtenu par la transformée de Fourier Clifford couleur de Batard *et al.* [Batard 10] qui généralise les transformées de Fourier quaternioniques pré-existantes [Ell 07]. Deux approches sont développées. La première, se fonde sur les descripteurs de Fourier définis sur les images en niveaux de gris et les étend à la couleur. En particulier, les descripteurs définis par Smach *et al.* [Smach 08] sont mis en avant parce qu'ils possèdent par construction de “bonnes” propriétés pour la classification d'images. La deuxième approche généralise la classique corrélation de phase [Reddy 96] aux images couleur. Dans les deux cas, les résultats obtenus sur des bases de données de référence montrent la supériorité des méthodes utilisant les algèbres de Clifford, ou les quaternions, sur les méthodes marginales. Le bénéfice est visible en termes de taux de classification comme en termes de complexité algorithmique.

Contenu

Cette thèse comprend trois chapitres, les deux premiers introduisent les notions théoriques, le dernier les valident par des applications. Nous commençons dans le premier chapitre par un état de l'art des transformées de Fourier utilisées pour la reconnaissance de formes, domaine dans lequel la notion d'invariance aux transformations géométriques, particulièrement les similitudes, est centrale.

Deux parties composent ce chapitre : les transformées en niveaux de gris et leurs extensions en couleur. Après avoir rappelé les définitions des séries de Fourier et de la transformée de Fourier usuelle, la transformée de Fourier-Mellin ainsi que sa version analytique sont présentées. Celles-ci permettent d’obtenir, après une normalisation adéquate, des caractéristiques invariantes en translation, rotation et échelle. Nous notons que la méthode de construction de ces transformées repose sur des étapes *ad hoc*.

Par la suite, nous mettons en avant la définition de la transformée de Fourier abstraite à l’aide de la théorie des groupes [Vilenkin 68]. Cette notion est incontournable pour comprendre la M_2 -transformée de Bornard *et al.* [Bornard 86] qui permet de définir des invariants en translation et en rotation de manière rigoureuse. Enfin, pour compléter cette étude, nous introduisons la transformée de Fourier quaternionique de Bülow [Bülow 99] qui permet d’étudier les symétries d’un signal.

Ces transformées sont comparées dans l’objectif d’en comprendre les avantages et les inconvénients ainsi que de justifier les choix effectués.

Si cette utilisation du domaine fréquentiel pour la reconnaissance d’images en niveaux de gris fait partie du bagage du traiteur d’images, il existe somme toute peu de travaux pour les images couleurs. C’est l’objet de la seconde partie. Les transformées de Fourier quaternioniques de Sangwine et Ell [Sangwine 00] ont été les premières tentatives en ce sens. Après les avoir présentées, nous détaillons la transformée de Fourier Clifford couleur de Batard *et al.* [Batard 10] qui les généralise. Nous faisons le lien entre les transformées quaternioniques et celles construites sur des algèbres de Clifford en montrant que les rotations qui sont réalisées dans ces transformées sont identiques et en proposant une transformée de Fourier couleur analogue à la transformée de Fourier quaternionique à droite. Dans le texte, nous montrons que toutes ces transformées dépendent d’une direction d’analyse. Ce paramètre, qui doit être choisi par l’utilisateur, permet de construire une base dans laquelle la transformée est décomposée en une partie parallèle et une partie orthogonale. Les visualisations que nous proposons montrent l’influence de ce choix. Finalement, un algorithme rapide est proposé qui montre le gain en complexité relativement aux méthodes marginales.

Le second chapitre traite de la construction de descripteurs de Fourier à partir des transformées décrites précédemment. Là encore, deux parties décrivent respectivement les approches niveaux de gris puis couleur. Après les descripteurs usuels définis sur le contour des formes seul, d’autres descripteurs définis sur l’image entière sont étudiés. En particulier, les descripteurs génériques de Zhang *et al.* [Zhang 02a] se basent sur une transformée de Fourier polaire *ad hoc* pour obtenir des descripteurs invariants en translation et en rotation. Toutefois, le passage dans le domaine polaire pose un problème d’interpolation dans le cas des images qui sont définies sur une grille discrète. Les descripteurs de Fourier-Mellin et Mellin analytique, invariants en translation, en rotation et en échelle, souffrent du même problème auquel s’ajoute celui de l’estimation du centre de l’image, là encore à l’aide d’une méthode dédiée. Finalement, les seuls descripteurs qui ne supposent pas de prétraitements heuristiques sont les descripteurs de Fourier généralisés *GFD* de Gauthier *et al.* [Gauthier 91] qui sont invariants par construction en translation et rotation. Ce sont ceux que nous avons privilégiés par la suite. Le passage aux images couleurs est étudié dans la seconde partie du chapitre et constitue l’essentiel de notre contribution. Après les méthodes marginales ou par projection de l’image sur un plan couleur, nous définissons les moments de

Fourier-Mellin quaternioniques couleur à partir de la transformée du même nom. Sur la base de la transformée de Fourier Clifford couleur et des descripteurs de Fourier généralisés de Gauthier *et al.*, nous proposons trois ensembles de descripteurs couleur (*GCFD1*, *GCFD2*, *GCFD3*). Par construction, ils sont invariants en translation et en rotation, comme les *GFD* mais ils utilisent l'information colorimétrique de façon plus efficace que les méthodes marginales.

Le troisième, et dernier chapitre, est consacré à la validation expérimentale de nos descripteurs sur des bases d'images de référence (COIL-100, color FERET, etc.) ainsi qu'à la définition d'une nouvelle corrélation de phase couleur.

Les descripteurs de Fourier couleur *GCFD* sont comparés aux méthodes marginales en termes de taux de reconnaissance. Une partie de nos expérimentations est dédiée à la recherche du choix de la direction d'analyse qui maximise le taux de reconnaissance pour un classificateur de type *SVM*.

Nous étendons la corrélation de phase, usuellement utilisée pour des images en niveaux de gris, à la couleur à partir de la transformée de Fourier Clifford couleur et nous proposons deux méthodes de calcul. Là encore, une série d'expérimentations, qui la compare à la corrélation de phase quaternionique, montre la validité de l'approche ainsi que l'impact du choix de la direction d'analyse.

Enfin, nous concluons ce manuscrit en rappelant nos contributions et en proposant des perspectives tant méthodologiques qu'applicatives.

Remarques

Les descripteurs utilisés en reconnaissance d'images forment une famille bien plus large que les descripteurs de Fourier décrits dans ce document. Citons les moments de Hu [Hu 62], de Zernike [Zernike 34] et de Legendre [Teague 80] qui sont très utilisés avec de bonnes performances [Flusser 09]. Comme les descripteurs de Fourier, ils font partie des méthodes globales car l'ensemble de l'image est considéré pour leur calcul. Toutefois ils sont calculés dans le domaine spatial. Le lecteur intéressé trouvera des éléments de comparaison, notamment en termes de performances dans [Smach 08, Choksuriwong 08]. Les taux de reconnaissance affichés donnent un avantage en faveur des méthodes fréquentielles, particulièrement en cas de changements d'illuminant.

D'autres méthodes plus récentes reposent sur l'extraction de points d'intérêts dans les images (méthodes *SIFT* [Lowe 99], *FAST* [Rosten 10], *GLOH* [Mikolajczyk 05], etc.). Ces méthodes sont locales car les descripteurs sont extraits à partir de points caractéristiques comme les coins des objets présents dans l'image. Elles sont actuellement plus performantes [Choksuriwong 08] que les méthodes globales excepté, là encore, dans le cas de changements d'illuminant.

Enfin, signalons que ces méthodes, globales ou locales, sont développées en niveaux de gris et étendues à la couleur par des approches *ad-hoc*. Pour notre part, nous nous sommes focalisés sur les approches fréquentielles et plus particulièrement sur l'apport du contenu fréquentiel couleur.

CHAPITRE 1

LES TRANSFORMÉES DE FOURIER : UN ÉTAT DE L'ART

Les transformées de Fourier : un état de l'art

Sommaire

1.1	Les transformées de Fourier pour les images en niveaux de gris	18
1.1.1	La transformée de Fourier complexe	19
1.1.1.1	Définition	19
1.1.1.2	Propriétés	20
1.1.1.3	Interprétation	22
1.1.2	La transformée de Fourier-Mellin classique et sa version analytique	25
1.1.2.1	Définitions	25
1.1.2.2	Propriétés	28
1.1.3	La transformée de Fourier abstraite	29
1.1.4	La M_2 -transformée	31
1.1.4.1	Définition	31
1.1.4.2	Propriétés	33
1.1.5	La transformée de Fourier quaternionique de Bülow [Bülow 99]	34
1.1.5.1	Définition	35
1.1.5.2	Propriétés	36
1.1.6	Discussion	37
1.2	Les transformées de Fourier pour les images couleur	38
1.2.1	Les généralisations de la transformée de Fourier	38
1.2.2	La transformée de Fourier quaternionique de Sangwine et Ell [Sangwine 00]	40
1.2.2.1	Définitions	40
1.2.2.2	Interprétation en termes de rotation	42
1.2.2.3	Propriétés de la transformée de Fourier quaternionique	43
1.2.2.4	Choix de μ_1 et construction de la base $\{\mu_1, \mu_2, \mu_3\}$	44
1.2.2.5	Algorithme de calcul	46
1.2.2.6	Visualisation et interprétation générale	47
1.2.3	La transformée de Fourier Clifford couleur de Batard <i>et al.</i> [Batard 10]	49
1.2.3.1	Définitions	49
1.2.3.2	Interprétation en termes de rotation	53
1.2.3.3	Propriétés de la transformée de Fourier Clifford couleur	54
1.2.3.4	Construction de B	55
1.2.3.5	Construction de la base $\{c, e_4, v, w\}$	56
1.2.3.6	Algorithme de calcul	57
1.2.3.7	Visualisation et interprétation générale	59
1.2.4	Discussion	61
1.3	Conclusion	63

L'objectif de cette thèse est de définir à partir du domaine fréquentiel d'une image, autrement dit de sa transformée de Fourier, des descripteurs permettant de reconnaître des objets dans les images. Le contexte est donc celui du traitement d'images et de la reconnaissance de formes. Dans la littérature, les descripteurs construits à partir de la transformée de Fourier sont issus d'images en niveaux de gris ou d'images couleur traitées plan par plan. Cette approche est usuellement qualifiée de marginale. Une de nos contributions est de traiter l'image couleur dans sa globalité, sans traitement *ad hoc*, pour obtenir des descripteurs dépendants de la couleur.

Dans ce chapitre, un état de l'art original des transformées de Fourier pour les images en niveaux de gris est présenté. L'objectif est d'étudier leurs propriétés au regard des transformations usuelles des images (translation, rotation et homothétie). Celles-ci sont liées directement aux déplacements dans l'espace des objets (par exemple, la translation d'une image correspond à la translations des objets qui la composent). L'étude de ces transformations est donc un préalable à la définition de descripteurs de formes, quel que soit le modèle mathématique considéré.

L'extension de telles transformées aux images nD est un problème difficile obligeant à revenir sur la construction mathématique de ces transformées. Notre intérêt se tourne donc vers une autre famille de transformées de Fourier couleur. Ces dernières traitent l'information couleur dans un cadre résolument géométrique et donc de manière non marginale, comme la transformée de Fourier quaternionique de Sangwine et Ell [Sangwine 00] et la transformée de Fourier Clifford de Batard *et al.* [Batard 10]. Cette dernière est au coeur de ce travail de thèse car, étant définie à l'aide d'un modèle mathématique général, elle permet de généraliser les transformées de Fourier quaternioniques couleur.

Après une définition complète de ces deux transformées, une définition alternative à la transformée de Fourier Clifford couleur est formalisée pour obtenir une forme équivalente à la transformée de Fourier quaternionique dite à droite.

Nous observons alors qu'une transformée de Fourier adaptée aux images couleur implique nécessairement un paramétrage lié à la direction d'analyse (vecteur ou plan) de l'image. Dans le cadre de la transformée de Fourier Clifford couleur, il s'agit du bivecteur B . Ce paramètre fondamental définit la direction d'analyse de cette transformée et influe directement sur la pertinence des descripteurs qui peuvent être extraits.

La construction et le choix de ce bivecteur apparaissent alors comme des phases déterminantes pour l'utilisation de la transformée de Fourier Clifford couleur. Ainsi, dans ce chapitre, notre construction de ce paramètre est détaillée, puis, dans le chapitre consacré aux expérimentations, nous aborderons son choix. De plus, nous détaillons un algorithme original de calcul rapide de la transformée de Fourier Clifford couleur.

1.1 Les transformées de Fourier pour les images en niveaux de gris

Dans cette partie, nous nous intéressons à quatre transformées de Fourier pour les images en niveaux de gris : la transformée de Fourier usuelle [Bracewell 86], la transformée de Fourier-Mellin [Sheng 86b] ainsi que sa version analytique [Ghorbel 93], la M_2 -transformée [Bornard 86]

et la transformée de Fourier quaternionique de Bülow [Bülow 99].

La définition de ces transformées ainsi que leurs propriétés concernant les effets de certaines transformations spatiales de l'image (translation, rotation et homothétie) dans le domaine fréquentiel sont étudiées.

1.1.1 La transformée de Fourier complexe

C'est dans un article de 1822 que le mathématicien et physicien français Joseph Fourier publie ses travaux [Fourier 22] qui ont permis de résoudre les équations de propagation de la chaleur dans les corps solides. Ceux-ci reposent sur la décomposition de fonctions périodiques $f(x)$ réelles (intégrables sur $[-\pi, \pi]$) en une somme infinie de fonctions en sinus et en cosinus

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{+\infty} a_n \cos(nx) + \sum_{n=1}^{+\infty} b_n \sin(nx) \quad (1.1)$$

avec

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \end{aligned}$$

Les coefficients a_n et b_n sont appelés les coefficients de Fourier. Cette méthode correspond au calcul des séries de Fourier qui utilisent les relations d'orthogonalité entre les fonctions sinus et cosinus. Pour obtenir une forme simplifiée, ces séries peuvent être réécrites sous la forme d'une somme infinie de fonctions exponentielles dont les parties réelles et imaginaires correspondent à a_n et b_n :

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{i2\pi \frac{n}{T} x} \quad (1.2)$$

où T est la période de f et $c_n(f)$ sont les coefficients de Fourier qui s'écrivent

$$c_n(f) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i2\pi \frac{n}{T} t} dt \quad (1.3)$$

La transformée de Fourier est une généralisation des séries de Fourier complexes pour les fonctions non périodiques.

1.1.1.1 Définition

Soit f une fonction de carré intégrable sur \mathbb{R} , sa transformée de Fourier [Bracewell 86] est définie par

$$\hat{f}(u) = \int_{-\infty}^{+\infty} f(x) e^{-2i\pi ux} dx \quad (1.4)$$

où x représente la coordonnée spatiale et u la coordonnée fréquentielle. Cette transformée est inversible et son inverse \check{f} s'écrit

$$\check{f}(x) = \int_{-\infty}^{+\infty} \hat{f}(u) e^{2i\pi ux} du \quad (1.5)$$

La transformée de Fourier peut être étendue aux fonctions de \mathbb{R}^n dans \mathbb{C} . Elle est définie comme

$$\hat{f}(\mathbf{u}) = \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-2i\pi\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \quad (1.6)$$

où \mathbf{x} et \mathbf{u} sont des vecteurs, et $\langle \cdot \rangle$ est le produit scalaire usuel.

Cette intégrale peut se réécrire comme

$$\hat{f}(u_1, u_2, \dots, u_n) = \int_{\mathbb{R}} \left(\dots \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2, \dots, x_n) e^{-2i\pi u_1 x_1} dx_1 \right) e^{-2i\pi u_2 x_2} dx_2 \right) \dots \right) e^{-2i\pi u_n x_n} dx_n \quad (1.7)$$

Cela permet, d'un point de vue pratique, de calculer cette transformée en composant des transformées de Fourier 1D relativement à chacune des dimensions de l'espace de départ de la fonction f .

Dans le cas des images en niveaux de gris, le signal à traiter est défini de \mathbb{R}^2 dans \mathbb{C} . La transformée de Fourier 2D d'une image en niveaux de gris s'écrit

$$\hat{f}(u_1, u_2) = \int_{\mathbb{R}^2} f(x_1, x_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \quad (1.8)$$

avec x_1, x_2 les coordonnées spatiales des pixels dans l'image et u_1, u_2 les coordonnées fréquentielles. Notons que dans l'équation (1.8), le terme $f(x_1, x_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)}$ peut s'interpréter comme une rotation dans le plan complexe de la valeur de la fonction f en (x_1, x_2) ¹. L'angle de cette rotation, donné par $2\pi(u_1 x_1 + u_2 x_2)$, dépend à la fois des coordonnées spatiales et fréquentielles. Nous verrons dans la partie 1.2.3 que cette rotation peut être interprétée dans le cadre des actions et des représentations de groupe.

1.1.1.2 Propriétés

Le but est ici d'étudier le comportement de la transformée de Fourier 2D d'une image face à certaines transformations dans le domaine spatial (translations, rotations, changements d'échelle). La première propriété abordée est connue sous le nom de "Shift-Theorem". Celui-ci montre l'effet d'une translation de l'image dans le domaine fréquentiel.

Translation Une translation, définie par un vecteur $(\Delta_1, \Delta_2)^t$, de l'image produit un déphasage dans le domaine fréquentiel, soit la multiplication par un complexe de module 1 dépendant de la translation.

Démonstration. Soit f une fonction de $L^2(\mathbb{R}^2, \mathbb{C})$, $x'_1 = x_1 - \Delta_1$, $x'_2 = x_2 - \Delta_2$ et $g(x_1, x_2) = f(x_1 - \Delta_1, x_2 - \Delta_2)$ alors

$$\begin{aligned} \hat{g}(u_1, u_2) &= \int_{\mathbb{R}^2} f(x_1 - \Delta_1, x_2 - \Delta_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} f(x'_1, x'_2) e^{-2i\pi(u_1(x'_1 + \Delta_1) + u_2(x'_2 + \Delta_2))} dx'_1 dx'_2 \\ &= \left(\int_{\mathbb{R}^2} f(x'_1, x'_2) e^{-2i\pi(u_1 x'_1 + u_2 x'_2)} dx'_1 dx'_2 \right) e^{-2i\pi(u_1 \Delta_1 + u_2 \Delta_2)} \\ &= \hat{f}(u_1, u_2) e^{-2i\pi(u_1 \Delta_1 + u_2 \Delta_2)} \end{aligned} \quad (1.9)$$

□

1. En effet, $f(\mathbf{x}) e^{-i\theta} = f(\mathbf{x}) (\cos(\theta) - i \sin(\theta)) = f(\mathbf{x}) \cos(\theta) - i f(\mathbf{x}) \sin(\theta)$

Cette propriété peut être utilisée pour estimer la meilleure translation entre deux images. Cette méthode, appelée corrélation de phase, sera abordée et généralisée dans le chapitre applicatif. Nous pouvons aussi remarquer que $\forall u_1, u_2, |\widehat{f}(u_1, u_2)| = |\widehat{g}(u_1, u_2)|$ et donc que le module de la transformée de Fourier est invariant en translation.

Rotation Une rotation d'angle θ de l'image, par rapport à l'origine, induit une rotation identique des coordonnées fréquentielles.

Démonstration. Soit θ un angle de rotation et $g_\theta(x_1, x_2) = f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta))$ alors

$$\widehat{g}_\theta(u_1, u_2) = \int_{\mathbb{R}^2} f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta)) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \quad (1.10)$$

Par changement de variables,

$$\begin{cases} x'_1 = x_1 \cos(\theta) + x_2 \sin(\theta) \\ x'_2 = -x_1 \sin(\theta) + x_2 \cos(\theta) \end{cases} \Rightarrow \begin{cases} x_1 = x'_1 \cos(\theta) - x'_2 \sin(\theta) \\ x_2 = x'_1 \sin(\theta) + x'_2 \cos(\theta) \end{cases}$$

Les formes différentielles de x_1 et x_2 deviennent

$$\begin{cases} dx_1 = \cos(\theta) dx'_1 - x'_2 \sin(\theta) dx'_2 \\ dx_2 = \sin(\theta) dx'_1 + x'_2 \cos(\theta) dx'_2 \end{cases}$$

Le déterminant de la Jacobienne de ce changement de coordonnées est 1, ce qui implique que $dx'_1 dx'_2 = dx_1 dx_2$.

L'équation suivante est obtenue

$$\begin{aligned} \widehat{g}_\theta(u_1, u_2) &= \int_{\mathbb{R}^2} g(x_1, x_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} f(x'_1, x'_2) e^{-2i\pi(u_1 x'_1 \cos(\theta) - u_1 x'_2 \sin(\theta) + u_2 x'_1 \sin(\theta) + u_2 x'_2 \cos(\theta))} dx'_1 dx'_2 \\ &= \int_{\mathbb{R}^2} f(x'_1, x'_2) e^{-2i\pi[(u_1 \cos(\theta) + u_2 \sin(\theta))x'_1 + (-u_1 \sin(\theta) + u_2 \cos(\theta))x'_2]} dx'_1 dx'_2 \\ &= \widehat{f}(u_1 \cos(\theta) + u_2 \sin(\theta), -u_1 \sin(\theta) + u_2 \cos(\theta)) \end{aligned} \quad (1.11)$$

□

Homothétie Une homothétie d'une image de rapport k dans la direction x_1 se traduit par une homothétie de rapport $\frac{1}{k}$ dans la direction u_1 .

Démonstration. Si on contracte $f(x_1, x_2)$ dans la direction x_1 (c'est-à-dire $g(x_1, x_2) = f(kx_1, x_2)$) alors

$$\widehat{g}(u_1, u_2) = \int_{\mathbb{R}^2} f(kx_1, x_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2$$

En effectuant le changement de variable $x'_1 = kx_1$ alors $dx_1 = \frac{1}{k}dx'_1$ et

$$\begin{aligned}\widehat{g}(u_1, u_2) &= \int_{\mathbb{R}^2} f(x'_1, x_2) e^{-2i\pi(\frac{u_1 x'_1}{k} + u_2 x_2)} \frac{1}{k} dx'_1 dx_2 \\ &= \frac{1}{k} \widehat{f}\left(\frac{u_1}{k}, u_2\right)\end{aligned}\quad (1.12)$$

□

Dans le cas où cette même homothétie de rapport k est effectuée dans les deux directions x_1 et x_2 (i.e. $g(x_1, x_2) = f(kx_1, kx_2)$: contraction sans déformation de l'image), il y a une homothétie de rapport $\frac{1}{k}$ dans le domaine fréquentiel ainsi que la multiplication des coefficients par $\frac{1}{k^2}$:

$$\widehat{g}(u_1, u_2) = \frac{1}{k^2} \widehat{f}\left(\frac{u_1}{k}, \frac{u_2}{k}\right) \quad (1.13)$$

D'après les équations précédentes, nous pouvons noter que la seule invariance qui peut être facilement obtenue (à l'aide du module) avec la transformée de Fourier $2D$ classique est celle relative à la translation. Nous verrons par la suite que les autres types d'invariances peuvent être obtenus, mais en considérant d'autres transformées de Fourier.

1.1.1.3 Interprétation

La transformée de Fourier $2D$ est un outil fondamental en traitement d'images qui permet de passer du domaine spatial au domaine fréquentiel. Une étude de ce dernier permet d'identifier les fréquences caractéristiques des images. Dans le cas de la définition de descripteurs, cela contribue à la sélection des coefficients de Fourier nécessaires à la reconnaissance. Par exemple, dans le cas de la reconnaissance d'objets, il faut s'intéresser tout particulièrement aux basses fréquences (plus représentatives des formes) et pour la reconnaissance de textures, ce sont les hautes fréquences qui sont privilégiées [Smach 08].

Le résultat de la transformée de Fourier $2D$ est une image à valeurs complexes. Les basses fréquences sont présentes aux quatre coins de l'image résultante et les hautes fréquences sont concentrées au centre de l'image.

Dans l'objectif de faciliter son interprétation visuelle, l'image de la transformée de Fourier est usuellement réarrangée en repositionnant son origine au centre de cette image ("FFTshift"). Les basses fréquences sont alors ramenées au voisinage du centre de l'image et les hautes fréquences se retrouvent éloignées de celui-ci. Ici et par la suite, nous considérerons toujours les transformées de Fourier dans cette configuration pour en obtenir une visualisation conforme aux habitudes des traiteurs d'images.

La visualisation de l'image complexe formée par les coefficients de la transformée de Fourier $\widehat{f}(u_1, u_2)$ n'est pas directe. Ses coefficients sont considérés sous leur forme polaire pour visualiser un module d'une part et une phase d'autre part :

$$\widehat{f}(u_1, u_2) = r(u_1, u_2) e^{i\theta(u_1, u_2)} \quad (1.14)$$

où $r(u_1, u_2)$ est le module de $\widehat{f}(u_1, u_2)$ et $\theta(u_1, u_2)$ sa phase.

Le module de la transformée de Fourier (aussi appelé “spectre” de la transformée de Fourier [Gonzalez 01]) nous informe de l’amplitude d’une fréquence présente dans l’image. Il fournit une intensité qui peut être commodément visualisée sous la forme d’une image en niveaux de gris où chaque pixel correspond au logarithme du module. Cela permet d’étaler l’histogramme du spectre de la transformée de Fourier 2D qui, dans le cas des images naturelles, décroît exponentiellement en s’éloignant de l’origine [Torralba 03].

Dans l’exemple de la “maison de Fourier” donné en figure 1.1, des “pics” sont présents sur l’image spectrale. Ils sont représentatifs de fréquences particulières et permettent une interprétation visuelle de l’image que l’on veut analyser. La figure 1.1 montre les modules de la transformée de Fourier de “la maison de Fourier” à différentes étapes de sa construction. Nous remarquons alors que :

- A l’étape 1, il n’y a que des stries horizontales dans l’image. Cela induit des pics sur l’axe des ordonnées dans l’image du module de la transformée de Fourier.
- A l’étape 2, nous remarquons des pics sur l’axe des abscisses qui sont dus aux stries verticales.
- Aux étapes 3 et 4, les pics sur les diagonales nous indiquent la présence de stries dans les diagonales de l’image.

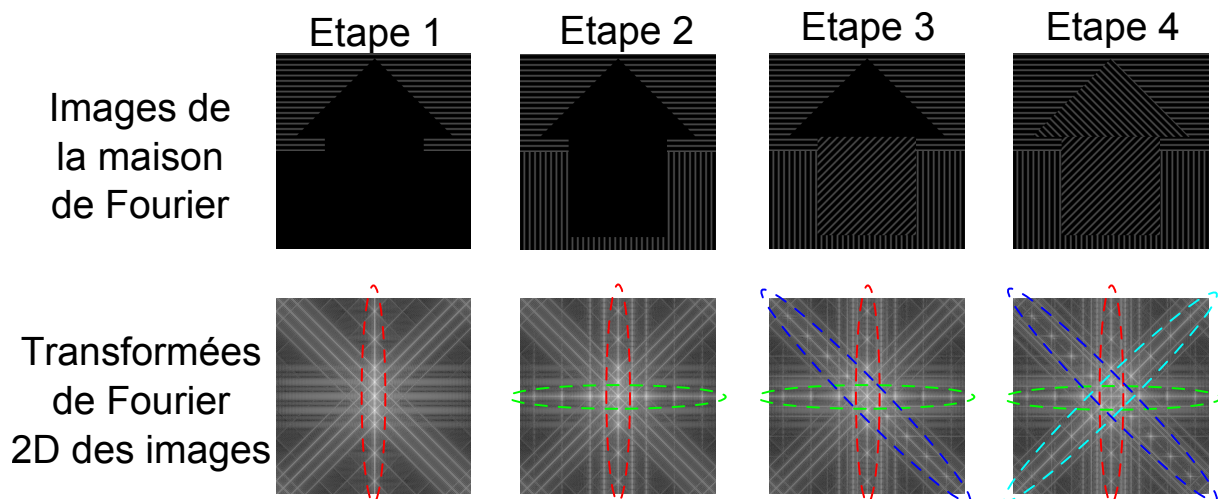


FIGURE 1.1 – Illustration du log-module de la transformée de Fourier à chaque étape de la construction de “la maison de Fourier”

Notons que le module nous informe seulement de la présence de fréquences dans l’image, mais pas sur leur localisation. Cette observation permet de rappeler que la transformée de Fourier est globale et non locale.

L’interprétation visuelle du module devient moins évidente dans le cas d’une image naturelle comme celle de Lenna dans la figure 1.3 où le mélange des fréquences ne permet pas de détecter de pics significatifs.

La phase contient la “localisation” d’une fréquence dans l’image et elle peut être représentée par une image TSL (Teinte Saturation Luminance) dont la teinte est égale à la phase (la saturation et la luminance sont fixées à 1). Elle contient la majeure partie de l’information contour.

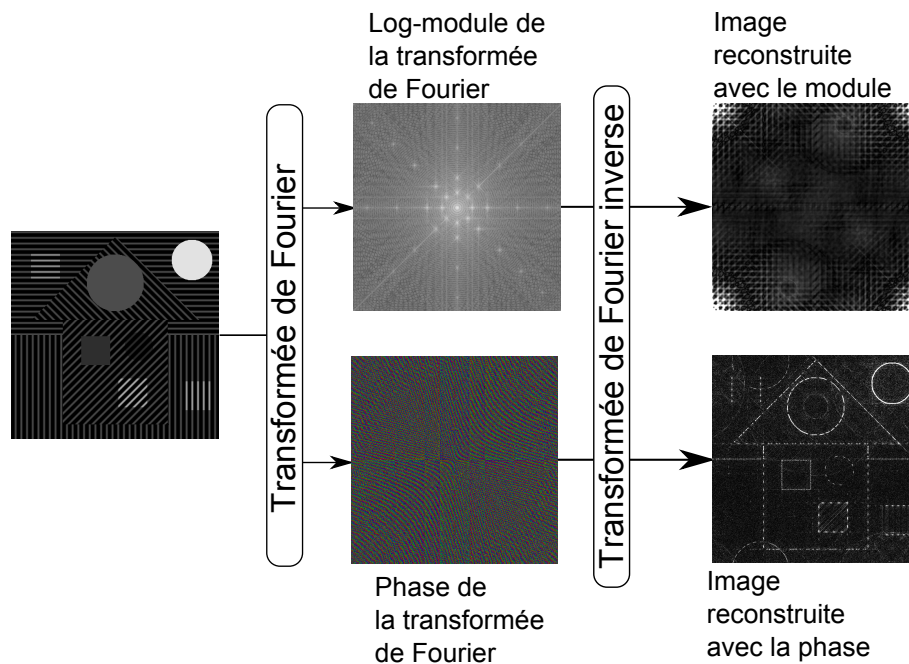


FIGURE 1.2 – Reconstruction de l'image de la "maison de Fourier" en conservant uniquement le module (en haut) et la phase (en bas)

Ainsi, Kovési [Kovési 03] l'utilise dans le concept de congruence de phase pour détecter les coins et les contours.

Cependant, l'image représentant la phase n'a pas d'interprétation visuelle simple, quelle que soit l'image originale, car son aspect est très bruité. Une méthode pour visualiser les informations que contient la phase (respectivement le module) est de reconstruire l'image originale à l'aide d'une transformée de Fourier inverse en négligeant le module en fixant $r(u_1, u_2) = 1$ (respectivement la phase en fixant $\theta(u_1, u_2) = 0$). Dans les figures 1.2 et 1.3, nous remarquons qu'une grande partie des contours est retrouvée. Par contre, l'image reconstruite uniquement avec le module est difficilement interprétable visuellement.

L'origine de la transformée de Fourier $\hat{f}(0, 0)$ représente la plus petite fréquence du spectre de Fourier. Elle est appelée composante continue de l'image et correspond à la moyenne des niveaux de gris de l'image. En s'éloignant du centre, nous retrouvons les basses fréquences qui représentent les faibles variations des composantes de l'image. Celles-ci concernent principalement les régions homogènes. Loin du centre, les hautes fréquences sont dues aux variations rapides des composantes de l'image. Elles sont donc représentatives des contours, des changements brusques d'intensité et du bruit dans les images.

En conclusion, la transformée de Fourier usuelle n'est invariante à aucune des transformations considérées. Nous verrons toutefois plus tard qu'il est possible de calculer, en considérant d'autres transformées de Fourier, des descripteurs respectant certaines invariances.

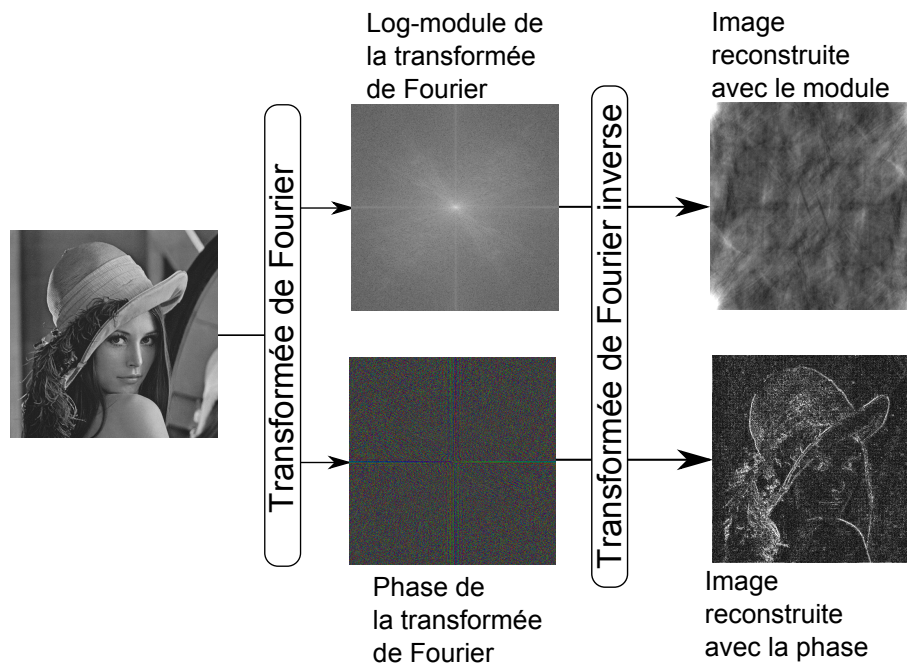


FIGURE 1.3 – Reconstruction de l'image de Lenna en conservant uniquement le module (en haut) et la phase (en bas)

Dans la littérature, on trouve d'autres transformées de Fourier pour les images en niveaux de gris. Dans la partie suivante, nous traitons l'une d'entre elles : la transformée de Fourier-Mellin.

1.1.2 La transformée de Fourier-Mellin classique et sa version analytique

Originellement, la transformée de Fourier-Mellin a été définie dans le but de mettre en oeuvre des méthodes invariantes aux similitudes directes². Dans [Casasent 76], la transformée de Fourier-Mellin est utilisée dans le domaine de l'optique pour définir une corrélation de phase qui ne varie pas si l'image subit une translation, une rotation et un changement d'échelle. Plus tard, elle a été utilisée dans le domaine du document numérique [Sheng 86b, Grace 91] pour définir des descripteurs invariants, ce qui sera abordé dans le chapitre 2.3.2.1.

1.1.2.1 Définitions

La transformée de Fourier-Mellin, très utilisée en traitement d'images [Sheng 86b, Derrode 01], a été définie pour obtenir des descripteurs invariants en translation (à l'aide d'un pré-traitement), en rotation et en échelle (par construction). Pour une image en niveaux de gris f de \mathbb{R}^2 dans \mathbb{C} passée en coordonnées polaires, elle s'écrit

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \hat{f}(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{-iv} e^{-ik\theta} d\theta \frac{dr}{r} \quad (1.15)$$

2. translation, rotation et homothétie

où f est supposée intégrable sur le groupe $\mathbb{R}_+^* \times S^1$,

$$\int_0^\infty \int_0^{2\pi} \left| f(r, \theta) r^{-iv} e^{-ik\theta} \right| d\theta \frac{dr}{r} = \int_0^\infty \int_0^{2\pi} \frac{1}{r} f(r, \theta) d\theta dr < \infty \quad (1.16)$$

Dans [Sheng 86a], les grandes étapes du calcul de cette transformée sont :

1. Le centre de l'image est redéfini comme étant le centre de l'objet d'intérêt pour obtenir une invariance à la translation. Cette étape est réalisée à l'aide d'une méthode *ad hoc* comme l'auto-corrélation ou, pour un objet isolé, par le calcul du centre d'inertie de l'objet avec les moments cartésiens [Gonzalez 01].
2. La représentation polaire de l'image, appelée souvent image polaire (voir figure 1.4), est utilisée pour convertir une rotation de l'image dans le domaine spatial en une simple translation suivant l'axe angulaire θ . Une transformée de Fourier 1D est ensuite appliquée suivant l'axe angulaire de l'image polaire. Cela correspond aux "fonctions harmoniques circulaires" d'une image polaire [Sheng 86a]. La translation dans l'image polaire se traduit alors par un déphasage dans le domaine fréquentiel comme pour la transformée de Fourier usuelle.
3. Une transformée de Mellin est appliquée suivant l'axe radial r dans l'objectif de gérer les changements d'échelle de l'image. En effet, dans le domaine polaire, un changement d'échelle n'affecte pas l'axe angulaire mais induit une homothétie suivant l'axe radial. La transformée de Mellin permet alors d'extraire le changement d'échelle sous la forme d'un coefficient multiplicatif.

En pratique, la transformée de Fourier-Mellin ainsi définie soulève un certain nombre de problèmes. Le premier est celui de la définition de l'origine du repère pour obtenir l'invariance en translation car elle doit être estimée par une méthode *ad hoc* qui peut échouer (à cause du bruit notamment).

Le second concerne le passage de l'image dans le domaine polaire, qui implique une étape d'interpolation car l'image est définie sur une grille discrète. Il est donc impossible d'obtenir un calcul exact de cette transformée pour une image.

De plus, l'hypothèse d'intégrabilité de f sur le groupe $\mathbb{R}_+^* \times S^1$ n'est pas garantie si f est une image. En effet, l'intégrale diverge quand r tend vers 0 si $f(0, \theta)$ est différent de 0. Or, $f(0, \theta)$ représente l'origine de l'image qui est un pixel dont le niveau de gris est non nul en général. Par contre, quand r tend vers l'infini, l'existence de l'intégrale à l'infini est évidente car les images ont une étendue spatiale finie [Derrode 99]. En dehors de l'image, les pixels sont supposés de valeur nulle.

Pour remédier au problème de divergence de cette transformée, Zwicke et Kiss proposent dans [Zwicke 83] d'annuler les valeurs de l'image au voisinage de l'origine sur un disque de rayon r . Malheureusement, cela engendre une perte d'information sur le contenu de l'image non négligeable pour le calcul de la transformée de Fourier-Mellin. En effet, au voisinage de $r = 0$, la valeur de $\frac{1}{r}$ est très grande. Le rayon du disque doit donc être le plus petit possible. Celui-ci est difficile à déterminer. Par exemple, si une image g correspond à une image f ayant subi une homothétie de rapport inconnu, et si le rayon minimal est constant pour toutes les images, alors la perte d'information sera différente pour les deux images.

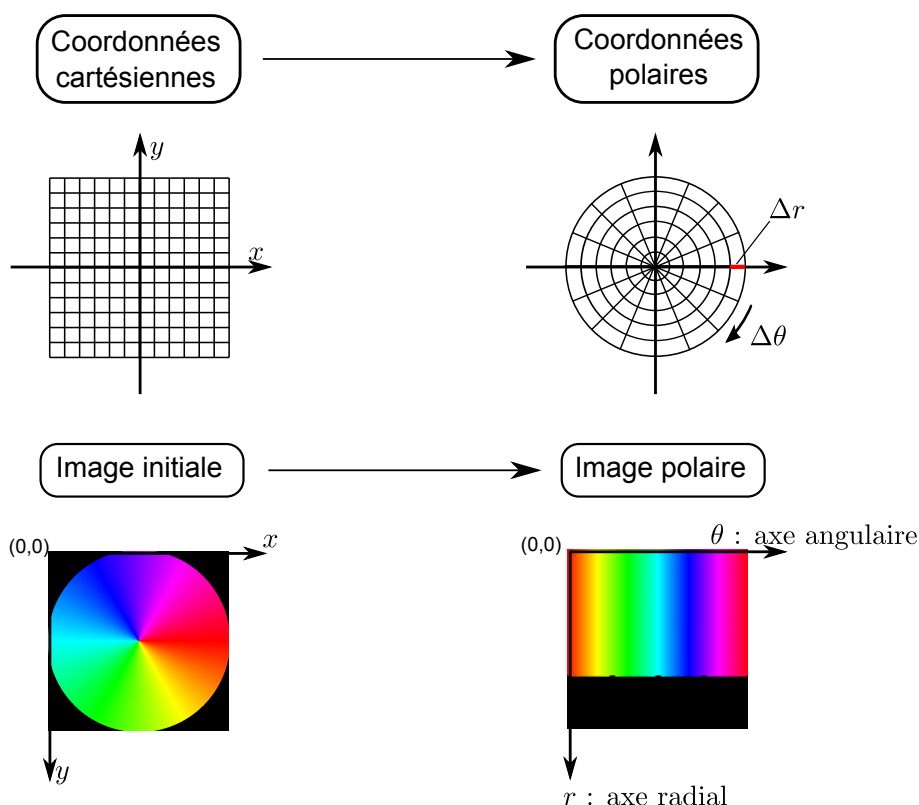


FIGURE 1.4 – Passage d'une image en coordonnées polaires

Toujours pour gérer la divergence de la transformée de Fourier-Mellin, [Ghorbel 93] propose de considérer son prolongement analytique, qui revient à traiter l'image $r^\sigma f(r, \theta)$ au lieu de $f(r, \theta)$ avec σ un nombre réel strictement positif fixé. Cette transformée est appelée la transformée de Fourier-Mellin analytique.

Soit f une fonction représentant une image en niveaux de gris, la transformée de Fourier-Mellin analytique s'écrit

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \hat{f}_\sigma(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r} \quad (1.17)$$

avec $\sigma > 0$.

Une visualisation du module de la transformée de Fourier-Mellin analytique³ d'une image de la base COIL-100 [Nene 96] est présentée dans la figure 1.5.

Théoriquement, cette transformée est inversible et donne une représentation unique des images pour une méthode d'estimation du centre et un σ donnés. Malheureusement, l'étape d'interpolation de f dans le domaine polaire fait perdre en pratique cette propriété ($\hat{f}_\sigma \neq f$). On trouve dans [Derrode 01] trois algorithmes d'approximation pour calculer la transformée de Fourier-Mellin analytique pratiquement. La première méthode repose sur le passage de l'image

3. calculée à l'aide de l'algorithme rapide de [Derrode 01] avec $\sigma = 0.5$

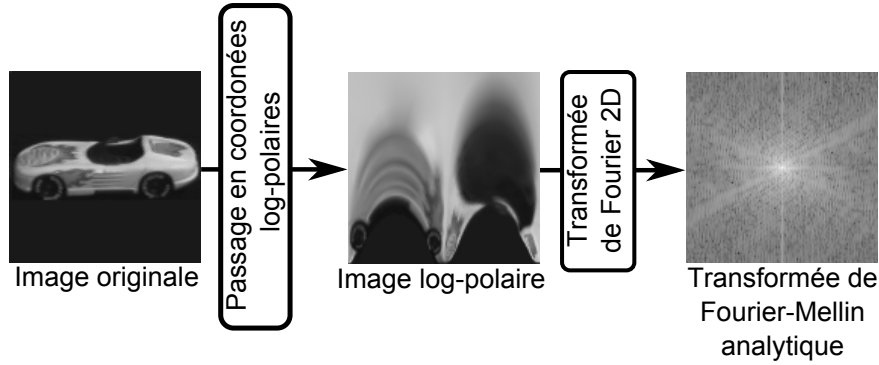


FIGURE 1.5 – Transformée de Fourier-Mellin analytique d'une image

en coordonnées polaires, suivi d'une estimation de la transformée de Fourier-Mellin Analytique (équation 1.17). La seconde revient à rééchantillonner l'image en coordonnées log-polaires (image log-polaire), puis à calculer une transformée de Fourier 2D. La troisième se base directement sur les coordonnées polaires de l'image.

1.1.2.2 Propriétés

Dans cette partie, nous étudions l'effet sur la transformée de Fourier-Mellin analytique de la translation, de la rotation et du changement d'échelle de l'image dans l'objectif de définir des invariants à ces transformations.

Translation Comme il a été exposé précédemment, l'invariance est obtenue en redéfinissant l'origine de l'image polaire qui est par défaut placée au centre de l'image originale mais qui peut aussi être définie à l'aide d'une fonction *ad hoc*. Ce n'est pas une propriété inhérente à la transformée de Fourier-Mellin et donc, par construction, la transformée de Fourier-Mellin n'est pas invariante en translation.

Rotation et homothétie Nous avons vu précédemment qu'une rotation d'angle β de l'image dans le domaine spatial induit une translation de l'image polaire et donc un déphasage dans le domaine fréquentiel. Un changement d'échelle de rapport α fait apparaître un coefficient multiplicatif.

Démonstration. Soit l'image g correspondant à l'image f ayant subi une rotation d'angle β et un changement d'échelle de facteur α , c'est-à-dire si $g(r, \theta) = f(\alpha r, \theta + \beta)$ alors

$$\begin{aligned}
 \forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \hat{g}_\sigma(k, v) &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(\alpha r, \theta + \beta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r} \\
 &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(\xi, \varphi) \left(\frac{\xi}{\alpha}\right)^{\sigma-iv} e^{-ik(\varphi-\beta)} d\varphi \frac{d\xi}{\xi} \\
 &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(\xi, \varphi) \xi^{\sigma-iv} \alpha^{-\sigma+iv} e^{-ik\varphi} e^{ik\beta} d\varphi \frac{d\xi}{\xi} \\
 &= \left(\frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(\xi, \varphi) \xi^{\sigma-iv} e^{-ik\varphi} d\varphi \frac{d\xi}{\xi} \right) \alpha^{-\sigma+iv} e^{ik\beta} \\
 &= \hat{f}_\sigma(k, v) \alpha^{-\sigma+iv} e^{ik\beta}
 \end{aligned} \tag{1.18}$$

avec le changement de variable $\xi = \alpha r$ et $\varphi = \theta + \beta$. □

À partir de ce résultat, la transformée de Fourier-Mellin analytique peut être utilisée pour obtenir des invariants en rotation et en échelle à l'aide d'une normalisation adéquate (voir la partie concernant les descripteurs de Fourier-Mellin).

Dans la figure 1.6, les modules des transformées de Fourier-Mellin analytique de deux images (dont l'une a subi une rotation et un changement d'échelle) sont présentés. Nous pouvons remarquer que les spectres des transformées de Fourier-Mellin analytiques, après une normalisation adéquate, sont semblables à l'exception de quelques différences mineures qui sont certainement dues à l'estimation de l'image log-polaire.

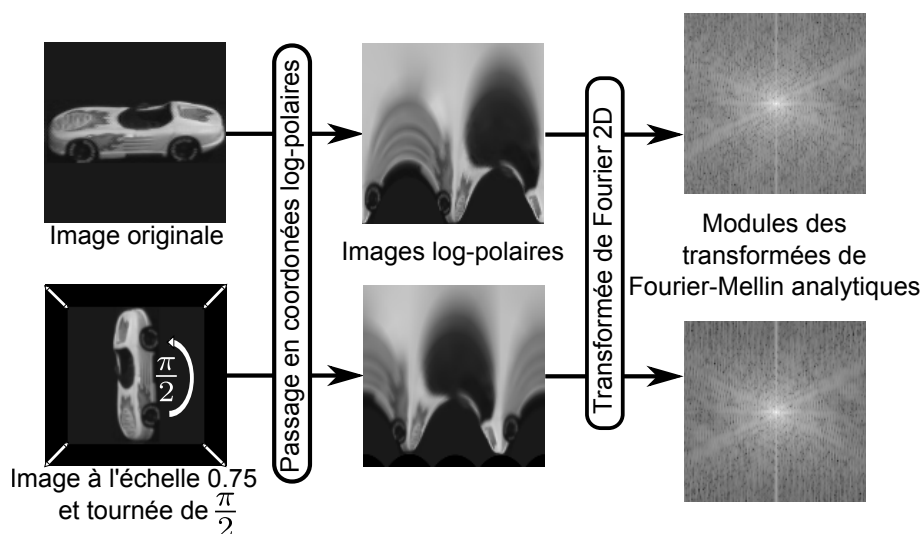


FIGURE 1.6 – Transformée de Fourier-Mellin analytique d'une image ayant subi une rotation d'angle $\frac{\pi}{2}$ et une homothétie de rapport 0.75

Soulignons le rôle de l'estimation du centre de l'image qui est très sensible au bruit. Si le centre de l'image est mal approximé, alors l'invariance en translation n'est pas obtenue. Le calcul de la transformée de Fourier-Mellin est dépendante de cette invariance en translation, ce qui implique que si le centre est mal estimé alors les descripteurs qui seraient extraits de cette transformée ne respecteraient plus aucune invariance.

Les transformées de Fourier que nous venons d'aborder ne sont pas les seules qui existent. En effet, à l'aide d'outils mathématiques de la théorie des groupes, nous verrons que les transformées de Fourier que nous avons définies précédemment ainsi que de nouvelles peuvent être construites.

1.1.3 La transformée de Fourier abstraite

La transformée de Fourier abstraite est fortement liée à la théorie de représentation des groupes [Vilenkin 68]. Dans cette partie, nous donnons des pistes pour comprendre les éléments de cette théorie qui nous intéressent. Nous l'abordons donc volontairement de manière simplifiée car nous l'utilisons principalement comme un outil. Commençons par rappeler des définitions importantes sur lesquelles nous nous appuyerons.

Un morphisme de groupe est une application entre deux groupes qui respecte la structure et la loi de composition des groupes. Considérons deux groupes $(G_1, +)$ et $(G_2, *)$, alors l'application $f : G_1 \mapsto G_2$ est un morphisme de groupe si

$$\forall (x_1, x_2) \in (\mathbb{G}_1 \times \mathbb{G}_1), f(x_1 + x_2) = f(x_1) * f(x_2) \quad (1.19)$$

Pour un espace vectoriel V , le groupe général linéaire $GL(V)$ est le groupe des isomorphismes de V dans V . Un élément a de $GL(V)$ est, par définition, une application linéaire de V dans V qui a comme inverse a^{-1} .

Une représentation linéaire d'un groupe (G, \cdot) est un morphisme de groupe $\varphi : G \rightarrow GL(V)$. Notons que la représentation $\varphi : G \rightarrow GL(V)$ définit une action de groupe de G sur l'espace vectoriel V .

Toute représentation unitaire irréductible d'un groupe G abélien (commutatif) est de dimension 1, *i.e.* un morphisme de groupe de G dans S^1 (le groupe multiplicatif des complexes de module 1). Cela correspond précisément à la définition d'un caractère de groupe.

Dans le cas général (à condition que G soit unimodulaire), la transformée de Fourier abstraite d'une fonction $f \in L^2(G; \mathbb{C})$ (*i.e.* une fonction sur G de carré sommable) est définie par

$$\hat{f}(\varphi) = \int_G f(g)\varphi(g^{-1})d\nu(g) \quad (1.20)$$

où G est un groupe abélien additif, ν une mesure de Haar sur G et φ représente les morphismes correspondants aux caractères du groupe G .

Dans [Vilenkin 68], les caractères du groupe \mathbb{R}^m sont les morphismes de groupe

$$\begin{aligned} \varphi_{\mathbf{u}} : \mathbb{R}^m &\longrightarrow S^1 \\ \mathbf{x} &\longmapsto e^{i\langle \mathbf{x}, \mathbf{u} \rangle} \end{aligned} \quad (1.21)$$

où $\mathbf{x} = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$ et $\mathbf{u} = (u_1, u_2, \dots, u_m) \in \mathbb{R}^m$.

Pour des fonctions $f \in L^2(SO(2); \mathbb{C})$ périodiques, les caractères du groupe $SO(2)$ ($= \mathbb{R}/\mathbb{Z}$) doivent être étudiés. Ils sont donnés par

$$\begin{aligned} \varphi_u : SO(2) &\longrightarrow S^1 \\ \theta &\longmapsto e^{iu\theta} \end{aligned} \quad (1.22)$$

pour tout $u \in \mathbb{Z}$.

En appliquant l'équation 1.20 au cas $G = \mathbb{R}^m$ (équation 1.21), alors la transformée de Fourier usuelle est retrouvée. De même, en considérant le groupe $SO(2)$, les séries de Fourier sont obtenues.

En ce qui concerne le traitement des images en niveaux de gris, les signaux traités sont définis de \mathbb{R}^2 dans \mathbb{C} . Les caractères correspondants sont alors déduits des morphismes de l'équation 1.21 :

$$\begin{aligned} \varphi_{u_1, u_2} : \mathbb{R}^2 &\longrightarrow S^1 \\ (x_1, x_2) &\longmapsto e^{i(u_1 x_1 + u_2 x_2)} \end{aligned} \quad (1.23)$$

pour tout $x_1, x_2 \in \mathbb{R}$. En calculant la transformée de Fourier abstraite dans le cas où $G = \mathbb{R}^2$, la transformée de Fourier $2D$ est alors retrouvée. En terme d'action de groupe, la transformée de Fourier $2D$ correspond à l'action du groupe S^1 sur \mathbb{C} ou, par identification, à l'action de $SO(2)$ sur \mathbb{R}^2 .

A partir de la transformée de Fourier abstraite et de la théorie des groupes, la M_2 -transformée a été proposée par [Bornard 86].

1.1.4 La M_2 -transformée

Cette transformée a été définie dans l'objectif de définir des descripteurs invariants par l'action du groupe M_2 (groupe contenant les translations et les rotations dans le plan). Mais celle-ci reste peu appliquée car elle nécessite un fort investissement théorique. Dans cette partie, les principales étapes de construction de cette transformée sont données (pour plus de détails, voir [Bornard 86, Gourd 89]).

Dans [Bornard 86], Bornard *et al.* définissent une transformée de Fourier en utilisant le groupe $M_2 = \mathbb{R}^2 \rtimes SO(2)$ (où \rtimes est le produit semi-direct) des déplacements du plan dans le but de définir des descripteurs de Fourier invariants par l'action de ce groupe. Le groupe M_2 est constitué des rotations d'angle θ et des translations (x, y) dans le plan et il est noté (θ, x, y) . La loi interne du groupe M_2 s'écrit

$$(\theta_1, x_1, x_2) \cdot (\theta_2, x_2, y_2) = (\theta_1 + \theta_2, \cos(\theta_1)x_2 - \sin(\theta_1)y_2 + x_1, \sin(\theta_1)x_2 + \cos(\theta_1)y_2 + y_1) \quad (1.24)$$

Elle définit la composition des déplacements du plan. Notons que ce groupe n'est clairement pas abélien.

1.1.4.1 Définition

Dans ce cadre là, Bornard *et al.* [Bornard 86] interprètent une image $f(x_1, x_2)$ comme une fonction du groupe M_2 :

$$\begin{aligned} f : M_2 &\rightarrow \mathbb{R}^+ \\ (\theta, x_1, x_2) &\mapsto f(x_1, x_2) \end{aligned} \quad (1.25)$$

Notons que f est indépendante de θ .

La fonction f est une image, elle est donc intégrable sur \mathbb{R}^2 :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy < \infty \quad (1.26)$$

Donc, l'intégrale

$$\int_0^{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy d\theta < \infty \quad (1.27)$$

converge.

Comme le groupe M_2 n'est pas abélien, deux types de représentations irréductibles peuvent être considérés :

- des représentations de dimension 1.
- des représentations de dimension infinie, notées $T_r(\theta, x_1, x_2)$.

Cependant, ce sont les représentations de dimension infinie qui sont utilisées pour définir la M_2 -transformée. D'après [Smach 08], pour définir une transformée de Fourier (dans le cas des groupes non abéliens) raisonnablement, on ne peut considérer seulement les caractères (les représentations de dimension 1). Plus précisément, la mesure de Plancherel s'appuie, dans le cas du groupe M_2 , sur les représentations de dimension infinie. Dans ce cas, les caractères n'interviennent pas.

Les représentations $T_r(\theta, x_1, x_2)$ ont été calculées pour la première fois dans [Vilenkin 68]. Elles agissent sur des fonctions de carré intégrable sur le cercle S^1 , notées $\psi(u)$, comme

$$[T_r(\theta, x_1, x_2) \perp \psi](u) = e^{ir(x_1 \cos(u) + x_2 \sin(u))} \psi(u + \theta) \quad \forall \psi \in L^2(S^1) \quad (1.28)$$

où r est un réel strictement positif et \perp décrit l'action de $T_r(\theta, x_1, x_2)$ sur ψ .

En remplaçant les caractères de groupe par ces représentations dans la transformée de Fourier abstraite définie dans la partie 1.1.3, la M_2 -transformée s'écrit

$$\widehat{f}(r) = \int_{M_2} f(x_1, x_2) T_r^{-1}(\theta, x_1, x_2) dx_1 dx_2 d\theta \quad \forall r > 0 \quad (1.29)$$

Cette transformée peut être réécrite en utilisant l'équation 1.28 :

$$\begin{aligned} \left[\widehat{f}(r) \perp \psi \right] (u) &= \int_{M_2} f(x_1, x_2) [T_r^{-1}(\theta, x_1, x_2) \perp \psi](u) dx_1 dx_2 d\theta \quad \forall r > 0 \\ &= \int_{M_2} f(x_1, x_2) e^{-ir(x_1 \cos(u-\theta) + x_2 \sin(u-\theta))} \psi(u - \theta) dx_1 dx_2 d\theta \end{aligned} \quad (1.30)$$

Cette expression est alors passée en coordonnées polaires ($x_1 = \rho \cos(\omega)$ et $x_2 = \rho \sin(\omega)$) et en posant $\theta' = \theta - u$, l'équation suivante est obtenue :

$$\left[\widehat{f}(r) \perp \psi \right] (u) = \int_{M_2} f(\rho, \omega) e^{-ir\rho \cos(\theta' + \omega)} \psi(-\theta') \rho d\rho d\omega d\theta' \quad (1.31)$$

L'expression de $\hat{f}(r)$ est alors calculée dans la base $\{e^{in\theta}, n \in \mathbb{Z}\}$

$$\hat{f}(r)_n = \int_{M_2} f(\rho, \omega) e^{-ir\rho \cos(\theta'+\omega)} e^{-in\theta'} \rho d\rho d\omega d\theta' \quad (1.32)$$

$$= \int_{\rho, \omega} f(\rho, \omega) \left[\int_{\theta'} e^{-ir\rho \sin(\theta'+\omega+\frac{\pi}{2})} e^{-in\theta'} d\theta' \right] \rho d\rho d\omega$$

en posant $k = \theta' + \omega + \frac{\pi}{2}$ alors $\theta' = k - \omega - \frac{\pi}{2}$ et

$$= \int_{\rho, \omega} f(\rho, \omega) \left[\int_k e^{-ir\rho \sin(k)} e^{-in(k-\omega-\frac{\pi}{2})} dk \right] \rho d\rho d\omega$$

$$= 2\pi \int_{\rho, \omega} f(\rho, \omega) e^{in(\omega+\frac{\pi}{2})} \frac{1}{2\pi} \left[\int_k e^{-ir\rho \sin(k)-ink} dk \right] \rho d\rho d\omega$$

$$= 2\pi e^{in\frac{\pi}{2}} \int_{\rho, \omega} f(\rho, \omega) e^{in\omega} J_n(-r\rho) \rho d\rho d\omega \quad (1.33)$$

où $J_n(x)$ est la n-ième fonction de Bessel qui s'écrit

$$J_n(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n\theta - x \sin(\theta))} d\theta. \quad (1.34)$$

L'algorithme de calcul de $\hat{f}(r)_n$ est donné (à une constante près) dans [Bornard 86] :

Algorithme 1 Algorithme de la M_2 -transformée

Tache : Calculer la M_2 -transformée

Paramètres :

– f : une image en niveaux de gris

Sorties :

– $\hat{f}(r)_n$: la M_2 -transformée de l'image f

Algorithme :

// Calculer la transformée de Fourier \hat{f} de f à l'aide d'une FFT 2D

$$\hat{f}(u_1, u_2) = \int_{x_1, x_2} f(x_1, x_2) e^{-2i\pi(x_1 u_1 + x_2 u_2)} dx_1 dx_2$$

// Exprimer $\hat{f}(u_1, u_2)$ dans le domaine polaire (voir les équations 1.28, 1.30, 1.31)

$$\hat{f}(r, \theta) = \int_{\rho, \omega} f(\rho, \omega) e^{-2i\pi r\rho \cos(\theta+\omega)} \rho d\rho d\omega$$

// Calculer les séries de Fourier discrètes usuelles par rapport à θ (voir équation 1.32)

$$\hat{f}(r)_n = \sum_{\theta=0}^{2\pi} \hat{f}(r, \theta) e^{-2i\pi n\theta}$$

1.1.4.2 Propriétés

Translation et rotation L'effet d'une translation et d'une rotation est décrit par les représentations du groupe M_2 , c'est-à-dire $T_r(\theta, \rho, \omega)$ (voir équation 1.28). En choisissant la base $\{e^{in\theta}, n \in \mathbb{Z}\}$, elles s'écrivent

$$T_r(\theta, x_1, x_2) = e^{ir(x_1 \cos(u) + x_2 \sin(u))} e^{in(u+\theta)} \quad (1.35)$$

Si l'on calcule les séries de Fourier standard et que l'on passe en coordonnées polaires ($x_1 = \rho \cos(\omega)$ et $x_2 = \rho \sin(\omega)$), on obtient

$$\begin{aligned}
T_r(\theta, x_1, x_2) &= e^{ir(x_1 \cos(u) + x_2 \sin(u))} e^{in(u+\theta)} \\
&= \frac{1}{2\pi} \int_0^{2\pi} e^{ir(x_1 \cos(u) + x_2 \sin(u))} e^{in(u+\theta)} e^{-imu} du \\
T_r(\theta, \rho, \omega) &= \frac{1}{2\pi} \int_0^{2\pi} e^{ir(\rho \cos(\omega) \cos(u) + \rho \cos(\omega) \cos(u))} e^{in(u+\theta)} e^{-imu} du \\
&= \frac{1}{2\pi} e^{in\theta} \int_0^{2\pi} e^{ir\rho \cos(\omega-u)} e^{iu(n-m)} du \\
&= \frac{1}{2\pi} e^{in\theta} \int_0^{2\pi} e^{ir\rho \sin(\omega-u+\frac{\pi}{2})} e^{iu(n-m)} du \\
&\quad \text{en posant } \theta' = u - \omega + \frac{\pi}{2} \text{ alors } u = \theta' + \omega - \frac{\pi}{2} \text{ et} \\
&= \frac{1}{2\pi} e^{in\theta} \int_0^{2\pi} e^{ir\rho \sin(\theta')} e^{i(\theta'+\omega-\frac{\pi}{2})(n-m)} du \\
&= e^{i(n\theta+(\omega-\frac{\pi}{2})(n-m))} \frac{1}{2\pi} \int_0^{2\pi} e^{ir\rho \sin(\theta')} e^{i\theta'(n-m)} du \\
T_r(\theta, \rho, \omega)_{(m,n)} &= e^{i(n\theta+(n-m)(\omega-\frac{\pi}{2}))} J_{m-n}(r\rho) \tag{1.36}
\end{aligned}$$

Soit $g(x, y) = f((\theta_1, \rho_1, \omega_1) \perp (x_1, x_2))$ une image ayant subi une translation de vecteur ($\Delta_1 = \rho_1 \cos(\omega_1)$, $\Delta_2 = \rho_1 \sin(\omega_1)$) et une rotation d'angle θ_1 alors

$$\widehat{g}(r)_n = \widehat{f}(r)_n \perp T_r(\theta_1, \rho_1, \omega_1) \tag{1.37}$$

Changement d'échelle Soit $g(x_1, x_2) = f(kx_1, kx_2)$ alors

$$\widehat{g}(r)_n = \frac{1}{k^2} \widehat{f}\left(\frac{r}{k}\right)_n \tag{1.38}$$

Dans [Bornard 86], il est montré que l'expression $MD(r) = \sum_n \widehat{f}(r)_n \overline{\widehat{f}(r)_n}$ est invariante par l'action du groupe M_2 .

En comparaison avec la transformée de Fourier-Mellin, cette transformée ne nécessite ni de redéfinir le centre de l'image, ni d'estimer l'image log-polaire.

Cependant, cette transformée de Fourier nécessite de calculer les séries de Fourier sur un cercle. Dans le cas des images, cela implique une intégration des valeurs de la transformée de Fourier sur un cercle discret, amenant une difficulté au calcul de cette transformée.

1.1.5 La transformée de Fourier quaternionique de Bülow [Bülow 99]

Dans [Ell 92], la première transformée de Fourier quaternionique a été définie dans le but de généraliser la transformée de Fourier complexe. Plus récemment, Bülow [Bülow 99] a défini une autre transformée de Fourier pour les images en niveaux de gris à l'aide des quaternions (voir l'annexe A). L'objectif de cette transformée est d'étudier les symétries dans les signaux $2D$ (par exemple, une image en niveaux de gris). L'intérêt de l'utilisation des quaternions est d'obtenir dans le spectre quaternionique toutes les symétries spatiales possibles en $2D$. C'est ce que nous allons démontrer dans cette partie.

Commençons par rappeler le lien entre la transformée de Fourier 1D usuelle et les symétries du signal que l'on veut analyser [Bracewell 86]. En effet, il est bien connu que la transformée de Fourier d'un signal pair est paire alors que celle d'un signal impair est impaire. Nous pouvons alors décomposer un signal complexe en composantes paire f_p et impaire f_i et les lier aux composantes paire \hat{f}_p et impaire \hat{f}_i de la transformée de Fourier 1D. De plus, la transformée de Fourier d'un signal réel (respectivement imaginaire) pair est réelle (respectivement imaginaire) et paire. La transformée de Fourier d'un signal réel (respectivement imaginaire) impair est imaginaire (respectivement réelle) et impaire. Ces observations sont résumées dans le schéma suivant :

$$\begin{array}{ccccccc}
 f(x) = f_p(x) + f_i(x) & = & \Re(f_p(x)) + \Im(f_p(x)) & + & \Re(f_i(x)) + \Im(f_i(x)) \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 \hat{f}(u) = \hat{f}_p(u) + \hat{f}_i(u) & = & \Re(\hat{f}_p(u)) + \Im(\hat{f}_p(u)) & + & \Re(\hat{f}_i(u)) + \Im(\hat{f}_i(u))
 \end{array}$$

Dans ce schéma, $\Re(C)$ (respectivement $\Im(C)$) correspond à la partie réelle (respectivement imaginaire) d'un complexe C et

$$\begin{aligned}
 \hat{f}_p(u) &= \int_{\mathbb{R}} f(x) \cos(2\pi ux) dx \\
 \hat{f}_i(u) &= -i \int_{\mathbb{R}} f(x) \sin(2\pi ux) dx
 \end{aligned} \tag{1.39}$$

L'objectif de Bülow est alors d'utiliser une transformée de Fourier quaternionique afin d'étendre l'analyse des symétries aux signaux 2D et plus particulièrement aux images en niveaux de gris. Les quaternions possédant trois parties imaginaires et une partie scalaire, chacune de ces parties permet d'analyser un seul type de symétrie.

1.1.5.1 Définition

Une fonction $f(x_1, x_2)$, typiquement une image, peut être décomposée en composantes paires et impaires comme les signaux 1D. Dans ce cas, il faut considérer les symétries sur les deux axes x_1 et x_2 . La décomposition correspondante est alors

$$f(x_1, x_2) = f_{pp}(x_1, x_2) + f_{ii}(x_1, x_2) + f_{ip}(x_1, x_2) + f_{pi}(x_1, x_2) \tag{1.40}$$

La fonction $f_{pp}(x_1, x_2)$ (respectivement $f_{ii}(x_1, x_2)$) correspond alors à la partie de $f(x_1, x_2)$ qui est paire (respectivement impaire) le long des axes x_1 et x_2 . La fonction $f_{pi}(x_1, x_2)$ (respectivement $f_{ip}(x_1, x_2)$) est paire (respectivement impaire) le long de l'axe x_1 et impaire (respectivement paire) le long de l'axe x_2 .

Pour une fonction f de \mathbb{R}^2 dans \mathbb{H} , la transformée de Fourier quaternionique de Bülow est définie par

$$\hat{f}^q(\mathbf{u}) = \int_{\mathbb{R}^2} e^{-2\pi i x_1 u_1} f(\mathbf{x}) e^{-2\pi j u_2 x_2} dx_1 dx_2 \tag{1.41}$$

où i et j peuvent être remplacés par un quaternion pur arbitraire. Le choix de cette formule est justifié par l'égalité suivante :

$$\hat{f}^q(\mathbf{u}) = \hat{f}_{pp}^q(\mathbf{u}) - i \hat{f}_{ip}^q(\mathbf{u}) - j \hat{f}_{pi}^q(\mathbf{u}) + k \hat{f}_{ii}^q(\mathbf{u}) \tag{1.42}$$

où

$$\begin{aligned} \hat{f}_{pp}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} f(\mathbf{x}) \cos(2\pi u_1 x_1) \cos(2\pi u_2 x_2) d\mathbf{x} & \hat{f}_{ii}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} f(\mathbf{x}) \sin(2\pi u_1 x_1) \sin(2\pi u_2 x_2) d\mathbf{x} \\ \hat{f}_{ip}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} f(\mathbf{x}) \sin(2\pi u_1 x_1) \cos(2\pi u_2 x_2) d\mathbf{x} & \hat{f}_{pi}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} f(\mathbf{x}) \cos(2\pi u_1 x_1) \sin(2\pi u_2 x_2) d\mathbf{x} \end{aligned} \quad (1.43)$$

La fonction $\hat{f}_{pp}^q(\mathbf{u})$ (respectivement $\hat{f}_{ii}^q(\mathbf{u})$) est associée aux variations paires (respectivement impaires) du signal suivant les axes x_1 et x_2 . Les fonctions $\hat{f}_{ip}^q(\mathbf{u})$ et $\hat{f}_{pi}^q(\mathbf{u})$ quant à elles correspondent aux variations paires dans une dimension et impaires dans l'autre.

Une caractérisation de toutes les symétries spatiales est donc obtenue dans le domaine fréquentiel de la transformée de Fourier quaternionique et cela se résume par le schéma suivant :

$f_1(x) = \Re(f_{pp}(x)) + \Re(f_{ip}(x)) + \Re(f_{pi}(x)) + \Re(f_{ii}(x))$	$f_2(x) = \mathcal{I}(f_{pp}(x)) + \mathcal{I}(f_{ip}(x)) + \mathcal{I}(f_{pi}(x)) + \mathcal{I}(f_{ii}(x))$
$\hat{f}_1^q(u) = \mathcal{S}(\hat{f}_{pp}^q(u)) + \mathcal{I}(\hat{f}_{ip}^q(u)) + \mathcal{J}(\hat{f}_{pi}^q(u)) + \mathcal{K}(\hat{f}_{ii}^q(u))$	$\hat{f}_2^q(u) = \mathcal{I}(\hat{f}_{pp}^q(u)) + \mathcal{S}(\hat{f}_{ip}^q(u)) + \mathcal{K}(\hat{f}_{pi}^q(u)) + \mathcal{J}(\hat{f}_{ii}^q(u))$
$f_3(x) = \mathcal{J}(f_{pp}(x)) + \mathcal{J}(f_{ip}(x)) + \mathcal{J}(f_{pi}(x)) + \mathcal{J}(f_{ii}(x))$	$f_4(x) = \mathcal{K}(f_{pp}(x)) + \mathcal{K}(f_{ip}(x)) + \mathcal{K}(f_{pi}(x)) + \mathcal{K}(f_{ii}(x))$
$\hat{f}_3^q(u) = \mathcal{J}(\hat{f}_{pp}^q(u)) + \mathcal{K}(\hat{f}_{ip}^q(u)) + \mathcal{S}(\hat{f}_{pi}^q(u)) + \mathcal{I}(\hat{f}_{ii}^q(u))$	$\hat{f}_4^q(u) = \mathcal{K}(\hat{f}_{pp}^q(u)) + \mathcal{J}(\hat{f}_{ip}^q(u)) + \mathcal{I}(\hat{f}_{pi}^q(u)) + \mathcal{S}(\hat{f}_{ii}^q(u))$

Dans ce schéma, $\mathcal{S}(Q)$ (respectivement $\mathcal{I}(Q)$, $\mathcal{J}(Q)$, $\mathcal{K}(Q)$) correspond à la composante scalaire (respectivement i, j, k) d'un quaternion Q .

1.1.5.2 Propriétés

Translation Une translation induit un déphasage dans la transformée de Fourier quaternionique de Bülow.

Démonstration. Soit f, g deux fonctions de $L^2(\mathbb{R}^2, \mathbb{H})$, $\Delta = (\Delta_1, \Delta_2)^t$ contient les paramètres définissant une translation, et $g(\mathbf{x}) = f(\mathbf{x} - \Delta)$ alors

$$\begin{aligned} \hat{g}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} e^{-2\pi i x_1 u_1} f(\mathbf{x} - \Delta) e^{-2\pi j u_2 x_2} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{-2\pi i (x'_1 + \Delta_1) u_1} f(\mathbf{x}') e^{-2\pi j u_2 (x'_2 + \Delta_2)} dx'_1 dx'_2 \\ &= e^{-2\pi i \Delta_1} \left(\int_{\mathbb{R}^2} e^{-2\pi i x'_1 u_1} f(\mathbf{x}') e^{-2\pi j u_2 x'_2} dx'_1 dx'_2 \right) e^{-2\pi j \Delta_2} \\ &= e^{-2\pi i \Delta_1} \hat{f}^q(\mathbf{u}) e^{-2\pi j \Delta_2} \end{aligned} \quad (1.44)$$

avec $\mathbf{x} = (x_1, x_2)$ et $\mathbf{x}' = \mathbf{x} - \Delta$. □

En écrivant $\hat{f}^q(\mathbf{u})$ dans sa forme polaire, soit

$$\hat{f}^q(\mathbf{u}) = |\hat{f}^q(\mathbf{u})| e^{i\phi(\mathbf{u})} e^{k\psi(\mathbf{u})} e^{j\theta(\mathbf{u})} \quad (1.45)$$

on obtient la forme polaire de $\hat{g}^q(\mathbf{u})$:

$$\hat{g}^q(\mathbf{u}) = |\hat{f}^q(\mathbf{u})| e^{-2\pi i \Delta_1} e^{i\phi(\mathbf{u})} e^{k\psi(\mathbf{u})} e^{j\theta(\mathbf{u})} e^{-2\pi j \Delta_2} \quad (1.46)$$

Il est clair qu'une translation implique un déphasage de $\phi(\mathbf{u})$ ainsi que de $\theta(\mathbf{u})$.

Rotation Soit θ un angle de rotation et $g_\theta(x_1, x_2) = f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta))$ alors

$$\begin{aligned}\hat{g}_\theta^q(\mathbf{u}) &= \int_{\mathbb{R}^2} e^{-2\pi i x_1 u_1} f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta)) e^{-2\pi j u_2 x_2} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{-2\pi i u_1 (x'_1 \cos(\theta) - x'_2 \sin(\theta))} f(x'_1, x'_2) e^{-2\pi j u_2 (x'_1 \sin(\theta) + x'_2 \cos(\theta))} dx'_1 dx'_2\end{aligned}\quad (1.47)$$

avec $x'_1 = x_1 \cos(\theta) + x_2 \sin(\theta)$ et $x'_2 = -x_1 \sin(\theta) + x_2 \cos(\theta)$. Dans [Bülow 01], une forme simplifiée montrant les rotations qui sont impliquées dans le domaine fréquentiel est donnée. Soit la matrice de rotation A ,

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

alors,

$$\hat{g}_\theta^q(\mathbf{u}) = \frac{1}{2} (\hat{f}^q(A\mathbf{u}) + \hat{f}^q(A^{-1}\mathbf{u}) + i(\hat{f}^q(A\mathbf{u}) - \hat{f}^q(A^{-1}\mathbf{u}))j) \quad (1.48)$$

Homothétie Une homothétie de rapport k dans le domaine spatial provoque une homothétie de rapport $\frac{1}{k}$ dans le domaine fréquentiel de la transformée de Fourier quaternionique.

Démonstration. Si on contracte $f(x_1, x_2)$ dans la direction x_1 (c'est-à-dire $g(x_1, x_2) = f(kx_1, x_2)$) alors,

$$\begin{aligned}\hat{g}^q(\mathbf{u}) &= \int_{\mathbb{R}^2} e^{-2\pi i x_1 u_1} f(kx_1, x_2) e^{-2\pi j u_2 x_2} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{-2\pi i \frac{x'_1 u_1}{k}} f(x'_1, x_2) e^{-2\pi j u_2 x_2} dx'_1 dx_2 \\ &= \frac{1}{k} \hat{f}^q\left(\frac{u_1}{k}, u_2\right)\end{aligned}\quad (1.49)$$

avec $x'_1 = kx_1$. □

Cette transformée de Fourier quaternionique est utilisée plus particulièrement par Bülow dans le cadre de la définition du signal analytique quaternionique. Cela permet de définir les amplitudes et les phases instantanées de ce signal ainsi que des filtres de Gabor quaternioniques (pour plus de détails, voir [Bülow 99]).

Rappelons que cette transformée de Fourier quaternionique a été décrite après celle de Ell [Ell 92].

1.1.6 Discussion

Dans la première partie de ce chapitre, nous avons rappelé les transformées de Fourier pour les images en niveaux de gris de la plus classique (la transformée de Fourier usuelle) à la plus récente (la transformée de Fourier quaternionique de Bülow [Bülow 99]). Définies à partir de la transformée de Fourier usuelle, toutes ces transformées ont été construites chacune dans un objectif précis.

La transformée de Fourier-Mellin a été proposée pour définir des méthodes (extraction de descripteurs, corrélation de phase, etc.) invariantes en translation, rotation et en échelle. Cependant, nous avons relevé plusieurs difficultés pour le calcul de cette transformée comme la définition de l'origine de l'image log-polaire ainsi que l'interpolation de l'image dans le domaine log-polaire.

La M_2 -transformée a été proposée pour obtenir dans les images des caractéristiques invariantes en translation et rotation en utilisant la formulation de la transformée de Fourier abstraite. Elle ne nécessite pas l'estimation du centre de l'image mais elle implique une intégration sur un cercle discret dans le domaine fréquentiel.

Le but de la transformée de Fourier quaternionique de Bülow est différent. L'objectif de cette transformée est de mieux décrire les symétries dans les images, ce qui permet de réaliser des opérations complexes (comme des opérations de filtrage [Bülow 99] ou de détection de coins [Moya-Sánchez 10]) plus simplement qu'avec la transformée de Fourier usuelle.

Ces transformées sont définies pour les images en niveaux de gris. Il n'est donc pas possible d'utiliser directement ces transformées pour traiter des images couleur qui sont définies de \mathbb{R}^2 dans \mathbb{R}^3 . La prochaine partie concerne donc la généralisation des transformées de Fourier pour les images couleur.

1.2 Les transformées de Fourier pour les images couleur

Dans cette partie, la problématique est de définir une transformée de Fourier pour les images multicanaux en évitant un traitement marginal des différents canaux (par exemple *RVB*). Pour cela, nous étudierons deux méthodologies différentes qui ont été employées pour étendre la transformée de Fourier classique à la couleur. La première se base sur une extension des nombres complexes à l'aide de nombres hypercomplexes (les quaternions). La seconde est définie à l'aide des algèbres de Clifford, des actions et des représentations de groupe. Dans cette partie, nous faisons un état de l'art des transformées de Fourier couleur et nous explorons leurs propriétés concernant les transformations usuelles de l'image.

Nous commençons par étudier les différentes généralisations de la transformée de Fourier à l'aide des quaternions et des algèbres de Clifford.

1.2.1 Les généralisations de la transformée de Fourier

Nous avons abordé dans les parties précédentes les transformées de Fourier pour les images en niveaux de gris. Nous avons vu que ces transformées de Fourier ne sont définies que pour des fonctions $f \in L^2(\mathbb{R}^2, \mathbb{C})$. Ces définitions sont très restrictives. Elles ne permettent pas de traiter des signaux plus complexes comme les images couleur ou multi-spectrales (fonctions de $L^2(\mathbb{R}^2; \mathbb{R}^m)$, $m > 2$). La solution la plus immédiate et aussi la plus utilisée revient à calculer m transformées de Fourier indépendamment sur chaque canal de l'image. Cette "généralisation" marginale pose un certain nombre de problèmes. Par exemple, si l'on réalise une opération de convolution dans le domaine fréquentiel pour filtrer une image couleur, alors la méthode marginale fera apparaître des "fausses couleur" dans l'image résultante [Serra 09].

Certains auteurs ont alors proposé d'étendre la définition de la transformée de Fourier à l'aide des quaternions et des algèbres de Clifford (dont une introduction est proposée dans les annexes).

La première transformée de Fourier quaternionique a été définie par Ell en 1992 [Ell 92] et a été utilisée par Sangwine en 1996 [Sangwine 96] pour les images couleur. Puis, une seconde permettant de calculer une corrélation couleur a été proposée par Sangwine et Ell [Sangwine 00]. L'approche qui est proposée est de coder l'espace couleur RVB par l'espace des quaternions purs \mathbb{H}_0 . Cette transformée sera abordée dans la partie suivante.

Les quaternions sont un cas particulier de structures algébriques plus générales : les algèbres de Clifford. Ces algèbres sont principalement utilisées en physique. Dans ce cadre, Hestenes [Hestenes 66] est l'un des principaux contributeurs. Elles sont aussi utilisées en vision par ordinateur et en robotique [Sommer 01, Bayro-Corrochano 10, Dorst 07]. Cela apporte un cadre générique pour les traitements géométriques des images nD comprenant la segmentation ou la diffusion [Batard 09]. L'un des sujets principaux de ces approches est de construire rigoureusement des transformées de Fourier et de donner une nouvelle vision des précédents travaux concernant les transformées de Fourier quaternioniques.

Une définition d'une transformée de Fourier est proposée par Felsberg en utilisant les algèbres $\mathbb{R}_{2,0}$ et $\mathbb{R}_{3,0}$ pour introduire le concept de signal analytique $2D$ [Felsberg 02] :

$$\hat{f}_{\mathbf{e}_1\mathbf{e}_2}(\mathbf{u}) = \int_{\mathbb{R}^2} e^{-2\pi\mathbf{e}_1\mathbf{e}_2\langle\mathbf{u},\mathbf{x}\rangle} f(\mathbf{x}) d\mathbf{x} \quad (1.50)$$

pour une fonction $f(\mathbf{x}) = f(x\mathbf{e}_1) = f(x)\mathbf{e}_2$ de \mathbb{R} dans \mathbb{R} et

$$\hat{f}_{\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3}(\mathbf{u}) = \int_{\mathbb{R}^3} e^{-2\pi\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\langle\mathbf{u},\mathbf{x}\rangle} f(\mathbf{x}) d\mathbf{x} \quad (1.51)$$

pour une fonction $f(\mathbf{x}) = f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = f(x_1, x_2)\mathbf{e}_3$ de \mathbb{R}^2 dans \mathbb{R} . Le coefficient $\mathbf{e}_1\mathbf{e}_2$ (respectivement $\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3$) est le pseudo-scalaire de l'algèbre de Clifford $\mathbb{R}_{2,0}$ (respectivement $\mathbb{R}_{3,0}$). Cette transformée n'a été définie que pour un besoin technique : la définition du signal analytique.

Ebling et Scheuermann proposent dans [Ebling 05] une définition qui est principalement appliquée pour analyser les fréquences des champs de vecteurs et le comportement des filtres à valeurs vectorielles. Pour cela, le précédent noyau est utilisé pour les fonctions $f \in \mathbb{R}^3 \rightarrow \mathbb{R}_{3,0}$:

$$\hat{f}_{\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3}(\mathbf{u}) = \int_{\mathbb{R}^3} f(\mathbf{x}) e^{-2\pi\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\langle\mathbf{u},\mathbf{x}\rangle} d\mathbf{x} \quad (1.52)$$

Remarquons que si l'on note

$$\begin{aligned} f(\mathbf{x}) &= f_0(\mathbf{x}) + f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2 + f_3(\mathbf{x})\mathbf{e}_3 + f_{23}(\mathbf{x})I_3\mathbf{e}_1 + f_{31}(\mathbf{x})I_3\mathbf{e}_2 + f_{12}(\mathbf{x})I_3\mathbf{e}_3 + f_{123}(\mathbf{x})I_3 \\ &= (f_0(\mathbf{x}) + f_{123}(\mathbf{x})I_3) 1 + (f_1(\mathbf{x}) + f_{23}(\mathbf{x})I_3) \mathbf{e}_1 + (f_2(\mathbf{x}) + f_{31}(\mathbf{x})I_3) \mathbf{e}_2 + (f_3(\mathbf{x}) + f_{12}(\mathbf{x})I_3) \mathbf{e}_3 \end{aligned} \quad (1.53)$$

avec $I_3 = \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3$, cette transformée peut être écrite comme la somme de quatre transformées de Fourier complexes en identifiant I_3 à l'imaginaire complexe i . En particulier, pour une fonction

f ayant des valeurs seulement dans la partie vectorielle de l'algèbre de Clifford, cela revient à un calcul marginal.

Brackx *et al.* proposent dans [Brackx 06] des applications aux filtres de Gabor à partir d'une construction utilisant l'opérateur de Dirac Γ . La transformée de Fourier correspondante s'écrit

$$\hat{f}_{\pm}(\mathbf{u}) = \left(\frac{1}{\sqrt{2\pi}} \right)^n \int_{\mathbb{R}^n} e^{\mp i \frac{\pi}{2} \Gamma \mathbf{u}} \times e^{-i \langle \mathbf{u}, \mathbf{x} \rangle} f(\mathbf{x}) d\mathbf{x} \quad (1.54)$$

Dans le cas particulier de f de \mathbb{R}^2 dans $\mathbb{C}_2 = \mathbb{R}_{0,2} \otimes \mathbb{C}$ (la complexification de l'algèbre $\mathbb{R}_{0,2}$), le noyau peut être défini explicitement et

$$\hat{f}_{\pm}(\mathbf{u}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{\pm \mathbf{u} \wedge \mathbf{x}} f(\mathbf{x}) d\mathbf{x} \quad (1.55)$$

où $\exp(\pm \mathbf{u} \wedge \mathbf{x})$ est un spineur, c'est-à-dire l'exponentielle d'un bivecteur (voir l'annexe B) et \otimes est le produit tensoriel. Cette construction permet d'introduire les filtres bidimensionnels de Gabor Clifford (voir [Brackx 06] pour plus de détails).

Dans [Batard 09], Batard *et al.* proposent une nouvelle transformée de Fourier couleur à partir des actions des groupes $Spin(3)$ et $Spin(4)$. Cette transformée est étudiée plus particulièrement dans la partie 1.2.3 car elle joue un rôle central dans notre travail.

Par la suite, nous nous intéressons plus particulièrement aux transformées de Fourier couleur qui sont : la transformée de Fourier Clifford couleur ainsi que la transformée de Fourier quaternionique couleur, qui est le sujet de la partie suivante.

1.2.2 La transformée de Fourier quaternionique de Sangwine et Ell [Sangwine 00]

La transformée de Fourier quaternionique couleur est la première qui a été proposée pour étendre la transformée de Fourier classique aux images couleur. L'étude de ses propriétés permettra d'identifier les effets des transformations usuelles sur la transformée de Fourier quaternionique. Une interprétation générale et en termes de rotation sont données pour nous permettre de mieux appréhender la transformée de Fourier quaternionique.

1.2.2.1 Définitions

La première transformée de Fourier quaternionique est introduite par Ell dans [Ell 92] et utilisée pour les images couleur en 1996 par Sangwine [Sangwine 96]. Elle s'écrit

$$\hat{f}_{jk}(u_1, u_2) = \int_{\mathbb{R}^2} e^{-2j\pi u_1 x_1} f(x_1, x_2) e^{-2k\pi u_2 x_2} dx_1 dx_2 \quad (1.56)$$

Sangwine démontre dans [Sangwine 96] que cette équation peut s'appliquer aux images couleur, mais il indique que les tentatives d'implémentation de la convolution ou de la corrélation utilisant cette généralisation ont échoué.

En effet, une première difficulté est la non-commutativité de l'algèbre des quaternions (voir l'annexe A). La seconde difficulté exposée dans [Sangwine 97] est que cette transformée de Fourier ne possède pas une propriété importante permettant de calculer la corrélation : en effet, dans le cas de la transformée de Fourier classique, pour une image complexe $f(\mathbf{x})$ et son conjugué

$g(\mathbf{x}) = \overline{f(\mathbf{x})}$, l'égalité $\widehat{f}(\mathbf{x}) = \overline{\widehat{g}(-\mathbf{x})}$ est vérifiée. Cela permet de calculer une corrélation dans le domaine fréquentiel à l'aide du théorème de la "cross-correlation" de la transformée de Fourier qui s'écrit pour deux fonctions complexes f et g :

$$Corr(x) = \int_{-\infty}^{\infty} \overline{f(x)}g(x+y)dy = \left(\widehat{\overline{f(u)}}\widehat{g}(u)\right) \quad (1.57)$$

Dans le cas de la transformée de Fourier \widehat{f}_{jk} , cette propriété n'est plus valable [Sangwine 97].

C'est pourquoi une définition alternative a été proposée par Sangwine et Ell [Sangwine 00]. Comme il a été mentionné dans la partie 1.2.1, l'information couleur est codée à l'aide de quaternions purs. Une image couleur est alors considérée comme une fonction de \mathbb{R}^2 dans \mathbb{H}_0 , c'est-à-dire

$$f(x, y) = r(x, y)i + v(x, y)j + b(x, y)k \quad (1.58)$$

où $r(x, y)$, $v(x, y)$ et $b(x, y)$ sont les canaux rouge, vert et bleu de l'image.

Pour définir la transformée de Fourier quaternionique, Sangwine et Ell proposent de remplacer l'imaginaire complexe i par un quaternion pur et unitaire μ_1 . Il est facilement vérifié que $\mu_1^2 = -1$. Un choix typique pour μ_1 est $\mu_1 = (i + j + k)/\sqrt{3}$ qui correspond dans RVB à l'axe des niveaux de gris. La transformée de Fourier quaternionique dite à gauche $\widehat{f}_{\mu_1}^L$ et son inverse $\widehat{f}_{\mu_1}^{-L}$ sont données par la formule suivante :

$$\widehat{f}_{\mu_1}^{\pm L}(\mathbf{u}) = \int_{\mathbb{R}^2} e^{\mp\mu_1 2\pi\langle\mathbf{x},\mathbf{u}\rangle} f(\mathbf{x})d\mathbf{x} \quad (1.59)$$

Dans [Sangwine 00], une autre transformée semblable à celle définie précédemment mais dont l'exponentielle est placée à droite (la transformée de Fourier quaternionique à droite) est définie par

$$\widehat{f}_{\mu_1}^{\pm R}(\mathbf{u}) = \int_{\mathbb{R}^2} f(\mathbf{x})e^{\mp\mu_1 2\pi\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{x} \quad (1.60)$$

Nous verrons dans la partie 3.2.2.1 qu'une corrélation de phase couleur peut être définie à l'aide de ces deux transformées.

Pour implémenter les transformées de Fourier quaternioniques couleur, Ell et Sangwine [Ell 07] utilisent la décomposition symplectique des quaternions (voir l'annexe A). Elle est donnée pour une image $f \in \mathbb{H}_0$ par

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})\mu_2 \quad (1.61)$$

où $f_1(\mathbf{x})$ (respectivement $f_2(\mathbf{x})\mu_2$) correspond à une projection de l'image $f(\mathbf{x})$ dans le plan $(1, \mu_1)$ (respectivement $(\mu_2, \mu_1\mu_2)$).

Cette forme est très utile car elle permet de décomposer une image en une partie *parallèle* et une partie *orthogonale* relative à un quaternion pur μ_1 . Cela permet d'implémenter efficacement la transformée de Fourier quaternionique dans la partie 1.2.2.5.

Nous avons vu dans la partie dédiée à la transformée de Fourier usuelle que la multiplication d'une fonction complexe par une exponentielle complexe correspond à une rotation des valeurs de cette fonction. Dans la partie suivante, nous nous intéressons aux rotations qui sont réalisées dans la transformée de Fourier quaternionique. L'objectif est de les interpréter et de comprendre leur action dans le domaine fréquentiel.

1.2.2.2 Interprétation en termes de rotation

L'action d'un quaternion unitaire $u = e^{\mu\phi} = \cos(\phi) + \mu\sin(\phi)$ sur \mathbb{H} engendre une rotation $4D$ donnée par

$$q \mapsto uq = e^{\mu\phi}q \quad (1.62)$$

avec μ un quaternion unitaire et q un quaternion. La transformée de Fourier quaternionique (équation 1.59) montre clairement la multiplication de l'image f par l'exponentielle d'un quaternion unitaire et pur μ . Remarquons alors que la transformation de Fourier quaternionique réalise un ensemble de rotations. Ces rotations peuvent plus facilement être décrites si la décomposition symplectique de cette transformée est réalisée [Eli 07].

Pour la transformée de Fourier à gauche $\widehat{f}_{\mu_1}^L(\mathbf{u})$, cette décomposition s'écrit

$$\widehat{f}_{\mu_1}^L(\mathbf{u}) = \int_{\mathbb{R}^2} e^{-\mu_1 2\pi\langle \mathbf{x}, \mathbf{u} \rangle} f_1(\mathbf{x}) d\mathbf{x} + \int_{\mathbb{R}^2} e^{-\mu_1 2\pi\langle \mathbf{x}, \mathbf{u} \rangle} f_2(\mathbf{x}) \mu_2 d\mathbf{x} \quad (1.63)$$

Les rotations qui sont réalisées dans la partie parallèle s'écrivent

$$\begin{aligned} e^{-\mu_1\phi} f_1(\mathbf{x}) &= (\cos(-\phi) + \mu_1 \sin(-\phi))(f_{1,1}(\mathbf{x}) + \mu_1 f_{1,2}(\mathbf{x})) \\ &= (f_{1,1}(\mathbf{x}) \cos(-\phi) - f_{1,2}(\mathbf{x}) \sin(-\phi)) + \mu_1 (f_{1,1}(\mathbf{x}) \sin(-\phi) + f_{1,2}(\mathbf{x}) \cos(-\phi)) \end{aligned} \quad (1.64)$$

avec $\phi = 2\pi\langle \mathbf{x}, \mathbf{u} \rangle$.

Dans la partie parallèle, il est clair que les valeurs de $f_1(\mathbf{x})$ subissent des rotations anti-horaires d'angle $-\phi$ (des rotations horaires d'angle ϕ) dans le plan $(1, \mu_1)$. Pour le choix de $\mu_1 = (i + j + k)/\sqrt{3}$ et en identifiant le plan $(1, \mu_1)$ au plan complexe, cela correspond à une analyse fréquentielle de l'axe achromatique de l'image.

Concernant la partie orthogonale, les rotations suivantes sont considérées

$$\begin{aligned} e^{-\mu_1\phi} f_2(\mathbf{x}) \mu_2 &= (\cos(-\phi) + \mu_1 \sin(-\phi))(f_{2,1}(\mathbf{x}) + \mu_1 f_{2,2}(\mathbf{x})) \mu_2 \\ &= ((f_{2,1}(\mathbf{x}) \cos(-\phi) - f_{2,2}(\mathbf{x}) \sin(-\phi)) + \mu_1 (f_{2,1}(\mathbf{x}) \sin(-\phi) + f_{2,2}(\mathbf{x}) \cos(-\phi))) \mu_2 \\ &= \mu_2 (f_{2,1}(\mathbf{x}) \cos(-\phi) - f_{2,2}(\mathbf{x}) \sin(-\phi)) + \mu_1 \mu_2 (f_{2,1}(\mathbf{x}) \sin(-\phi) + f_{2,2}(\mathbf{x}) \cos(-\phi)) \end{aligned} \quad (1.65)$$

Comme pour la partie parallèle, $f_2(\mathbf{x}) \mu_2$ subit des rotations anti-horaires d'angle $-\phi$. Mais cette fois-ci, cette rotation a lieu dans le plan (μ_2, μ_3) où $\mu_3 = \mu_1 \mu_2$. Dans le cas où μ_1 est choisi comme l'axe achromatique (l'axe des niveaux de gris), le plan (μ_2, μ_3) correspond au plan de teinte. Nous obtenons alors une analyse fréquentielle dans le plan de chrominance de l'image.

Nous pouvons donc conclure que la transformée de Fourier quaternionique à gauche réalise des rotations dans deux plans orthogonaux, de même angle, et dans le même sens (anti-horaire). Ces rotations sont alors qualifiées d'isoclines [Lounesto 01]. Si l'on choisit $\mu_1 = (i + j + k)/\sqrt{3}$, alors une analyse luminance/chrominance est réalisée. Dans le cas de la transformée de Fourier quaternionique inverse, le même type de rotations agit mais l'angle est ϕ .

La différence entre les transformées de Fourier quaternioniques à gauche et à droite réside dans le sens des rotations qui sont effectuées dans la partie orthogonale. En effet, concernant la partie parallèle, l'exponentielle hypercomplexe et $f_1(\mathbf{x})$ commutent car leurs parties vectorielles sont colinéaires. Les rotations sont donc les mêmes pour la transformée de Fourier à droite et à gauche. Par contre, les rotations dans la partie orthogonale de la transformée de Fourier quaternionique à droite s'écrivent

$$\begin{aligned}
f_2(\mathbf{x})\mu_2e^{-\mu_1\phi} &= (f_{2,1}(\mathbf{x}) + \mu_1f_{2,2}(\mathbf{x}))\mu_2(\cos(-\phi) + \mu_1\sin(-\phi)) \\
&= (\mu_2f_{2,1}(\mathbf{x}) + \mu_1\mu_2f_{2,2}(\mathbf{x}))(\cos(-\phi) + \mu_1\sin(-\phi)) \\
&= \mu_2f_{2,1}(\mathbf{x})\cos(-\phi) + \mu_2\mu_1f_{2,1}(\mathbf{x})\sin(-\phi) + \\
&\quad \mu_1\mu_2f_{2,2}(\mathbf{x})\cos(-\phi) + \mu_1\mu_2\mu_1f_{2,2}(\mathbf{x})\sin(-\phi) \\
&= \mu_2(f_{2,1}(\mathbf{x})\cos(-\phi) + f_{2,2}(\mathbf{x})\sin(-\phi)) + \mu_1\mu_2(f_{2,1}(\mathbf{x})\cos(-\phi) - f_{2,2}(\mathbf{x})\sin(-\phi)) \\
&= e^{\mu_1\phi}f_2(\mathbf{x})\mu_2
\end{aligned} \tag{1.66}$$

La preuve est essentiellement basée sur le fait que deux quaternions purs et orthogonaux entre eux anti-commutent (voir l'annexe A).

La partie orthogonale de la transformée de Fourier quaternionique à droite effectue donc des rotations horaires d'angle $-\phi$ des valeurs de $f_2(\mathbf{x})\mu_2$. Les rotations réalisées dans la transformée de Fourier quaternionique à droite ne sont alors pas isoclines car les rotations réalisées dans la partie parallèle et dans la partie orthogonale n'ont pas le même sens de rotation (anti-horaire et horaire).

Maintenant que nous avons donné une interprétation aux rotations qui sont réalisées par les transformées de Fourier quaternioniques, nous nous intéressons aux propriétés de ces transformées concernant les transformations usuelles d'une image.

1.2.2.3 Propriétés de la transformée de Fourier quaternionique

Les effets de la translation, rotation et du changement d'échelle de l'image dans le domaine fréquentiel de la transformée de Fourier quaternionique sont analogues à celles de la transformée de Fourier classique.

Translation Une translation induit un déphasage dans la transformée de Fourier quaternionique.

Démonstration. Soit f, g deux fonctions de $L^2(\mathbb{R}^2, \mathbb{H}_0)$, $\Delta = (\Delta_1, \Delta_2)^t$ contient les paramètres définissant une translation, et $g(\mathbf{x}) = f(\mathbf{x} - \Delta)$ alors

$$\begin{aligned}
\widehat{g}_{\mu_1}^{\pm L}(\mathbf{u}) &= \int_{\mathbb{R}^2} e^{\mp\mu_1 2\pi\langle \mathbf{x}, \mathbf{u} \rangle} f(\mathbf{x} - \Delta) d\mathbf{x} \\
&= \int_{\mathbb{R}^2} e^{\mp\mu_1 2\pi\langle \mathbf{x}' + \Delta, \mathbf{u} \rangle} f(\mathbf{x}') d\mathbf{x}' \\
&= e^{\mp\mu_1 2\pi\langle \mathbf{u}, \Delta \rangle} \int_{\mathbb{R}^2} e^{\mp 2\mu_1 \pi\langle \mathbf{x}', \mathbf{u} \rangle} f(\mathbf{x}') d\mathbf{x}' \\
&= e^{\mp\mu_1 2\pi\langle \mathbf{u}, \Delta \rangle} \widehat{f}_{\mu_1}^{\pm L}(\mathbf{u})
\end{aligned} \tag{1.67}$$

avec $\mathbf{x} = (x_1, x_2)$ et $\mathbf{x}' = \mathbf{x} - \Delta$.

□

Rotation Une rotation d'angle θ dans le domaine spatial implique une rotation identique dans le domaine fréquentiel.

Démonstration. Soit $g(x_1, x_2) = f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta))$ alors

$$\begin{aligned}\widehat{g}_{\mu_1}^{\pm L}(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\mp 2\mu_1 \pi (u_1 x_1 + u_2 x_2)} f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta)) dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{\mp 2\mu_1 \pi [(u_1 \cos(\theta) + u_2 \sin(\theta))x'_1 + (-u_1 \sin(\theta) + u_2 \cos(\theta))x'_2]} f(x'_1, x'_2) dx'_1 dx'_2 \\ &= \widehat{f}_{\mu_1}^{\pm L}(u_1 \cos(\theta) + u_2 \sin(\theta), -u_1 \sin(\theta) + u_2 \cos(\theta))\end{aligned}\quad (1.68)$$

avec $x'_1 = x_1 \cos(\theta) + x_2 \sin(\theta)$ et $x'_2 = -x_1 \sin(\theta) + x_2 \cos(\theta)$. □

Homothétie Une homothétie de rapport k dans le domaine spatial provoque une homothétie de rapport $\frac{1}{k}$ dans le domaine fréquentiel de la transformée de Fourier quaternionique.

Démonstration. Si on contracte $f(x_1, x_2)$ dans la direction x_1 (c'est-à-dire $g(x_1, x_2) = f(kx_1, x_2)$) alors,

$$\begin{aligned}\widehat{g}_{\mu_1}^{\pm L}(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\mp 2\mu_1 \pi (u_1 x_1 + u_2 x_2)} f(kx_1, x_2) dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{\mp 2\mu_1 \pi (\frac{u_1 x'_1}{k} + u_2 x_2)} f(x'_1, x_2) \frac{1}{k} dx'_1 dx_2 \\ &= \frac{1}{k} \widehat{f}_{\mu_1}^{\pm L}\left(\frac{u_1}{k}, u_2\right)\end{aligned}\quad (1.69)$$

avec $x'_1 = kx_1$. □

Ces trois propriétés sont aussi valables pour la transformée de Fourier quaternionique à droite. Dans ce cas, l'exponentielle se trouve à droite de la fonction f .

Nous avons vu que la transformée de Fourier quaternionique dépend d'un paramètre μ_1 , un quaternion pur et unitaire. Celui-ci définit la direction d'analyse de cette transformée. Nous allons donc étudier quels sont les choix possibles décrits dans la littérature ainsi que la construction de la base $\{\mu_1, \mu_2, \mu_3\}$ qui dépend de μ_1 .

1.2.2.4 Choix de μ_1 et construction de la base $\{\mu_1, \mu_2, \mu_3\}$

Nous avons vu précédemment que le choix du paramètre μ_1 est très important. Il détermine l'analyse fréquentielle qui est réalisée par la transformée de Fourier quaternionique. La seule condition pour le choix de μ_1 est qu'il doit être un quaternion pur et unitaire. À notre connaissance, dans la littérature, le choix presque systématique est $\mu_1 = (i + j + k)/\sqrt{3}$ car il ne privilégie aucune composante [Ell 07, Moxey 03, Chen 10]. Comme nous l'avons vu précédemment, il en résulte une décomposition luminance/chrominance de l'image. La figure 1.7 illustre la décomposition symplectique de l'image de "la maison de Fourier couleur" avec le choix $\mu_1 = (i + j + k)/\sqrt{3}$. Il est clair que l'image couleur est décomposée en une image de luminance et une image de chrominance.

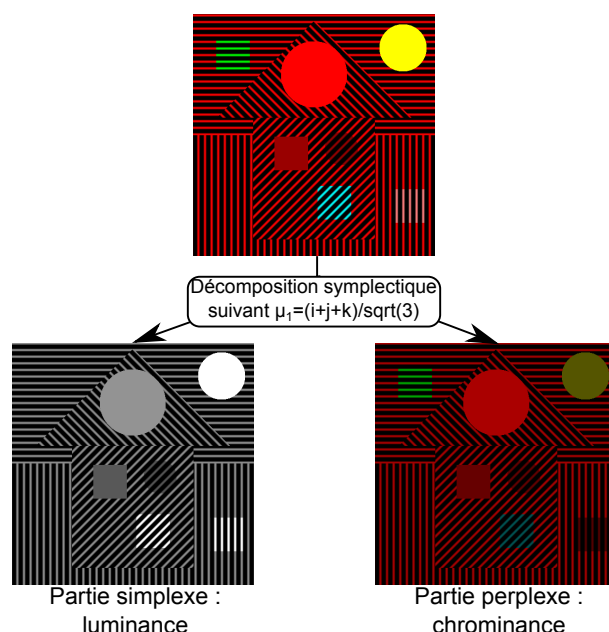


FIGURE 1.7 – *Partie simplexe (à gauche) et partie perplexe (à droite) d'une image selon le choix $\mu_1 = (i + j + k)/\sqrt{3}$*

Toutefois, la transformée de Fourier quaternionique n'est pas limitée à ce choix. D'autres choix comme $\mu_1 = i$ (couleur rouge), $\mu_1 = j$ (couleur verte) et $\mu_1 = k$ (couleur bleue) peuvent être faits. Dans [Guo 11], le choix $\mu_1 = \frac{\sqrt{2}}{2}(i + k)$ est pris pour calculer des moments de Fourier-Mellin quaternioniques. Cependant, ce choix n'est pas justifié par l'auteur.

Le choix de la direction d'analyse μ_1 est donc laissé à l'utilisateur. Une base $\{\mu_1, \mu_2, \mu_3\}$ doit alors être définie à partir de μ_1 pour calculer la transformée de Fourier quaternionique. Le quaternion μ_1 définit un plan dans un espace 3D qui est normal à μ_1 . Pour définir μ_2 dans ce plan, il faut commencer par choisir un quaternion pur q non parallèle à μ_1 . Le quaternion μ_2 correspond alors au produit vectoriel de μ_1 et q , qui est ensuite normalisé pour obtenir un quaternion unitaire. Soit μ_1 un quaternion unitaire et pur, μ_2 est alors calculé comme

$$\mu_2 = \frac{\mu_1 \times q}{|\mu_1 \times q|} = \frac{\mu_1 q - q \mu_1}{|\mu_1 q - q \mu_1|} \quad (1.70)$$

avec \times le produit vectoriel. Une fois μ_2 défini, le quaternion μ_3 est alors simplement $\mu_3 = \mu_1 \mu_2$. Comme μ_1 et μ_2 sont des quaternions purs, unitaires et orthogonaux entre eux, $\mu_3 = \mu_1 \times \mu_2$. L'image est ensuite projetée sur cette base et l'algorithme de la transformée de Fourier quaternionique utilisé.

A l'aide de cette base, nous allons définir dans la partie suivante l'algorithme de calcul rapide de la transformée de Fourier quaternionique.

1.2.2.5 Algorithme de calcul

Nous avons vu précédemment que la décomposition symplectique permet de décomposer un quaternion en deux plans orthogonaux. Ceux-ci s'intersectent seulement à l'origine d'un espace $4D$. Elle consiste donc à décomposer chaque pixel d'une image plongée dans \mathbb{H}_0 . Cela permet de réécrire la transformée de Fourier quaternionique de l'équation 1.59 comme

$$\widehat{f}_{\mu_1}^L(\mathbf{u}) = \widehat{f}_{1\mu_1}^L(\mathbf{u}) + \widehat{f}_{2\mu_1}^L(\mathbf{u})\mu_2 \quad (1.71)$$

avec

$$\widehat{f}_{i\mu_1}^L(\mathbf{u}) = \int_{\mathbb{R}^2} e^{-\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle} f_i(\mathbf{x}) d\mathbf{x} \quad (1.72)$$

pour $i \in \{1, 2\}$.

En suivant cette décomposition, $\widehat{f}_{\mu_1}^R(\mathbf{u})$ est définie par

$$\widehat{f}_{\mu_1}^R(\mathbf{u}) = \widehat{f}_{1\mu_1}^R(\mathbf{u}) + \widehat{f}_{2\mu_1}^R(\mathbf{u}) \quad (1.73)$$

avec

$$\begin{aligned} \widehat{f}_{1\mu_1}^R(\mathbf{u}) &= \int_{\mathbb{R}^2} f_1(\mathbf{x}) e^{-\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle} d\mathbf{x} = \int_{\mathbb{R}^2} e^{-\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle} f_1(\mathbf{x}) d\mathbf{x} = \widehat{f}_{1\mu_1}^L(\mathbf{u}) \\ \widehat{f}_{2\mu_1}^R(\mathbf{u}) &= \int_{\mathbb{R}^2} f_2(\mathbf{x}) \mu_2 e^{-\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle} d\mathbf{x} = \int_{\mathbb{R}^2} e^{\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle} f_2(\mathbf{x}) \mu_2 d\mathbf{x} = \widehat{f}_{2\mu_1}^{-L}(\mathbf{u}) \mu_2 \end{aligned} \quad (1.74)$$

Les équations $\widehat{f}_{i\mu_1}^L$ peuvent se calculer à l'aide de transformées de Fourier $2D$ rapides. En effet, le quaternion μ_1 est identifiable au complexe imaginaire i et $f_1(\mathbf{x})$ et $e^{-\mu_1 2\pi \langle \mathbf{x}, \mathbf{u} \rangle}$ commutent (ce sont deux quaternions dont les composantes vectorielles sont colinéaires).

À partir de cette observation, il est évident que $\widehat{f}_{1\mu_1}^R$ est identique à $\widehat{f}_{1\mu_1}^L$. Par contre, concernant la partie orthogonale, le signe de l'angle de l'exponentielle change quand l'exponentielle est passée à gauche de $f_2(\mathbf{x})$ (voir partie 1.2.2.2). Ce qui implique que $\widehat{f}_{2\mu_1}^R$ est égal à $\widehat{f}_{2\mu_1}^{-L}$ et donc qu'il faut appliquer une transformée de Fourier $2D$ inverse à f_2 pour calculer $\widehat{f}_{2\mu_1}^R$.

Pour calculer la transformée de Fourier à droite, l'algorithme est le même à la seule différence qu'il faut calculer une transformée de Fourier inverse sur $f_2(\mathbf{x})$. Une autre solution consiste à conjuguer $f_2(\mathbf{x})$ avant le calcul de la transformée de Fourier rapide et conjuguer le résultat. C'est une méthode classique très utilisée pour calculer une transformée de Fourier inverse à partir d'une transformée de Fourier usuelle.

L'algorithme pour calculer la transformée de Fourier quaternionique à gauche est alors

Algorithme 2 Algorithme de la transformée de Fourier quaternionique couleur à gauche

Tache : Calculer la transformée de Fourier quaternionique couleur

Paramètres :

– f : une image couleur

– μ_1 : un quaternion pur et unitaire (par défaut μ_1 est choisi comme l'axe achromatique *i.e.* $\mu_1 = (i + j + k)/\sqrt{3}$)

Sorties :

– $\widehat{f}_{\mu_1}^L$: la transformée de Fourier quaternionique à gauche de f

Algorithme :

// Décomposition de l'image selon μ_1 et la décomposition symplectique (équation 1.61)

$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})\mu_2$

// Développement de la partie simplexe et la partie perplexe

$f_1(\mathbf{x}) = s_{1,1}(\mathbf{x}) + s_{1,2}(\mathbf{x})\mu_1$

$f_2(\mathbf{x}) = s_{2,1}(\mathbf{x}) + s_{2,2}(\mathbf{x})\mu_1$

// où $s_{i,j}$ est une matrice de réels

// Identification de la partie simplexe et perplexe au plan complexe

$f'_1(\mathbf{x}) = s_{1,1}(\mathbf{x}) + s_{1,2}(\mathbf{x})i$

$f'_2(\mathbf{x}) = s_{2,1}(\mathbf{x}) + s_{2,2}(\mathbf{x})i$

// Calcul des transformées de Fourier 2D rapides de f'_1 et f'_2

$\widehat{f}'_1(\mathbf{x}) = \widehat{s}_{1,1}(\mathbf{x}) + \widehat{s}_{1,2}(\mathbf{x})i$

$\widehat{f}'_2(\mathbf{x}) = \widehat{s}_{2,1}(\mathbf{x}) + \widehat{s}_{2,2}(\mathbf{x})i$

// Identification de \widehat{f}'_1 et \widehat{f}'_2 au plan $(1, \mu_1)$

$\widehat{f}_1(\mathbf{x}) = \widehat{s}_{1,1}(\mathbf{x}) + \widehat{s}_{1,2}(\mathbf{x})\mu_1$

$\widehat{f}_2(\mathbf{x}) = \widehat{s}_{2,1}(\mathbf{x}) + \widehat{s}_{2,2}(\mathbf{x})\mu_1$

// Reconstruction de la transformée de Fourier quaternionique

$\widehat{f}_{\mu_1}^L(\mathbf{x}) = \widehat{f}_1(\mathbf{x}) + \widehat{f}_2(\mathbf{x})\mu_2$

1.2.2.6 Visualisation et interprétation générale

Le résultat de la transformée de Fourier quaternionique peut être vu comme une image hypercomplexe, *i.e.* une image où chaque pixel est un quaternion $q = a + bi + cj + dk$. Une méthode pour visualiser cette image est d'utiliser une forme polaire des quaternions :

$$q = |q|e^{\beta\Phi} = |q|(\cos\Phi + \beta\sin\Phi) \quad (1.75)$$

où β est un quaternion pur et unitaire et $0 \leq \Phi \leq \pi$. Le module d'un quaternion est égale à $|q| = q\bar{q}$ où \bar{q} est le conjugué de q qui s'écrit

$$\bar{q} = a - bi - cj - dk \quad (1.76)$$

Le logarithme du module, pour étaler la dynamique, est visualisé comme une image en niveaux de gris et s'écrit

$$M(q) = \frac{\log(1 + |q|)}{\log(1 + K)} \quad (1.77)$$

où $K = \max(|q|)$.

La phase est représentée par une image couleur dans un espace colorimétrique TSL . Dans cet espace, l'angle Φ est représenté par une teinte et s'écrit

$$\Phi = \text{atan} \left(\frac{|\mathcal{V}(q)|}{\mathcal{S}(q)} \right) \quad (1.78)$$

où $\mathcal{V}(q)$ (respectivement $\mathcal{S}(q)$) correspond à la partie vectorielle (respectivement scalaire) de q . L'angle Φ représente un angle de rotation autour de β dans le plan dont la normale est β . Une troisième quantité qui peut aussi être étudiée concerne le quaternion unitaire et pur noté β de l'équation 1.75. En effet, contrairement à la transformée de Fourier classique où l'imaginaire i reste constant, β ne l'est pas. Cette information est appelée *l'axe propre* ou encore plus simplement *l'axe* et il est égal à

$$\beta = \frac{\mathcal{V}(q)}{|\mathcal{V}(q)|} \quad (1.79)$$

L'axe est visualisé à l'aide de vecteurs unitaires dont l'origine est au centre du cube RVB . Il en résulte une image couleur (voir la figure 1.8).

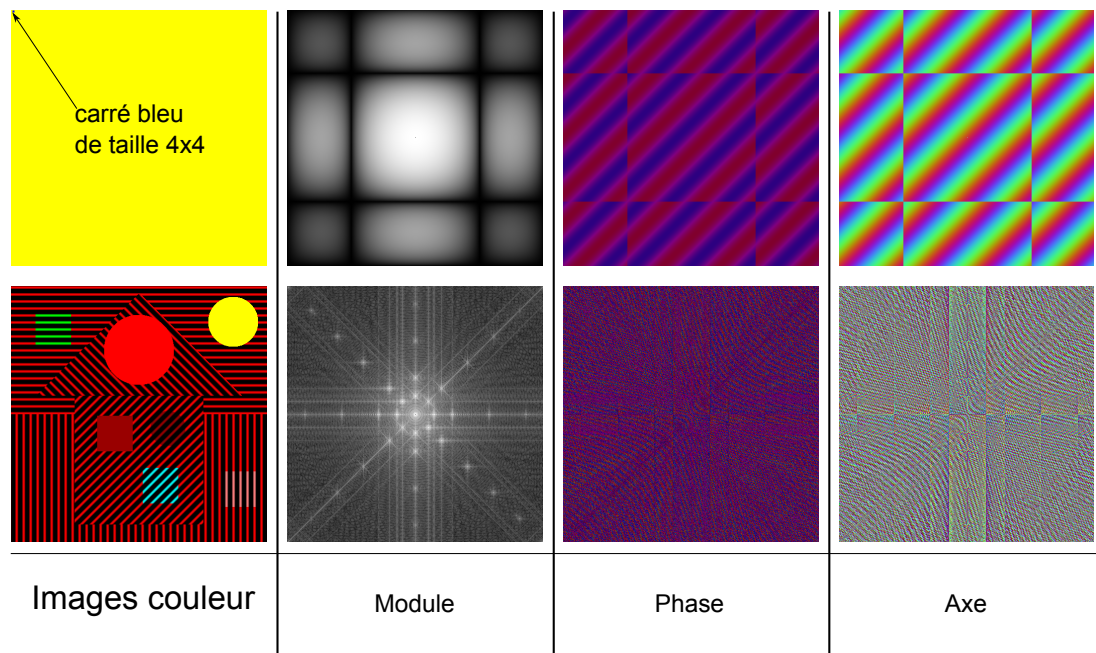


FIGURE 1.8 – *Transformée de Fourier quaternionique à gauche appliquée sur une image jaune contenant un carré bleu (en haut) et sur la maison de Fourier couleur (en bas).*

La première ligne reproduit une expérimentation que Ell et Sangwine ont menée dans [Ell 07]. Elle consiste à visualiser une image dont le fond est jaune avec un carré bleu de taille 4×4 .

La seconde ligne correspond à la visualisation de la transformée de Fourier quaternionique de l'image de la maison de Fourier couleur.

Concernant la première expérimentation, le carré bleu correspondant à une impulsion, le module est donc une image contenant un sinus cardinal. Pour l'image de “la maison de Fourier”, le module contient des pics représentatifs des fréquences présentes dans l'image originale, l'interprétation visuelle est donc analogue à celle réalisée avec la transformée de Fourier classique.

Concernant l'axe et la phase, leur interprétation visuelle est difficilement réalisable pour les deux images de la figure 1.8. Concernant l'image utilisée par Ell et Sangwine, les valeurs de la phase ainsi que de l'angle oscillent suivant la direction "Sud-Est". Ce résultat était attendu car cette direction correspond à l'axe sur lequel le carré bleu se trouve.

Dans [Denis 07], une expérimentation fait apparaître des pixels de couleur verte (angle de $\pi/2$ qui est la valeur limite de atan), dénotant une information de phase non pertinente.

Ell et Sangwine proposent une autre méthode pour interpréter la transformée de Fourier quaternionique : les chemins couleur. Ceux-ci permettent d'interpréter les rotations des valeurs des pixels dans la transformée de Fourier spatio-chromatique. Cette transformée est la transformée de Fourier 2D d'une image complexe dont les parties réelle et imaginaire correspondent aux canaux u et v d'une image CIE_Luv .

Cette notion de chemin couleur est étendue à la transformée de Fourier quaternionique dans [Ell 07] où une expérimentation est menée sur l'image de Lenna. Certaines fréquences de la transformée de Fourier quaternionique sont sélectionnées et une image couleur est reconstruite avec uniquement ces fréquences. Des observations comparables aux chemins couleur sont alors retrouvées.

Cette partie traite de la généralisation de la transformée de Fourier aux images couleur à l'aide des quaternions. Batard *et al.* ont défini une transformée de Fourier Clifford généralisant les transformées de Fourier quaternioniques de Bülow et de Sangwine et Ell. Celle-ci est traitée dans la partie suivante.

1.2.3 La transformée de Fourier Clifford couleur de Batard *et al.* [Batard 10]

Batard *et al.* [Batard 10] ont défini une transformée de Fourier pour les fonctions de $L^2(\mathbb{R}^m; \mathbb{R}^n)$. Cette transformée se différencie des autres transformées de Fourier couleur [Sangwine 00, Ell 92] car celle-ci s'appuie fortement sur la notion d'action de groupe et des algèbres de Clifford (voir l'annexe B). C'est précisément ce point de vue qui justifie la nécessité de choisir une direction d'analyse. Elle clarifie les relations entre la transformée de Fourier et l'action du groupe des translations à travers le groupe des spineurs. Il est démontré dans [Batard 10] que les généralisations précédemment proposées pour les images couleur sont des cas particuliers de cette définition. L'objectif de cette partie est de rappeler la transformée de Fourier Clifford couleur, de proposer un algorithme de calcul rapide, de discuter du choix de la direction d'analyse et d'étudier ses propriétés pour définir des descripteurs de Fourier couleur.

1.2.3.1 Définitions

Dans la partie 1.1.3, nous avons montré que la transformée de Fourier pour les fonctions complexes peut être généralisée par les actions de groupe et leur représentation.

Dans le cadre des algèbres de Clifford, remarquons qu'une fonction complexe $f = f_1(\mathbf{x}) + if_2(\mathbf{x})$ peut être plongée dans l'algèbre de Clifford $\mathbb{R}_{2,0}$

$$f(\mathbf{x}) = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2 \quad (1.80)$$

où $f_1(\mathbf{x})$ (respectivement $f_2(\mathbf{x})$) correspond à la partie réelle (respectivement imaginaire) de la fonction $f(\mathbf{x})$.

La transformée de Fourier d'une fonction $f \in \mathbb{C}$ peut alors se réécrire dans les algèbres de Clifford (à un facteur multiplicatif près) :

$$\begin{aligned}\widehat{f}(\mathbf{u}) &= \int_{\mathbb{R}^2} (f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2)e^{-\mathbf{e}_1\mathbf{e}_2\langle\mathbf{u},\mathbf{x}\rangle}d\mathbf{x} \\ &= \int_{\mathbb{R}^2} e^{\frac{1}{2}\mathbf{e}_1\mathbf{e}_2\langle\mathbf{u},\mathbf{x}\rangle}(f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2)e^{-\frac{1}{2}\mathbf{e}_1\mathbf{e}_2\langle\mathbf{u},\mathbf{x}\rangle}d\mathbf{x} \\ &= \int_{\mathbb{R}^2} [\cos(\langle\mathbf{u},\mathbf{x}\rangle/2)1 + \sin(\langle\mathbf{u},\mathbf{x}\rangle/2)\mathbf{e}_1\mathbf{e}_2](f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2) \\ &\quad [\cos(-\langle\mathbf{u},\mathbf{x}\rangle/2)1 + \sin(-\langle\mathbf{u},\mathbf{x}\rangle/2)\mathbf{e}_1\mathbf{e}_2]d\mathbf{x}\end{aligned}\quad (1.81)$$

où \langle , \rangle est le produit scalaire usuel et $(\mathbf{e}_1\mathbf{e}_2)^2 = -1$ le pseudo-scalaire de $\mathbb{R}_{2,0}$.

En utilisant le fait que $Spin(2) = \{a1 + b\mathbf{e}_1\mathbf{e}_2, a^2 + b^2 = 1\}$ est isomorphe au groupe S^1 des nombres complexes de module 1, on peut alors écrire la transformée de Fourier sous la forme

$$\widehat{f}(\mathbf{u}) = \int_{\mathbb{R}^2} (f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2)\perp\varphi_{\mathbf{u}}(-\mathbf{x})d\mathbf{x}\quad (1.82)$$

où $\varphi_{\mathbf{u}}$ est un morphisme de \mathbb{R}^2 dans $Spin(2)$ qui envoie \mathbf{x} dans $e^{\langle\mathbf{u},\mathbf{x}\rangle/2}(\mathbf{e}_1\mathbf{e}_2)$ et \perp décrit l'action $\mathbf{v}\perp s = s^{-1}\mathbf{v}s$ de $Spin(2)$ sur $\mathbb{R}_{2,0}^1$, et plus généralement de $Spin(n)$ vers $\mathbb{R}_{n,0}^1$ avec

$$Spin(n) = \{s \in \mathbb{R}_{n,0}^{(0)}, N(s) = 1, s^{-1}\mathbf{v}s \in \mathbb{R}_{n,0}^1, \forall \mathbf{v} \in \mathbb{R}_{n,0}^1\}\quad (1.83)$$

où $\mathbb{R}_{n,0}^{(0)}$ correspond à la sous-algèbre composée des éléments pairs $\mathbb{R}_{n,0}$ et $N(s) = s\bar{s}$. L'idée majeure qui prévaut pour cette transformée de Fourier est de remplacer l'action du groupe S^1 sur \mathbb{R}^2 par l'action de $spin(3)$ sur \mathbb{R}^3 , et plus généralement par $spin(n)$ sur \mathbb{R}^n .

Pour définir une transformée de Fourier Clifford couleur, il faut alors s'intéresser à l'action du groupe $Spin(3)$ sur \mathbb{R}^3 . Les morphismes de \mathbb{R}^2 dans $Spin(3)$ sont donnés par

$$(x_1, x_2) \mapsto e^{1/2(u_1x_1 + u_2x_2)B}\quad (1.84)$$

où u_1 et u_2 sont des réels et B est un bivecteur unitaire de $\mathbb{R}_{3,0}$. Les pixels de l'image couleur doivent alors être plongés dans l'algèbre $\mathbb{R}_{3,0}$, soit

$$\begin{aligned}f : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}_{3,0} \\ (x_1, x_2) &\mapsto r(x_1, x_2)\mathbf{e}_1 + v(x_1, x_2)\mathbf{e}_2 + b(x_1, x_2)\mathbf{e}_3\end{aligned}\quad (1.85)$$

où r, v et b correspondent aux canaux rouge, vert et bleu respectivement.

Mais, dans un souci d'implémentation (découpage de la transformée de Fourier Clifford en deux parties pour la calculer avec deux transformées de Fourier rapides), Batard *et al.* proposent de plonger l'image dans $\mathbb{R}_{4,0}$ et d'utiliser une partie du groupe $Spin(4)$, ce dernier étant isomorphe à $Spin(3) \times Spin(3)$. Les morphismes $\tilde{\Phi}_{u_1, u_2, B_1, u_3, u_4, B_2}$ de \mathbb{R}^2 dans $Spin(4)$ sont donnés par

$$(x_1, x_2) \mapsto e^{\frac{1}{8}(x_1(u_1+u_3)+x_2(u_2+u_4))(B_1+B_2+I(B_1-B_2))}e^{\frac{1}{8}(x_1(u_1-u_3)+x_2(u_2-u_4))(B_1-B_2+I(B_1-B_2))}\quad (1.86)$$

où $x_1, x_2, u_1, u_2, u_3, u_4 \in \mathbb{R}$, B_1 et B_2 sont des bivecteurs unitaires de $\mathbb{R}_{3,0}$ et I_4 est le pseudo-scalaire de $\mathbb{R}_{4,0}$.

Les détails du calcul des morphismes de \mathbb{R}^2 dans $Spin(4)$ (respectivement \mathbb{R}^2 dans $Spin(3)$) sont donnés dans [Batard 10]. Ces morphismes ont été calculés en utilisant le principe général de décomposition d'Hestenes et Sobczyk [Hestenes 84].

En notant $B = (1/4)(B_1 + B_2 + I_4(B_1 - B_2))$ et $I_4B = (1/4)(B_1 - B_2 + I_4(B_1 + B_2))$, le morphisme $\tilde{\Phi}_{u_1, u_2, B_1, u_3, u_4, B_2}$ qui est paramétré par deux bivecteurs unitaires de $\mathbb{R}_{4,0}$ est alors remplacé par le morphisme $\tilde{\Phi}_{u_1, u_2, u_3, u_4, B}$ qui ne compte plus qu'un paramètre B et qui s'écrit :

$$(x_1, x_2) \mapsto e^{\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))B} e^{\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))I_4B} \quad (1.87)$$

Il en découle une transformée de Fourier Clifford générale pour les signaux $4D$ qui est alors définie de la façon suivante :

Soit $f \in L^2(\mathbb{R}^2, (\mathbb{R}^4, Q))$ alors

$$\begin{aligned} \widehat{f}(u_1, u_2, u_3, u_4, B) &= \int_{\mathbb{R}^2} f(x_1, x_2) \perp \tilde{\Phi}_{u_1, u_2, u_3, u_4, B}(-x_1, -x_2) dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))B} e^{\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))I_4B} f(x_1, x_2) \\ &\quad e^{-\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))I_4B} e^{-\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))B} dx_1 dx_2 \end{aligned} \quad (1.88)$$

Cette transformée de Fourier est inversible, et son inverse notée \check{f} s'écrit

$$\check{f}(x_1, x_2) = \int_{\mathbb{R}^4 \times \mathbb{S}_{4,0}^2} f(u_1, u_2, u_3, u_4, B) \perp \tilde{\Phi}_{u_1, u_2, u_3, u_4, B}(x_1, x_2) dudvdwdzdv \quad (1.89)$$

où ν est une mesure unitaire sur les bivecteurs unitaires de $\mathbb{R}_{4,0}$. Remarquons que pour calculer les transformées inverses, on doit intégrer sur tous les bivecteurs B . Par la suite le bivecteur B est passé en paramètre de la transformée.

En plongeant les images couleur dans l'algèbre de Clifford $\mathbb{R}_{4,0}$, on obtient

$$f(x_1, x_2) = r(x_1, x_2)\mathbf{e}_1 + v(x_1, x_2)\mathbf{e}_2 + b(x_1, x_2)\mathbf{e}_3 + 0 \mathbf{e}_4 \quad (1.90)$$

L'équation 1.88 donne une formule générale pour la transformée de Fourier pour les signaux $4D$. Les images couleur étant à valeurs dans \mathbb{R}^3 , seule une partie des éléments de $Spin(4)$, ceux correspondants précisément aux morphismes de \mathbb{R}^2 dans $Spin(3)$, est utilisée :

$$\tilde{\Phi}_{u_1, u_2, 0, 0, B} : (x_1, x_2) \mapsto e^{\frac{1}{2}(u_1x_1+u_2x_2)(B+I_4B)} \quad (1.91)$$

Finalement, cette transformée de Fourier particulière est définie par

$$\begin{aligned} \widehat{f}_B(u_1, u_2) &= \int_{\mathbb{R}^2} f(x_1, x_2) \perp \tilde{\Phi}_{u_1, u_2, 0, 0, B}(-x_1, -x_2) dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1u_1+x_2u_2)B} e^{\frac{1}{2}(x_1u_1+x_2u_2)I_4B} f(x_1, x_2) e^{-\frac{1}{2}(x_1u_1+x_2u_2)I_4B} e^{-\frac{1}{2}(x_1u_1+x_2u_2)B} dx_1 dx_2 \\ &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1u_1+x_2u_2)(B+I_4B)} f(x_1, x_2) e^{-\frac{1}{2}(x_1u_1+x_2u_2)(B+I_4B)} dx_1 dx_2 \end{aligned} \quad (1.92)$$

Cette écriture est possible car B et I_4B sont des bivecteurs qui commutent. Ce cas particulier de la transformée de Fourier générale est appelée transformée de Fourier Clifford couleur. Cette transformée de Fourier couleur est inversible et son inverse s'écrit

$$\widetilde{f}_B(u_1, u_2) = \int_{\mathbb{R}^2} e^{-\frac{1}{2}(x_1u_1+x_2u_2)(B+I_4B)} f(x_1, x_2) e^{\frac{1}{2}(x_1u_1+x_2u_2)(B+I_4B)} dx_1 dx_2 \quad (1.93)$$

Une version alternative de la transformée de Fourier Clifford couleur, qui ne réalise pas les mêmes rotations, peut être définie. Cette transformée de Fourier Clifford couleur \widehat{f}'_B est analogue à la transformée de Fourier quaternionique à droite et comme pour cette transformée, les rotations qui sont réalisées ne sont pas isoclines :

$$\begin{aligned}
\widehat{f}'_B(u_3, u_4) &= \int_{\mathbb{R}^2} f(x_1, x_2) \perp \tilde{\Phi}_{0,0,u_3,u_4,B}(-x_1, -x_2) dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{\frac{1}{2}(-x_1 u_3 - x_2 u_4)I_4 B} f(x_1, x_2) e^{-\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{-\frac{1}{2}(-x_1 u_3 - x_2 u_4)I_4 B} dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{\frac{1}{2}(x_1 u_3 + x_2 u_4)(-I_4 B)} f(x_1, x_2) e^{-\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{-\frac{1}{2}(x_1 u_3 + x_2 u_4)(-I_4 B)} dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{-\frac{1}{2}(x_1 u_3 + x_2 u_4)I_4 B} f(x_1, x_2) e^{-\frac{1}{2}(x_1 u_3 + x_2 u_4)B} e^{\frac{1}{2}(x_1 u_3 + x_2 u_4)I_4 B} dx_1 dx_2
\end{aligned} \tag{1.94}$$

Dans le cadre des algèbres de Clifford, un vecteur peut être décomposé en une partie parallèle et une partie orthogonale par rapport à un bivecteur B (page 18 de [Hestenes 84]). Pour un vecteur \mathbf{v} et un bivecteur B , on a

$$\begin{aligned}
\mathbf{v} &= \mathbf{v} B B^{-1} = (\mathbf{v} \cdot B + \mathbf{v} \wedge B) B^{-1} \\
&= \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}
\end{aligned} \tag{1.95}$$

où

$$\begin{aligned}
\mathbf{v}_{\parallel} &= P_B(\mathbf{v}) \equiv (\mathbf{v} \cdot B) B^{-1} \\
\mathbf{v}_{\perp} &= P_B^{\perp}(\mathbf{v}) \equiv (\mathbf{v} \wedge B) B^{-1}
\end{aligned} \tag{1.96}$$

où $P_B(\mathbf{v})$ (respectivement $P_B^{\perp}(\mathbf{v})$) représente la projection parallèle (respectivement orthogonale) de \mathbf{v} sur un bivecteur B .

La fonction image f peut alors être décomposée en une partie parallèle $f_{\parallel B}$ (représentant la projection parallèle de f par rapport au bivecteur B) et une partie orthogonale $f_{\perp B}$ (représentant la projection orthogonale de f par rapport au bivecteur B).

L'équation 1.92 peut être réécrite selon cette décomposition :

$$\widehat{f}_B(\mathbf{u}) = \widehat{f}_{\parallel B}(\mathbf{u}) + \widehat{f}_{\perp B}(\mathbf{u}) \tag{1.97}$$

où

$$\begin{aligned}
\widehat{f}_{\parallel B}(\mathbf{u}) &= \int_{\mathbb{R}^2} f_{\parallel B}(x_1, x_2) \perp \tilde{\Phi}_{u_1, u_2, 0, 0, B}(-x_1, -x_2) dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} f_{\parallel B}(x_1, x_2) \perp e^{-\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} B} dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} B} f_{\parallel B}(\mathbf{x}) e^{-\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} B} d\mathbf{x} = \int_{\mathbb{R}^2} f_{\parallel B}(\mathbf{x}) e^{-\langle \mathbf{u}, \mathbf{x} \rangle B} d\mathbf{x} \\
\widehat{f}_{\perp B}(\mathbf{u}) &= \int_{\mathbb{R}^2} f_{\perp B}(\mathbf{x}) e^{-\langle \mathbf{u}, \mathbf{x} \rangle I_4 B} d\mathbf{x}
\end{aligned} \tag{1.98}$$

De même, l'équation 1.94 peut être réécrite selon cette décomposition :

$$\widehat{f'_B}(\mathbf{u}) = \widehat{f'_{\parallel B}}(\mathbf{u}) + \widehat{f'_{\perp B}}(\mathbf{u}) = \widehat{f_{\parallel B}}(\mathbf{u}) + \widehat{f_{\perp B}}(\mathbf{u}) \quad (1.99)$$

où

$$\widehat{f'_{\parallel B}}(\mathbf{u}) = \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} B} f_{\parallel B}(\mathbf{x}) e^{-\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} B} d\mathbf{x} = \int_{\mathbb{R}^2} f_{\parallel B}(\mathbf{x}) e^{-\langle \mathbf{u}, \mathbf{x} \rangle B} d\mathbf{x} \quad (1.100)$$

$$\widehat{f'_{\perp B}}(\mathbf{u}) = \int_{\mathbb{R}^2} e^{-\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} I_4 B} f_{\perp B}(\mathbf{x}) e^{\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} I_4 B} d\mathbf{x} = \int_{\mathbb{R}^2} f_{\perp B}(\mathbf{x}) e^{\langle \mathbf{u}, \mathbf{x} \rangle I_4 B} d\mathbf{x} \quad (1.101)$$

Plus tard, cette décomposition apportera une méthode pratique et efficace pour implémenter la transformée de Fourier Clifford couleur. Par la suite, nous n'utiliserons que la définition de la transformée de Fourier Clifford couleur $\widehat{f_B}$.

1.2.3.2 Interprétation en termes de rotation

Tout d'abord, rappelons que la transformée de Fourier Clifford générale de Batard *et al.* s'écrit

$$\widehat{f}(u_1, u_2, u_3, u_4, B) = \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))B} e^{\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))I_4 B} f(x_1, x_2) e^{-\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))B} e^{-\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))I_4 B} dx_1 dx_2$$

Il est clair que la transformée de Fourier générale implique des rotations générales (ou doubles rotations) dans \mathbb{R}^4 . Ces rotations agissent dans deux plans orthogonaux (générés par B et $I_4 B$) mais ne sont en général pas de même angle ($-\frac{1}{2}(x_1(u_1+u_3)+x_2(u_2+u_4))$ et $-\frac{1}{2}(x_1(u_1-u_3)+x_2(u_2-u_4))$).

Utilisées dans le cadre des images couleur, ces rotations pourraient amener d'autres propriétés mais le résultat de la transformée de Fourier générale est défini de \mathbb{R}^4 dans $\mathbb{R}_{4,0}^1$, ce qui la rend difficile à calculer et à traiter en pratique. En effet, l'espace mémoire pour stocker de cette transformée est important. Par exemple, pour une image de taille 128×128 , la taille de la transformée de Fourier Clifford générale correspondante a une taille de $128 \times 128 \times 4$. Un algorithme de calcul rapide doit aussi être implémenté pour calculer cette transformée.

La transformée de Fourier Clifford couleur est un cas particulier de la transformée de Fourier Clifford générale, elle s'écrit

$$\widehat{f_B}(u_1, u_2) = \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_1 + x_2 u_2)(B+I_4 B)} f(x_1, x_2) e^{-\frac{1}{2}(x_1 u_1 + x_2 u_2)(B+I_4 B)} dx_1 dx_2$$

Cette transformée implique quant à elle des rotations particulières dans \mathbb{R}^4 : des rotations isoclines [Lounesto 01]. Ces rotations sont de même angle $\theta = (x_1 u_1 + x_2 u_2)$ et agissent dans deux plans orthogonaux générés par B et $I_4 B$.

Le dernier type de rotation $4D$ correspond aux rotations simples qui agissent dans un seul plan invariant comme des rotations $2D$. Celles-ci sont considérées dans la partie 3.2.2 pour calculer une corrélation de phase couleur.

Dans l'objectif de définir dans le chapitre suivant des descripteurs couleur à partir de cette transformée, ses propriétés concernant la translation, la rotation et le changement d'échelle sont étudiées dans la prochaine partie.

1.2.3.3 Propriétés de la transformée de Fourier Clifford couleur

Comme pour la transformée de Fourier classique, il est important d'étudier les propriétés de la transformée de Fourier Clifford couleur en ce qui concerne la translation, la rotation et le changement d'échelle.

Translation Soit $f, g \in L^2(\mathbb{R}^2, \mathbb{R}_{4,0}^1)$ avec $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{\Delta})$, B un bivecteur unitaire dans $\mathbb{R}_{4,0}$ et $\mathbf{\Delta} = (\Delta_1, \Delta_2)^t$ le vecteur contenant les paramètres de la translation, alors

$$\widehat{g}_B(\mathbf{u}) = e^{\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} I_4 B} e^{\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} B} \widehat{f}_B(\mathbf{u}) e^{-\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} B} e^{-\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} I_4 B} \quad (1.102)$$

Démonstration. La preuve est essentiellement basée sur le fait que les rotations dans des plans orthogonaux commutent. Soit $\mathbf{x}' = \mathbf{x} - \mathbf{\Delta}$.

$$\begin{aligned} \widehat{g}_B(\mathbf{u}) &= \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} (B + I_4 B)} f(\mathbf{x} - \mathbf{\Delta}) e^{-\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{2} (B + I_4 B)} d\mathbf{x} \\ &= \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, (\mathbf{x}' + \mathbf{\Delta}) \rangle}{2} (B + I_4 B)} f(\mathbf{x}') e^{-\frac{\langle \mathbf{u}, (\mathbf{x}' + \mathbf{\Delta}) \rangle}{2} (B + I_4 B)} d\mathbf{x}' \\ &= \int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} e^{\frac{\langle \mathbf{u}, \mathbf{x}' \rangle}{2} (B + I_4 B)} f(\mathbf{x}') e^{-\frac{\langle \mathbf{u}, \mathbf{x}' \rangle}{2} (B + I_4 B)} e^{-\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} d\mathbf{x}' \\ &= e^{\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} \left(\int_{\mathbb{R}^2} e^{\frac{\langle \mathbf{u}, \mathbf{x}' \rangle}{2} (B + I_4 B)} f(\mathbf{x}') e^{-\frac{\langle \mathbf{u}, \mathbf{x}' \rangle}{2} (B + I_4 B)} d\mathbf{x}' \right) e^{-\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} \\ &= e^{\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} \widehat{f}_B(\mathbf{u}) e^{-\frac{\langle \mathbf{u}, \mathbf{\Delta} \rangle}{2} (B + I_4 B)} \end{aligned} \quad (1.103)$$

□

La translation de l'image dans le domaine spatial implique un déphasage de la transformée de Fourier Clifford couleur. L'équation (1.103) est plus compacte que l'équation (1.102), mais cette dernière souligne que deux rotations indépendantes agissent.

Rotation Une rotation dans le domaine spatial implique la même rotation dans le domaine fréquentiel.

Démonstration. Soit $f, g \in L^2(\mathbb{R}^2, \mathbb{R}_{4,0}^1)$ tel que $g(x_1, x_2) = f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) +$

$x_2 \cos(\theta)$) et θ un angle de rotation alors

$$\begin{aligned}
\widehat{g}_B(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_1 + x_2 u_2)(B+I_4 B)} f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta)) \\
&\quad e^{-\frac{1}{2}(x_1 u_1 + x_2 u_2)(B+I_4 B)} dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{1}{2}(u_1 x_1' \cos(\theta) - u_1 x_2' \sin(\theta) + u_2 x_1' \sin(\theta) + u_2 x_2' \cos(\theta))(B+I_4 B)} f(x_1', x_2') \\
&\quad e^{-\frac{1}{2}(u_1 x_1' \cos(\theta) - u_1 x_2' \sin(\theta) + u_2 x_1' \sin(\theta) + u_2 x_2' \cos(\theta))(B+I_4 B)} dx_1' dx_2' \\
&= \int_{\mathbb{R}^2} e^{-\frac{1}{2}([(u_1 \cos(\theta) + u_2 \sin(\theta))x_1' + (-u_1 \sin(\theta) + u_2 \cos(\theta))x_2'])(B+I_4 B)} f(x_1', x_2') \\
&\quad e^{-\frac{1}{2}([(u_1 \cos(\theta) + u_2 \sin(\theta))x_1' + (-u_1 \sin(\theta) + u_2 \cos(\theta))x_2'])(B+I_4 B)} dx_1' dx_2' \\
&= \widehat{f}_B(u_1 \cos(\theta) + u_2 \sin(\theta), -u_1 \sin(\theta) + u_2 \cos(\theta)) \tag{1.104}
\end{aligned}$$

avec $x_1' = x_1 \cos(\theta) + x_2 \sin(\theta)$ et $x_2' = -x_1 \sin(\theta) + x_2 \cos(\theta)$. □

Homothétie Une contraction de rapport k dans le domaine spatial s'exprime comme une dilatation dans le domaine fréquentiel.

Démonstration. Soit $f, g \in L^2(\mathbb{R}^2, \mathbb{R}_{4,0}^1)$ avec $g(x_1, x_2) = f(kx_1, x_2)$ alors,

$$\begin{aligned}
\widehat{g}_B(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(u_1 x_1 + u_2 x_2)(B+I_4 B)} f(kx_1, x_2) e^{-\frac{1}{2}(u_1 x_1 + u_2 x_2)(B+I_4 B)} dx_1 dx_2 \\
&= \int_{\mathbb{R}^2} e^{\frac{1}{2}(\frac{u_1 x_1'}{k} + u_2 x_2)(B+I_4 B)} f(kx_1, x_2) e^{-\frac{1}{2}(\frac{u_1 x_1'}{k} + u_2 x_2)(B+I_4 B)} \frac{1}{k} dx_1' dx_2 \\
&= \frac{1}{k} \widehat{f}_B\left(\frac{u_1}{k}, u_2\right) \tag{1.105}
\end{aligned}$$

avec $x_1' = kx_1$. □

1.2.3.4 Construction de B

Nous avons vu précédemment que le bivecteur B est un paramètre de la transformée de Fourier Clifford couleur. L'utilisateur a alors trois choix :

- L'utilisateur fournit directement un bivecteur unitaire. Dans ce cas, l'interprétation de la transformée de Fourier couleur devient difficile si l'on n'identifie pas le plan couleur qu'il génère. Pour le savoir, le bivecteur peut alors être factorisé comme le produit de vecteurs orthogonaux unitaires à l'aide d'un algorithme de factorisation [Fontijne 08].
- Il fournit un vecteur couleur \mathbf{c}_1 pour construire un bivecteur $B_1 = \mathbf{c}_1 \wedge e_4$.
- Il choisit deux vecteurs orthogonaux \mathbf{c}_2 et \mathbf{c}_3 . Le produit extérieur de ces deux vecteurs forment un bivecteur $B_2 = \mathbf{c}_2 \wedge \mathbf{c}_3$.

Il apparaît que les deux derniers choix sont analogues à un changement de signe près puisque la dualisation $I_4 B$ de B donne aussi un bivecteur de l'autre forme.

Démonstration.

Si $B_1 = \mathbf{c}_1 \wedge \mathbf{e}_4$, alors

$$\begin{aligned}
 B_1 &= c_{11}\mathbf{e}_1\mathbf{e}_4 + c_{12}\mathbf{e}_2\mathbf{e}_4 + c_{13}\mathbf{e}_3\mathbf{e}_4 \\
 I_4B_1 &= (\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4)(c_{11}\mathbf{e}_1\mathbf{e}_4 + c_{12}\mathbf{e}_2\mathbf{e}_4 + c_{13}\mathbf{e}_3\mathbf{e}_4) \\
 &= \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4c_{11}\mathbf{e}_1\mathbf{e}_4 + \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4c_{12}\mathbf{e}_2\mathbf{e}_4 + \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4c_{13}\mathbf{e}_3\mathbf{e}_4 \\
 &= -c_{13}\mathbf{e}_1\mathbf{e}_2 - c_{12}\mathbf{e}_3\mathbf{e}_1 - c_{11}\mathbf{e}_2\mathbf{e}_3
 \end{aligned} \tag{1.106}$$

Si $B_2 = \mathbf{c}_2 \wedge \mathbf{c}_3$, alors

$$\begin{aligned}
 B_2 &= (c_{21}\mathbf{e}_1 + c_{22}\mathbf{e}_2 + c_{23}\mathbf{e}_3) \wedge (c_{31}\mathbf{e}_1 + c_{32}\mathbf{e}_2 + c_{33}\mathbf{e}_3) \\
 &= (c_{21}c_{32} - c_{31}c_{22})\mathbf{e}_1\mathbf{e}_2 + (c_{31}c_{23} - c_{21}c_{33})\mathbf{e}_3\mathbf{e}_1 + (c_{22}c_{33} - c_{32}c_{23})\mathbf{e}_2\mathbf{e}_3 \\
 &= \alpha\mathbf{e}_1\mathbf{e}_2 + \beta\mathbf{e}_3\mathbf{e}_1 + \gamma\mathbf{e}_2\mathbf{e}_3 \\
 &\propto I_4B_1 \\
 I_4B_2 &= (\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4)(\alpha\mathbf{e}_1\mathbf{e}_2 + \beta\mathbf{e}_3\mathbf{e}_1 + \gamma\mathbf{e}_2\mathbf{e}_3) \\
 &= \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4\alpha\mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4\beta\mathbf{e}_3\mathbf{e}_1 + \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3\mathbf{e}_4\gamma\mathbf{e}_2\mathbf{e}_3 \\
 &= -\gamma\mathbf{e}_1\mathbf{e}_4 - \beta\mathbf{e}_2\mathbf{e}_4 - \alpha\mathbf{e}_3\mathbf{e}_4 \\
 &\propto B_1
 \end{aligned} \tag{1.107}$$

□

Par la suite, seul le choix $B = \mathbf{c} \wedge \mathbf{e}_4$ est considéré dans les expérimentations.

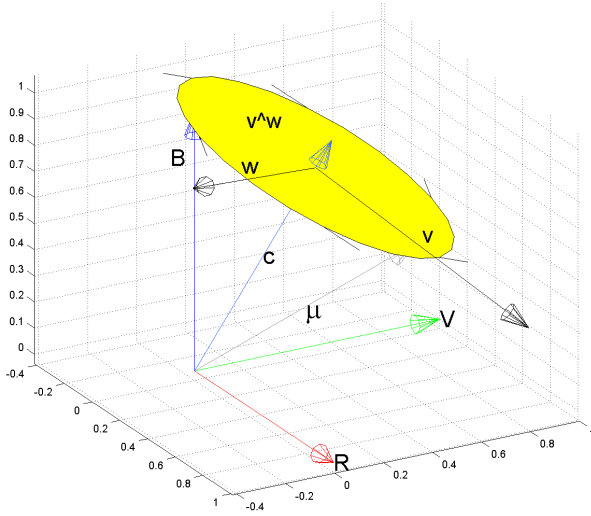
Remarquons que si \mathbf{c}_1 est choisi comme une couleur RVB , les vecteurs composant le bivecteur I_4B_1 ne sont pas forcément dans le cube RVB . De même, en choisissant deux vecteurs couleur \mathbf{c}_2 et \mathbf{c}_3 , les vecteurs composants le bivecteur I_4B_2 ne sont pas forcément dans le cube RVB . Par exemple, si $B_{\mathbf{c}_1} = \mathbf{e}_1 \wedge \mathbf{e}_4$, on peut construire le bivecteur $I_4B_{\mathbf{c}_1} = \mathbf{e}_2 \wedge \mathbf{e}_3$. Par contre, pour le choix $B_1 = (\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)/\sqrt{3}$, les vecteurs composant I_4B_1 ne sont forcément pas dans le cube RVB .

Un choix particulier $B_1 = \left(\frac{\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3}{\sqrt{3}} \wedge \mathbf{e}_4 \right)$ correspond à une analyse luminance-chrominance. On a alors une analyse fréquentielle analogue à celle de Sangwine *et al.* [Sangwine 00].

L'étape suivante pour le calcul de la transformée de Fourier Clifford couleur est de définir une base dans \mathbb{R}^4 à partir du bivecteur fourni par l'utilisateur.

1.2.3.5 Construction de la base $\{\mathbf{c}, \mathbf{e}_4, \mathbf{v}, \mathbf{w}\}$

La transformée de Fourier Clifford couleur peut être efficacement mise en oeuvre en utilisant deux transformées de Fourier rapides. Considérant que $\{\mathbf{c}, \mathbf{e}_4\}$ est une base pour le plan engendré par B , une base orthonormale $\{\mathbf{v}, \mathbf{w}\}$ pour le plan généré par I_4B doit être construite. Une solution possible est de choisir un vecteur unitaire $\boldsymbol{\mu}$ sans composante \mathbf{e}_4 et différent de \mathbf{c} . Le vecteur \mathbf{v} est défini comme la réjection de $\boldsymbol{\mu}$ sur \mathbf{c} et \mathbf{w} comme le vecteur orthogonal à \mathbf{v} dans le sous-espace représenté par le bivecteur I_4B (voir figure 1.9).



$$\mathbf{v} = -\mathbf{c}(\mathbf{c} \wedge \boldsymbol{\mu})^{-1} = (\boldsymbol{\mu} \wedge \mathbf{c})\mathbf{c}^{-1}$$

$$\mathbf{w} = -\mathbf{v}(I_4 B)^{-1} = \mathbf{v}(I_4 B)$$

FIGURE 1.9 – Illustration de la base $\{\mathbf{c}, \mathbf{e}_4, \mathbf{v}, \mathbf{w}\}$ de I_4 en utilisant la librairie GABLE [Dorst 02]

1.2.3.6 Algorithme de calcul

Nous proposons une mise en oeuvre de la transformée de Fourier Clifford couleur ainsi que de son algorithme de calcul.

Définition 1.2.3.1. Soit une fonction f , un bivecteur B et deux vecteurs \mathbf{c} et \mathbf{v} orthogonaux entre eux (définis comme ci-dessus), f peut être décomposée comme

$$f(\mathbf{x}) = \mathbf{c}[(f(\mathbf{x}) \cdot \mathbf{c}) + (f(\mathbf{x}) \cdot \mathbf{c}B)B] + \mathbf{v}[(f(\mathbf{x}) \cdot \mathbf{v}) + (f(\mathbf{x}) \cdot \mathbf{v}I_4B)(I_4B)] \quad (1.108)$$

Démonstration.

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x})BB^{-1} \\ &= (f(\mathbf{x}) \cdot B)B^{-1} + (f(\mathbf{x}) \wedge B)B^{-1} \\ &= (f(\mathbf{x}) \cdot (\mathbf{c} \wedge \mathbf{c}B))B^{-1} + (f(\mathbf{x}) \cdot I_4B)(I_4B)^{-1} \\ &= ((f(\mathbf{x}) \cdot \mathbf{c}) \wedge \mathbf{c}B - \mathbf{c} \wedge (f(\mathbf{x}) \cdot \mathbf{c}B))B^{-1} + (f(\mathbf{x}) \cdot (\mathbf{v} \wedge \mathbf{v}I_4B))(I_4B)^{-1} \\ &= ((f(\mathbf{x}) \cdot \mathbf{c})\mathbf{c}BB^{-1} - (f(\mathbf{x}) \cdot \mathbf{c}B)\mathbf{c}B^{-1}) + ((f(\mathbf{x}) \cdot \mathbf{v}) \wedge \mathbf{v}I_4B - \mathbf{v} \wedge (f(\mathbf{x}) \cdot \mathbf{v}I_4B))(I_4B)^{-1} \\ &= ((f(\mathbf{x}) \cdot \mathbf{c})\mathbf{c} + (f(\mathbf{x}) \cdot \mathbf{c}B)\mathbf{c}B) + ((f(\mathbf{x}) \cdot \mathbf{v})\mathbf{v}I_4B(I_4B)^{-1} - (f(\mathbf{x}) \cdot \mathbf{v}I_4B)\mathbf{v}(I_4B)^{-1}) \\ &= \mathbf{c}((f(\mathbf{x}) \cdot \mathbf{c}) + (f(\mathbf{x}) \cdot \mathbf{c}B)B) + ((f(\mathbf{x}) \cdot \mathbf{v})\mathbf{v} + (f(\mathbf{x}) \cdot \mathbf{v}I_4B)\mathbf{v}(I_4B)) \\ &= \mathbf{c}[(f(\mathbf{x}) \cdot \mathbf{c}) + (f(\mathbf{x}) \cdot \mathbf{c}B)B] + \mathbf{v}[(f(\mathbf{x}) \cdot \mathbf{v}) + (f(\mathbf{x}) \cdot \mathbf{v}I_4B)(I_4B)] \end{aligned} \quad (1.109)$$

□

Chacun des deux crochets peut être identifié à un nombre complexe puisque $BB = I_4BI_4B = -1$. Les expressions $\mathbf{c}B$ et $\mathbf{v}I_4B$ sont des vecteurs de $\mathbb{R}_{4,0}$. Le calcul de $\widehat{f_B}$ est alors réduit au

calcul de deux transformées de Fourier usuelles d'une fonction réelle et d'une fonction complexe. Le calcul de la transformée de Fourier Clifford couleur peut alors être résumé par l'algorithme 3.

Algorithme 3 Algorithme de la transformée de Fourier Clifford couleur

Tache : Calculer la transformée de Fourier Clifford couleur

Paramètres :

- f : une image couleur
- B : un bivecteur unitaire (par exemple $B = \mathbf{c} \wedge \mathbf{e}_4$)

Sorties :

- \widehat{f}_B : la transformée de Fourier Clifford couleur de f pour un bivecteur B

Algorithme :

```
// Décomposition de l'image selon le bivecteur B
f(x) = f||B(x) + f⊥B(x)
// Développement de la partie parallèle et de la partie orthogonale
f||B(x) = c[(f(x) · c) + (f(x) · cB)B]
f⊥B(x) = v[(f(x) · v) + (f(x) · vI4B)(I4B)]
// Identification de la partie parallèle et orthogonale au plan complexe
f'||B(x) = c[s1(x) + s2(x)i]
f'⊥B(x) = v[s3(x) + s4(x)i]
// s1(x) = (f(x) · c), s2(x) = (f(x) · vI4B), s3(x) = (f(x) · v) et s4(x) = (f(x) · vI4B) sont des matrices de scalaires
// Calcul des transformées de Fourier 2D rapides de f'||B et de f'⊥B
widehat{f}'||B(x) = c[s1(x) + s2(x)i]
widehat{f}'⊥B(x) = v[s3(x) + s4(x)i]
// Identification de widehat{f}'||B et widehat{f}'⊥B aux plans générés par B et I4B
widehat{f}||B(x) = c[s1(x) + s2(x)B]
widehat{f}⊥B(x) = v[s3(x) + s4(x)I4B]
// Reconstruction de la transformée de Fourier Clifford couleur
widehat{f}_B(x) = widehat{f}'||B(x) + widehat{f}'⊥B(x)
```

En fonction de l'application voulue, il n'est pas toujours nécessaire de reconstruire \widehat{f}_B à partir de $\widehat{f}'_{||B}$ et $\widehat{f}'_{\perp B}$ (plus précisément, les deux fonctions dans les crochets dans l'équation 1.109). Dans le cas où la reconstruction est désirée, les propriétés suivantes de la fonction vectorielle \widehat{f}_B donnent un ensemble de quatre équations linéaires à quatre inconnues :

$$\left(\widehat{f}_{||B}(\mathbf{x})\right)_B = \mathbf{c} \left[\left(\widehat{f}_B(\mathbf{x}) \cdot \mathbf{c}\right) + \left(\widehat{f}_B(\mathbf{x}) \cdot \mathbf{c}B\right)B \right] \quad (1.110)$$

$$\left(\widehat{f}_{\perp B}(\mathbf{x})\right)_B = \mathbf{v} \left[\left(\widehat{f}_B(\mathbf{x}) \cdot \mathbf{v}\right) + \left(\widehat{f}_B(\mathbf{x}) \cdot \mathbf{v}I_4B\right)I_4B \right] \quad (1.111)$$

L'écriture de ce système d'équation est résumé dans la figure 1.10.

Cela nous amène à l'équation de reconstruction suivante

$$\widehat{f}_B(\mathbf{x}) = \alpha(\mathbf{u})\mathbf{c} + \beta(\mathbf{u})\mathbf{c}B + \xi(\mathbf{u})\mathbf{v} + \psi(\mathbf{u})\mathbf{v}I_4B \quad (1.112)$$

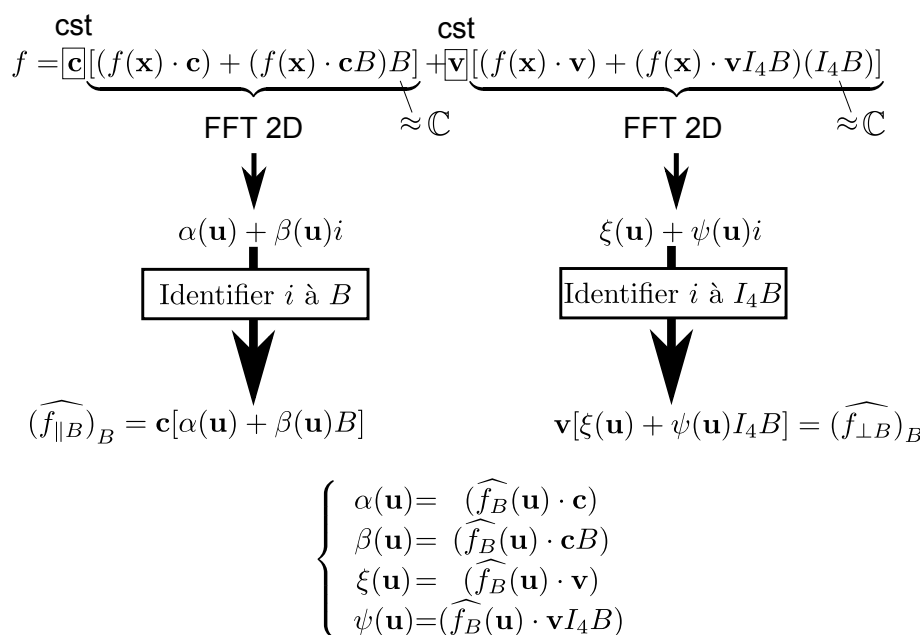


FIGURE 1.10 – Système d'équations permettant la reconstruction de la transformée de Fourier Clifford

1.2.3.7 Visualisation et interprétation générale

Dans cette partie, la visualisation et l'interprétation de cette transformée de Fourier est abordée. Deux solutions sont proposées : la première consiste à visualiser la transformée de Fourier Clifford couleur reconstruite et la seconde consiste à visualiser les parties parallèle et orthogonale séparément.

Visualisation de la transformée de Fourier Clifford couleur reconstruite

Les coefficients de la transformée de Fourier Clifford couleur reconstruite sont des vecteurs de $\mathbb{R}_{4,0}$. Une visualisation directe est donc difficile. Une solution est de visualiser sa magnitude comme une image en niveaux de gris. L'interprétation est alors analogue à celle du module de la transformée de Fourier usuelle (voir figure 1.11). Les pics présents dans l'image du module sont représentatifs des fréquences présentes dans l'image originale. La phase de la transformée de Fourier reconstruite n'est pas visualisable simplement car c'est un signal $4D$. Une autre solution est d'identifier chaque coefficient à un quaternion (voir partie 1.2.4) et une visualisation analogue à celle de la transformée de Fourier quaternionique (figure 1.8) peut alors être faite.

Une façon de visualiser les informations contenues dans le module et la phase couleur est de considérer la transformée de Fourier Clifford couleur inverse du module d'une part, et de la phase (les coefficients sont divisés par le module) d'autre part.

Dans la figure 1.12, l'image reconstruite avec le module de la transformée de Fourier Clifford couleur n'est pas facilement interprétable. Par contre, l'image reconstruite avec la phase semble

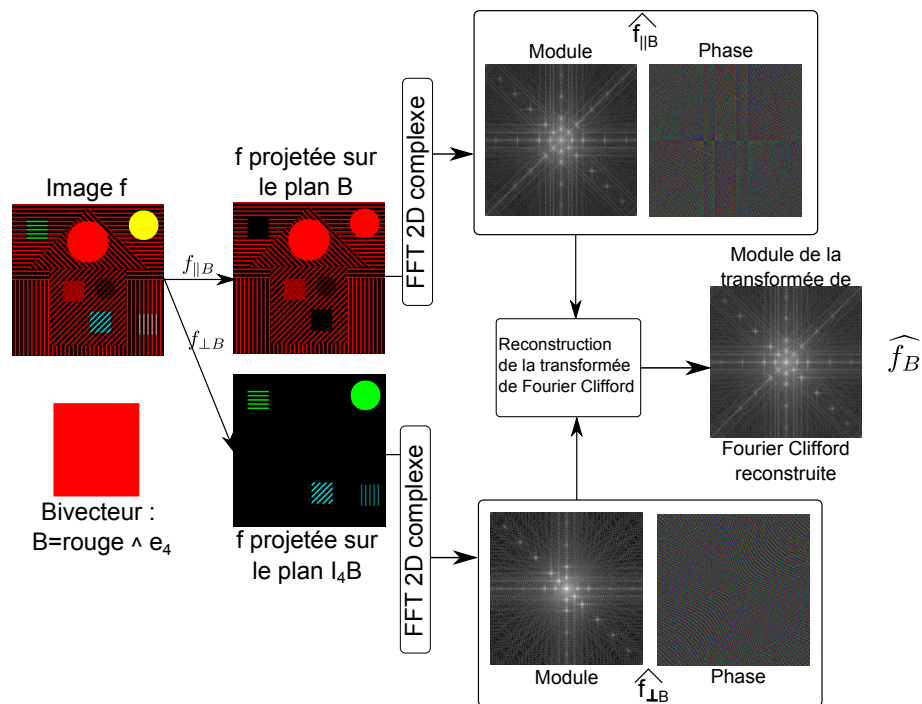


FIGURE 1.11 – Visualisation de la transformée de Fourier Clifford couleur avec le bivecteur $B_r = \text{rouge} \wedge e_4$

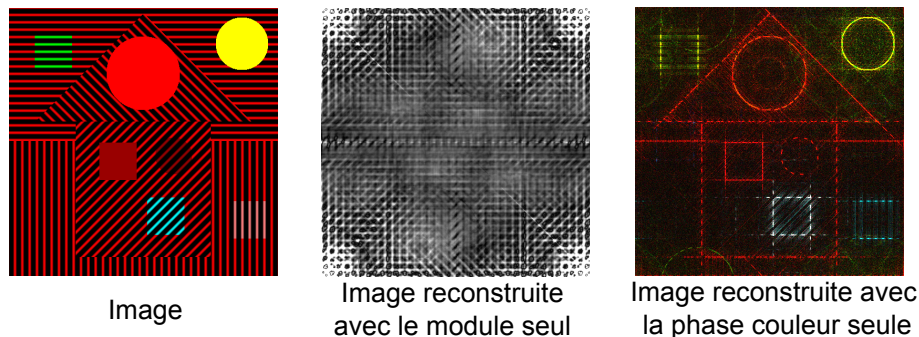


FIGURE 1.12 – Reconstruction de l'image à partir du module (à gauche) et de la phase (à droite) de la transformée de Fourier Clifford couleur

contenir majoritairement des contours couleur. Nous pouvons alors penser que la transformée de Fourier Clifford pourrait être utilisée dans le cadre de la détection de contours et de coins couleur.

Visualisation des parties parallèle et orthogonale de la transformée de Fourier Clifford couleur

Une autre solution est de considérer séparément les parties parallèle et orthogonale de la transformée de Fourier Clifford couleur. Deux modules et deux phases sont alors obtenues et concernent directement les projections parallèle et orthogonale de l'image originale. L'interpréta-

tion est alors analogue à celle d'images de phase et de module de la transformée de Fourier usuelle (voir partie 1.1.1.3). Cette visualisation permet aussi de réaliser plus facilement des filtres de la transformée de Fourier Clifford couleur par exemple [Batard 10].

Dans la figure 1.13, nous avons affiché les transformées de Fourier Clifford de la “maison de Fourier couleur” à plusieurs étapes de sa construction. Le bivecteur qui a été choisi est le bivecteur $B_r = \text{rouge} \wedge e_4$. On retrouve donc dans la partie parallèle les composantes de couleur rouge. Dans la partie orthogonale, nous pouvons remarquer qu'à chaque fois qu'une composante d'une couleur différente du rouge est ajoutée, les fréquences correspondantes se retrouvent dans la partie orthogonale. En ce qui concerne les transformées de Fourier Clifford reconstruites \widehat{f}_{B_r} , nous pouvons voir qu'elles sont une composition des parties parallèles et orthogonales.

1.2.4 Discussion

Les transformées de Fourier Clifford et quaternionique couleur ont un même objectif mais leur définition provient de théories mathématiques différentes.

Tout d'abord, la transformée de Fourier quaternionique a été définie en remplaçant le complexe imaginaire i dans l'exponentielle de la transformée de Fourier par un quaternion pur et unitaire μ_1 . Deux transformées de Fourier quaternioniques (à droite et à gauche) ont été rappelées. A notre connaissance, les applications de Sangwine et Ell considèrent uniquement le choix $\mu_1 = (i + j + k)/\sqrt{3}$, correspondant à une décomposition luminance-chrominance. Ces transformées peuvent être efficacement implémentées à l'aide de deux transformées de Fourier classiques. Concernant la visualisation, la décomposition polaire de Sangwine et Ell représente les quaternions avec un module, une phase et un axe.

La transformée de Fourier Clifford quant à elle s'appuie sur la théorie des représentations de groupe et les algèbres de Clifford. Elle généralise la transformée de Fourier quaternionique couleur [Batard 10]. Si dans l'équation 1.92, l'identification suivante est réalisée,

$$e_1 \leftrightarrow i \quad e_2 \leftrightarrow j \quad e_3 \leftrightarrow k \quad e_4 \leftrightarrow 1 \quad (1.113)$$

alors la rotation $4D$ générée par l'exponentielle d'un quaternion unitaire $e^{\mu\phi}$ sur \mathbb{H} peut être identifiée comme l'action du spineur $e^{-\frac{\phi}{2}(B+I_4B)}$ où B et I_4B sont deux bivecteurs unitaires orthogonaux. Par conséquent, l'action des morphismes de groupe $(x_1, x_2) \rightarrow e^{\mu(x_1u_1+x_2u_2)}$ de \mathbb{R}^2 dans \mathbb{H}_1 sur \mathbb{H} correspond à l'action des morphismes de groupe $(x_1, x_2) \rightarrow e^{-\frac{1}{2}(x_1u_1+x_2u_2)(B+I_4B)}$ de \mathbb{R}^2 dans $Spin(4)$ sur $\mathbb{R}_{4,0}^1$. La transformée de Fourier Clifford couleur définie dans [Batard 10] est analogue à $\widehat{f}_{\mu_1}^L$. La transformée \widehat{f}_B^L que nous proposons dans la partie 1.2.3.1 est analogue à $\widehat{f}_{\mu_1}^R$. Nous proposons un algorithme rapide pour calculer la transformée de Fourier Clifford couleur. Dans la partie 3, nous ne nous restreignons pas au choix $B_\mu = \frac{(i+j+k)}{\sqrt{3}} \wedge e_4$. Cela permet de choisir les plans sur lesquels on veut réaliser une analyse fréquentielle et par conséquent cela influe sur les résultats obtenus.

Enfin, la transformée de Fourier-Clifford générale peut être directement appliquée aux images de \mathbb{R}^2 dans \mathbb{R}^4 comme les images couleur+infrarouge ou autres images à 4 canaux en utilisant la forme générale (équation 1.88). De plus, les images à n-canaux pourraient être traitées à condition de calculer les morphismes de \mathbb{R}^2 dans $Spin(n)$.

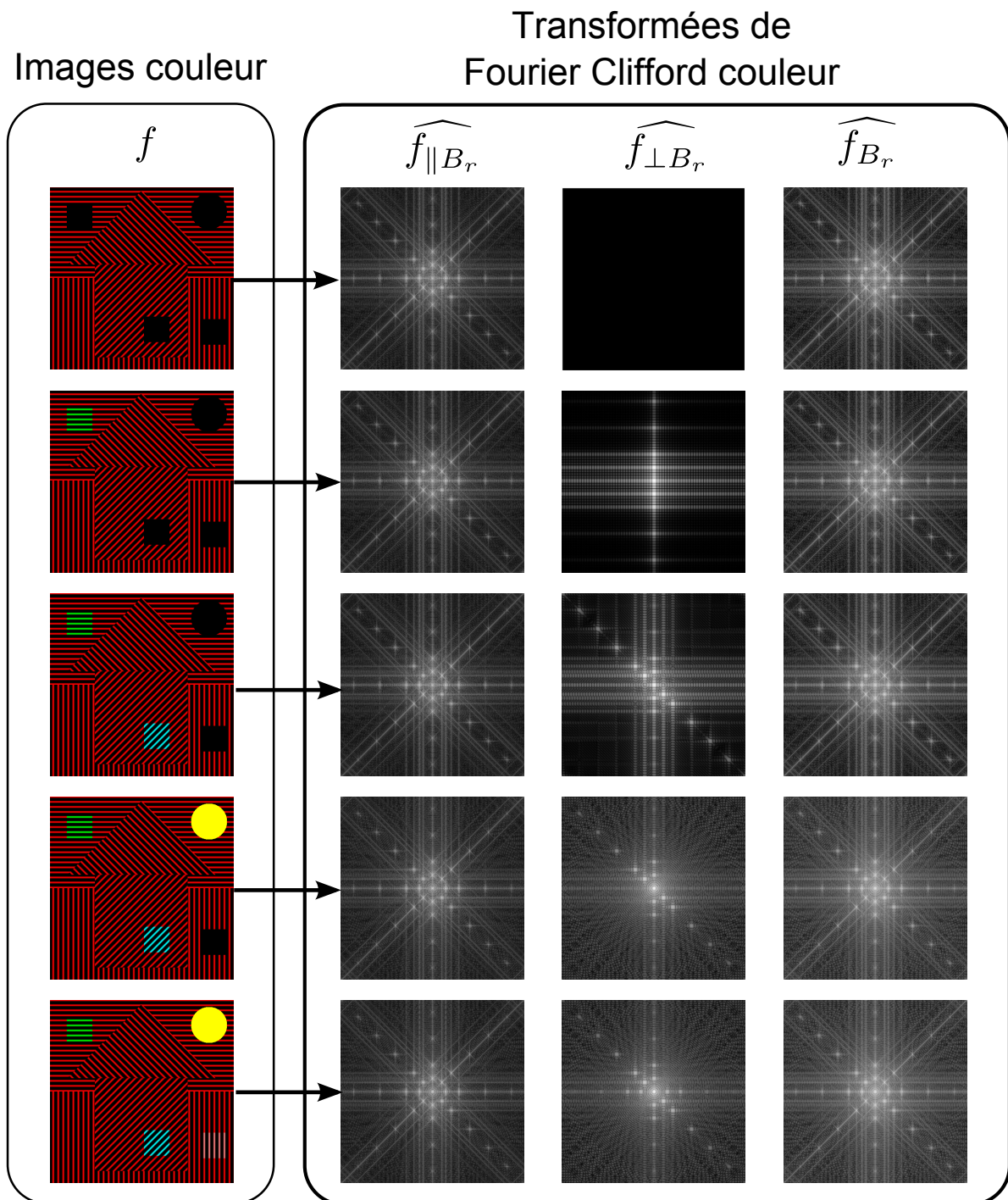


FIGURE 1.13 – Visualisation de la transformée de Fourier Clifford couleur à plusieurs étapes de la construction de la “maison de Fourier couleur”

1.3 Conclusion

Ce chapitre commence par un état de l'art original des transformées de Fourier pour les images en niveaux de gris. Nous observons que chacune d'elles est dérivée de la transformée de Fourier usuelle. Leur construction respective est motivée par un objectif précis : comme dans le cas de la transformée de Fourier-Mellin, construite pour l'obtention de caractéristiques invariantes en translation, rotation et échelle.

Motivés par le traitement d'images couleur de manière non marginale, nous nous sommes intéressés aux transformées de Fourier couleur. Par leurs approches géométriques, elles permettent de traiter l'information de couleur dans sa globalité. Ainsi, dans la deuxième partie de ce chapitre, deux transformées de Fourier couleur sont abordées : la transformée de Fourier quaternionique de Sangwine et Ell [Sangwine 00] et la transformée de Fourier Clifford de Batard *et al.* [Batard 10]. Le modèle mathématique de cette dernière généralise les transformées de Fourier quaternioniques couleur.

Suite à une définition de ces deux transformées, ce chapitre offre une première contribution en proposant une alternative à la transformée de Fourier Clifford couleur. Ceci permet l'obtention d'une forme équivalente à la transformée de Fourier quaternionique à droite.

L'étude de la transformée de Fourier Clifford couleur met en avant la présence d'un paramètre fondamental : le bivecteur B . Ce dernier définit une base dans \mathbb{R}^4 sur laquelle l'image couleur est projetée, correspondant ainsi à la direction d'analyse de cette transformée. Il apparaît clairement que son choix a un impact direct sur les coefficients de la transformée. Toutefois, à cette étape, nous pouvons nous demander quelle influence ce paramètre a sur les descripteurs que nous définirons. Ce verrou technique sera traité dans un chapitre consacré aux applications.

De part sa définition, la transformée de Fourier Clifford couleur est décomposable, dépendamment du choix de B , en une partie parallèle et une partie orthogonale. Suite à une identification au plan complexe, les deux parties peuvent être calculées à l'aide de deux transformées de Fourier usuelles. Cela permet de définir un nouvel algorithme qui réduit le temps de calcul.

La transformée de Fourier Clifford apparaît donc comme un outil prometteur dans le traitement d'images couleur pour la reconnaissance de formes. À l'aide de ses coefficients, nous pouvons nous demander comment construire des descripteurs pertinents. Ainsi, dans le chapitre suivant, nous allons nous intéresser aux descripteurs de formes niveaux de gris et couleur existants dans la littérature. Cet état de l'art nous permettra ainsi de définir des descripteurs de Fourier Clifford couleur.

D'autre part, une autre méthode dérivée de la transformée de Fourier pour la reconnaissance de forme sera abordée dans le chapitre applicatif : la corrélation de phase couleur.

CHAPITRE 2

LES DESCRIPTEURS DE FOURIER

Les descripteurs de Fourier

Sommaire

2.1	Propriétés générales des descripteurs	68
2.2	Les descripteurs de Fourier pour les images binaires	70
2.2.1	Les descripteurs de Fourier $1D$ usuels	71
2.2.1.1	Discussion	73
2.3	Les descripteurs de Fourier pour les images en niveaux de gris	74
2.3.1	Les descripteurs de Fourier génériques	74
2.3.1.1	Représentations polaires et transformée de Fourier	75
2.3.1.2	Obtention des invariances	76
2.3.1.3	Discussion	77
2.3.2	Les descripteurs de Fourier-Mellin	78
2.3.2.1	Les descripteurs de Fourier-Mellin classiques	78
2.3.2.2	Les descripteurs de Fourier-Mellin Analytiques	79
2.3.2.3	Discussion	80
2.3.3	Les descripteurs de Fourier généralisés	80
2.3.3.1	Définitions	80
2.3.3.2	Méthode de calcul	81
2.3.3.3	Discussion	82
2.3.4	Comparaison des descripteurs de Fourier	82
2.4	Les descripteurs de Fourier pour les images couleur	83
2.4.1	Méthode $1D$ utilisant des signatures dépendantes de la couleur	83
2.4.2	Méthode marginale $2D$	84
2.4.3	Méthode par projection de l'image sur un seul plan couleur	85
2.4.4	Les moments de Fourier-Mellin quaternioniques	86
2.4.5	De nouveaux descripteurs de Fourier couleur : GCFD1 et GCFD2	87
2.4.6	Un nouvel ensemble de descripteurs obtenu à partir de la transformée de Fourier Clifford reconstruite : les $GCFD3$	88
2.4.7	Comparaison des descripteurs couleur	90
2.5	Tableau récapitulatif	91
2.6	Conclusion	93

Dans le chapitre précédent, différentes transformées de Fourier niveaux de gris et couleur sont étudiées. Leurs propriétés respectives permettent d'obtenir un certain nombre d'invariants en traitant leurs coefficients.

Ainsi, dans le début de ce chapitre, un état de l'art original des descripteurs de Fourier est présenté pour chacune des transformées de Fourier définies précédemment (niveaux de gris et

couleur). Dans chacun des cas, les propriétés de ces descripteurs sont étudiées pour montrer leurs différentes invariances. De plus, un soin particulier est apporté à l'étude des avantages et des inconvénients.

Dans le cadre des images en niveaux de gris, l'étude faite sur les principaux descripteurs de Fourier permettra d'établir une comparaison inédite de ceux-ci. Cette confrontation est essentielle afin de mettre en évidence les mécanismes de constructions les plus pertinents. Ces derniers inspireront les nouveaux descripteurs de Fourier Clifford que nous proposerons.

Les descripteurs de Fourier généralisés de Smach *et al.* ont finalement retenu notre attention. Comme nous le verrons dans leur définition, ces descripteurs sont invariants en translation et en rotation par construction. Ils ne nécessitent donc pas de méthode *ad hoc* pour être calculés. Cette approche a guidé la définition de deux nouveaux ensembles de descripteurs de Fourier Clifford couleur. De la même manière que Smach *et al.* construisent deux ensembles de descripteurs de Fourier généralisés, nous définissons ces deux ensembles sur la transformée de Fourier Clifford.

Cette première approche intuitive permet d'obtenir des descripteurs à partir de la partie parallèle et orthogonale de la transformée de Fourier Clifford couleur. À cette étape, nous pouvons nous demander si la construction d'un nouvel ensemble de descripteurs obtenu à l'aide de la transformée de Fourier Clifford reconstruite permet l'obtention de descripteurs plus pertinents. Ainsi, nous définirons ce nouvel ensemble de descripteurs et la preuve des invariances sera démontrée. Ces trois ensembles de descripteurs seront comparés aux différents descripteurs de Fourier couleur de la littérature.

Pour clore ce chapitre, un tableau de synthèse original de tous les descripteurs abordés, permettra de mettre en avant les différences, tant du point de vue de la taille, que du coût algorithmique et des invariances.

2.1 Propriétés générales des descripteurs

L'objectif de la reconnaissance d'images est de reconnaître automatiquement une image quelle que soient les transformations qui lui sont appliquées. La définition de quantités réelles, appelées descripteurs, invariants par ces transformations est une étape cruciale du processus de reconnaissance. Ces descripteurs doivent respecter un certain nombre de propriétés.

Les invariances aux transformations géométriques Les transformations usuellement considérées sont les similitudes directes du plan (la translation, la rotation, l'homothétie), les transformations affines (qui préservent le parallélisme, les rapports de longueurs sur une droite, les coordonnées barycentriques) ou encore la réflexion. Ces transformations sont illustrées dans la figure 2.1. Par définition, si les descripteurs sont invariants par une transformation, alors leur valeur ne change pas si cette transformation est appliquée à l'image considérée.

Les descripteurs doivent respecter une ou plusieurs invariances à ces transformations pour être pertinents. Le besoin d'invariance dépend de l'application qui est concernée. En effet, une reconnaissance d'empreintes digitales ne nécessitera qu'une invariance en translation et en rotation. La taille de l'empreinte étant une caractéristique discriminante, les descripteurs ne doivent pas être invariants par changement d'échelle. La reconnaissance de personnes dans des vidéos de surveillance devra considérer des transformations plus générales comme les transformations affines par exemple.

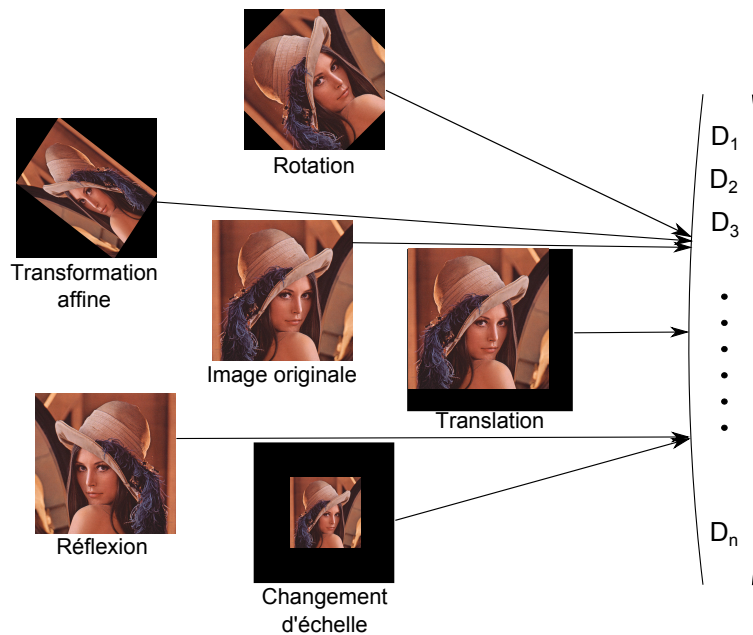


FIGURE 2.1 – Illustration de différentes transformations de l'image de Lenna

Les invariances photométriques Les invariances photométriques, contrairement aux invariances géométriques, sont définies sur les valeurs des pixels de l'image. L'objectif est d'obtenir une invariance à une modification de l'illuminant. Dans ce cadre, les changements de la luminance ou de la couleur des images n'influent pas sur la valeur des descripteurs [Zickler 08].

La complétude Un ensemble de descripteurs est qualifié de complet s'il décrit de manière unique une image pour l'ensemble des transformations prises en compte. Cela permet de discriminer totalement les images et de pouvoir les reconstruire parfaitement à partir de leurs seuls descripteurs. Dans [Smach 08], la notion de complétude faible est introduite. Elle signifie que les descripteurs définis dans cet article sont complets pour un large sous-ensemble de l'ensemble des images, ce qui est suffisant en général (pour plus de détails, voir [Smach 08]).

La vitesse de convergence La vitesse de convergence garantit une concentration de l'information concernant l'image dans les premiers descripteurs. Cela permet de réduire le nombre de

descripteurs nécessaires à la reconnaissance de formes. Dans ce cadre, cette propriété est primordiale car chaque descripteur supplémentaire ajoute une dimension en entrée de la méthode de classification. Dans le cadre de la classification, plus la dimensionnalité est élevée, plus le nombre d'images d'apprentissage doit être élevé (la malédiction de la dimension) [Duda 01]. De plus, le temps de calcul et l'espace mémoire utilisés sont aussi augmentés.

Le temps de calcul Le temps de calcul des descripteurs doit être pris en compte dans certaines applications, comme par exemple dans le cadre de la reconnaissance en temps réel.

Les coefficients de la transformée de Fourier respectent toutes ou partie de ces contraintes :

- Des invariants en translation, rotation et homothétie peuvent être obtenus à l'aide la transformée de Fourier-Mellin [Sheng 86b]. Des invariants en translation et en rotation peuvent être construits à partir de la M_2 -transformée [Smach 08].
- La complétude des descripteurs de Fourier-Mellin est prouvée dans [Ghorbel 93] et la complétude faible des descripteurs de Smach *et al.* dans [Smach 08].
- L'essentiel de l'information dans les différentes transformées de Fourier d'une image est contenu dans les basses fréquences, près de l'origine. Cela implique que seul un petit nombre de coefficients est nécessaire pour reconstruire convenablement l'image [Derrode 01] et donc pour reconnaître l'image.
- Les implémentations rapides des transformées de Fourier sont classiques et sont d'une complexité quasi-linéaire ($\mathcal{O}(n \log(n))$).

Dans ce cadre, nous étudions les descripteurs pour les images binaires dans la partie suivante.

2.2 Les descripteurs de Fourier pour les images binaires

La première définition des descripteurs de Fourier $1D$ apparaît au début des années 60 dans un article de Cosgriff [Cosgriff 60]. Ceux-ci sont calculés sur le contour d'un objet contenu dans une image binaire.

Une signature $1D$ est une représentation du contour qui peut être générée en utilisant différentes approches. L'une des plus simple à calculer est la distance entre le centre de la forme et chacun des points du contour. L'objectif est de représenter le contour comme une fonction $1D$ de longueur N (le nombre de points du contours) qui est *a priori* plus simple à traiter qu'un contour $2D$.

La transformée de Fourier $1D$ discrète de ces signatures peut être calculée. Les coefficients de Fourier obtenus sont alors traités pour obtenir des caractéristiques invariantes appelées descripteurs de Fourier. Il en découle de nombreuses déclinaisons où la signature diffère et/ou les coefficients subissent des traitements différents [Zahn 72, Lu 93].

Depuis les années 60, de nombreux types de descripteurs de Fourier ont été définis dans la littérature. Par exemple, dans [Mei 08], les signatures sont transformées dans une forme canonique (à l'aide d'une Analyse en Composantes Indépendantes ou d'une Analyse en Composantes Principales) afin d'obtenir une invariance aux transformations affines de l'objet.

Dans cette partie, nous nous intéressons aux descripteurs de Fourier $1D$ usuels parce qu'ils représentent la base des méthodes de Fourier pour la reconnaissance de formes et qu'ils apportent des idées applicables aux transformées de Fourier $2D$.

2.2.1 Les descripteurs de Fourier $1D$ usuels

La première étape du calcul des descripteurs de Fourier est l'extraction du contour de l'objet présent dans une image binaire. Les coordonnées cartésiennes du contour sont passées dans le plan complexe pour obtenir une signature 2π -périodique qui est classiquement utilisée dans la littérature [Gonzalez 01] :

$$s(k) = (x(k) - x_c) + (y(k) - y_c) i \quad (2.1)$$

où $k \in [1, N]$, N est le nombre de points du contour, $(x(k), y(k))$ sont les coordonnées spatiales du contour, (x_c, y_c) sont les coordonnées du centre de l'objet.

Les coefficients de Fourier $\hat{s}(k)$ sont obtenus en calculant la transformée de Fourier $1D$ discrète de la signature :

$$\hat{s}(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) e^{-2i\pi \frac{uk}{N}} \quad (2.2)$$

avec $s(k)$ le $k^{\text{ième}}$ point de la signature, N le nombre de points du contour et $u \in [-N/2 + 1, N/2]$. Après avoir défini les coefficients de Fourier de la signature $s(k)$, nous nous intéressons à leurs propriétés.

Translation Une translation de l'objet d'intérêt n'affecte que le coefficient $\hat{s}(0)$ (le centre de gravité de la forme).

Soit $s(k)$ la signature d'un objet et $s_t(k)$ la signature du même objet qui a subi une translation de vecteur $\mathbf{\Delta} = (\Delta_1, \Delta_2)^t$, *i.e.*, $s_t(k) = s(k) + (\Delta_1 + i\Delta_2)$, alors

$$\hat{s}_t(u) = \hat{s}(u) + (\Delta_1 + i\Delta_2)\delta(u) \quad (2.3)$$

où δ est une distribution de Dirac.

Le coefficient $\hat{s}(0)$ est donc dépendant de la translation. Pour obtenir l'invariance en translation, ce coefficient ne doit pas être pris en compte dans le vecteur de descripteurs.

Rotation La rotation de l'objet d'intérêt induit un déphasage des coefficients de Fourier de la signature.

Soit $s(k)$ la signature d'un objet et $s_r(k)$ la signature du même objet qui a subi une rotation d'angle θ , *i.e.* $s_r(k) = s(k)e^{i\theta}$, alors

$$\hat{s}_r(u) = \hat{s}(u)e^{i\theta} \quad (2.4)$$

On peut alors remarquer que les $|\hat{s}_r(u)|$ sont invariants en rotation.

Homothétie Si l'objet subit une homothétie de rapport α , alors les coefficients de Fourier de la signature sont multipliés par ce même coefficient.

Soit $s(k)$ la signature d'un objet et $s_h(k)$ la signature du même objet qui a subi une homothétie de rapport α , *i.e.* $s_h(k) = \alpha s(k)$, alors

$$\hat{s}_h(u) = \alpha \hat{s}(u) \quad (2.5)$$

Le cercle de rayon $|\hat{s}(1)|$ de centre $s_c = x_c + i y_c$ est une approximation du contour [Zhang 01]. L'invariance aux changements d'échelle est alors obtenue en divisant tous les coefficients de Fourier par $|\hat{s}(1)|$.

Autres propriétés La première est relative à la discrétisation de la courbe paramétrée représentant la signature. La seconde concerne la définition du point de départ dans le calcul de la signature qui s'écrit :

Soit $s_p(k), s(k)$ deux signatures du même objet mais dont le point de départ choisi pour leur calcul est différent, *i.e.* $s_p(k) = s(k - k_0)$, alors

$$\hat{s}_p(u) = \hat{s}(u)e^{2i\pi k_0 u/N} \quad (2.6)$$

où N est le nombre de points total du contours.

Il est clair que l'invariance au changement de point de départ n'est pas respectée car le coefficient multiplicateur dépend de u . Une méthode pour le sélectionner est de prendre comme point de départ le point le plus éloigné du centre. Une autre méthode consiste à sélectionner le point sur l'axe propre de la forme qui est le plus loin du centre [Gonzalez 01].

Un premier ensemble de descripteurs de Fourier FD invariant en translation, rotation et en échelle est défini par

$$FD = \left\{ \frac{|\hat{s}(-N/2 + 1)|}{|\hat{s}(1)|}, \frac{|\hat{s}(-N/2 + 2)|}{|\hat{s}(1)|}, \dots, \frac{|\hat{s}(-1)|}{|\hat{s}(1)|}, \frac{|\hat{s}(2)|}{|\hat{s}(1)|}, \dots, \frac{|\hat{s}(N/2)|}{|\hat{s}(1)|} \right\} \quad (2.7)$$

Comme seul le module est considéré, les coefficients $|\widehat{s}(-k)|$ et $|\widehat{s}(k)|$ sont égaux pour tout $k > 1$. Grâce à ces propriétés de symétrie, un second ensemble de descripteurs FD' qui ne contient que la moitié des descripteurs de Fourier est considéré :

$$FD' = \left\{ \frac{|\widehat{s}(-1)|}{|\widehat{s}(1)|}, \frac{|\widehat{s}(2)|}{|\widehat{s}(1)|}, \dots, \frac{|\widehat{s}(N/2)|}{|\widehat{s}(1)|} \right\} \quad (2.8)$$

Ces descripteurs respectent bien les invariances souhaitées mais ne sont pas complets car seule l'information du module est considérée.

Dans [Gourd 89], deux ensembles de descripteurs de Fourier sont définis à partir de la signature $\rho_c(\theta)$. Celle-ci est l'expression du contour en coordonnées polaires autour du centre de la forme. Les coefficients de Fourier $\widehat{\rho}_c(u)$ sont ensuite calculés sur cette signature. Le premier ensemble est alors le module noté P des coefficients de Fourier calculés sur $\rho_c(\theta)$:

$$P = \{|\widehat{\rho}_c(-N/2 + 1)|, \dots, |\widehat{\rho}_c(N/2)|\} \quad (2.9)$$

Comme seul le module est considéré, un deuxième ensemble R basé sur le décalage de phase est proposé :

$$R = \{(\varphi(-N/2 + 1)/(-N/2 + 1)) - (\varphi(u_0)/u_0), \dots, (\varphi(N/2)/(N/2)) - (\varphi(u_0)/u_0)\} \quad (2.10)$$

où u_0 est l'indice d'un coefficient de Fourier non nul et $e^{i\varphi(u)} = \widehat{\rho}_c(u)/|\widehat{\rho}_c(u)|$.

La réunion de ces deux ensembles de descripteurs forme un ensemble de descripteurs complet et invariant aux déplacements du plan (translations et rotations) [Gourd 89].

On trouve aussi dans la littérature d'autres ensembles de descripteurs comme les descripteurs de Fourier affines [Arbter 90, Mei 08] qui permettent d'obtenir une invariance aux transformations affines. D'autres descripteurs sont définis à partir d'une transformée de Fourier à court-terme [Zhang 05]. Cette transformée correspond à une transformée de Fourier calculée localement sur la signature. Elle est obtenue en appliquant une fonction de fenêtrage (par exemple une fenêtre gaussienne, rectangulaire, hamming etc.) sur le signal et permet de définir des descripteurs de Fourier locaux.

2.2.1.1 Discussion

Nous avons vu dans cette partie comment extraire des descripteurs de Fourier $1D$ à partir du contour d'une forme. Cette approche n'est valable que pour les images binaires, *i.e.*, des images en niveaux de gris ou en couleur qui ont été segmentées.

Ces descripteurs respectent une ou plusieurs invariances (en translation, en rotation, en échelle, aux transformations affines). Ils permettent d'obtenir de bons résultats en reconnaissance de formes [Zhang 05]. De plus, ils sont simples et rapides à calculer.

Les inconvénients de cette approche sont :

- ces descripteurs requièrent l’extraction du contour de la forme qui dépend fortement d’une étape de segmentation (séparation fond/forme). Cette étape est compliquée car la plupart des images naturelles sont difficilement segmentables.
- certaines de ces méthodes reposent sur le calcul du centre de la forme, ce qui ne garantit pas l’unicité de la signature.
- les descripteurs de Fourier $1D$ ne tiennent pas compte de l’intérieur de la forme [Zhang 02a]. Par exemple, les descripteurs de Fourier sont incapables de différencier deux panneaux de circulation triangulaires différents car l’essentiel de l’information se trouve à l’intérieur de la forme.
- la plupart de ces descripteurs ne sont applicables que sur des formes “star-shaped” qui garantissent que le centre de gravité des formes se trouvent bien à l’intérieur de celles-ci. Sinon, un choix de signature ne nécessitant pas d’estimation du centre de la forme doit être fait.

Pour répondre à ces difficultés, certains auteurs ont défini des descripteurs directement sur les images en niveaux de gris pour prendre en compte l’ensemble de l’image. Nous étudions ces approches dans la partie suivante.

2.3 Les descripteurs de Fourier pour les images en niveaux de gris

Dans la littérature, certains auteurs se sont intéressés à la définition de descripteurs directement sur les images en niveaux de gris ($f : \mathbb{R}^2 \mapsto \mathbb{C}$) en utilisant la transformée de Fourier $2D$. Nous avons montré dans le chapitre précédent que la transformée de Fourier $2D$ ne respecte aucune invariance. Toutefois, il est possible de définir à partir des coefficients des transformées de Fourier des descripteurs respectant différentes invariances. Nous présentons dans ce qui suit certains de ces descripteurs qui font partis des plus pertinents.

Parmi ces approches, les descripteurs de Fourier-Mellin [Sheng 86b, Ghorbel 94] sont invariants en translation, rotation et échelle. D’autres descripteurs issus de la M_2 -transformée [Bornard 86, Gauthier 91, Smach 08] sont invariants en translation et en rotation, tout comme les descripteurs génériques de Zhang et Lu [Zhang 02a].

Ces méthodes dédiées aux images en niveaux de gris sont étudiées dans cette partie dans l’objectif d’en proposer une extension pour les images couleur dans une partie ultérieure.

2.3.1 Les descripteurs de Fourier génériques

Les descripteurs de Fourier génériques de Zhang et Lu [Zhang 02a] sont définis sur les images en niveaux de gris. Pour obtenir des descripteurs invariants en translation et en rotation à partir de la transformée de Fourier $2D$, deux transformées de Fourier polaires *ad hoc* ont été définies.

2.3.1.1 Représentations polaires et transformée de Fourier

La première, \widehat{f}^{p1} , est construite en passant les coordonnées cartésiennes (respectivement fréquentielles) de l'image (respectivement de la transformée de Fourier) en coordonnées polaires. Rappelons que la transformée de Fourier $2D$ \widehat{f} d'une image f s'écrit

$$\widehat{f}(u_1, u_2) = \int_{\mathbb{R}^2} f(x_1, x_2) e^{-2i\pi(u_1 x_1 + u_2 x_2)} dx_1 dx_2. \quad (2.11)$$

Le changement de variables suivant est réalisé :

$$\begin{aligned} x_1 &= r \cos(\theta) & x_2 &= r \sin(\theta) \\ u_1 &= \rho \cos(\psi) & u_2 &= \rho \sin(\psi) \end{aligned}$$

où (r, θ) (respectivement (ρ, ψ)) correspondent aux coordonnées spatiales (x_1, x_2) (respectivement fréquentielles (u_1, u_2)) passées dans le domaine polaire. Les formes différentielles de x_1 et x_2 deviennent

$$\begin{aligned} dx_1 &= \cos(\theta) dr - r \sin(\theta) d\theta \\ dx_2 &= \sin(\theta) dr + r \cos(\theta) d\theta \end{aligned}$$

Le déterminant de la Jacobienne du passage en coordonnées polaires est donc r . La transformée de Fourier polaire \widehat{f}^{p1} de l'image dans le domaine polaire $f(r, \theta)$ s'écrit

$$\begin{aligned} \widehat{f}^{p1}(\rho, \Psi) &= \int_0^\infty \int_0^{2\pi} f(r, \theta) e^{-2i\pi(\rho \cos(\psi)r \cos(\theta) + \rho \sin(\psi)r \sin(\theta))} dr d\theta \\ &= \int_0^\infty \int_0^{2\pi} f(r, \theta) r e^{-2i\pi r \rho \cos(\theta - \Psi)} dr d\theta. \end{aligned} \quad (2.12)$$

Comme Ψ est présent dans la fonction cosinus $\cos(\theta - \Psi)$, cette variable perd son sens physique. En effet, Ψ ne correspond pas à une fréquence angulaire [Zhang 02b].

Ce constat a motivé la définition d'une seconde transformée polaire. La transformée \widehat{f}^{p2} est tout simplement la transformée de Fourier $2D$ de l'image dans le domaine polaire :

$$\widehat{f}^{p2}(u_r, u_\theta) = \int_0^\infty \int_0^{2\pi} f(r, \theta) e^{-2i\pi(r u_r + \theta u_\theta)} dr d\theta \quad (2.13)$$

où u_r correspond aux fréquences radiales et u_θ aux fréquences angulaires. Notons que pour cette transformée, les variables r et θ sont traités comme des coordonnées cartésiennes.

La figure 2.2 illustre le calcul de la transformée de Fourier polaire \widehat{f}^{p2} de deux images : une image de la base *COIL* – 100 et la même image après rotation. Nous pouvons remarquer que les modules de \widehat{f}^{p2} sont similaires.

Ces transformées ont été définies de manière *ad hoc* dans l'unique but de définir des descripteurs invariants en translation et en rotation.

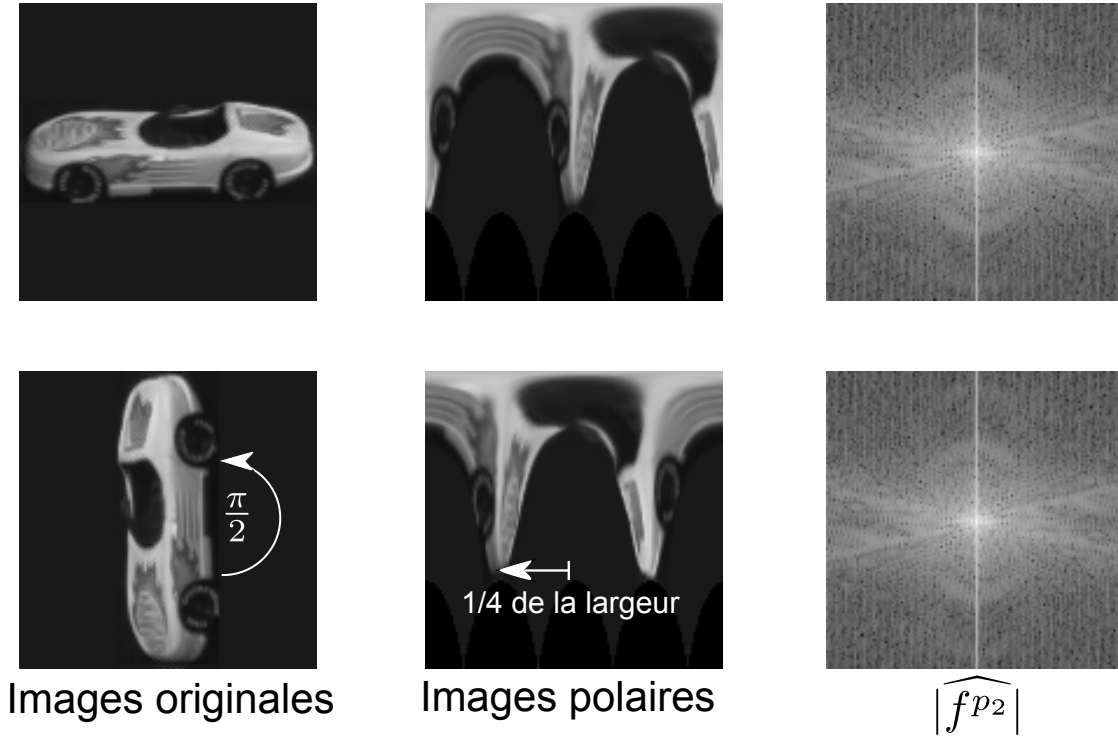


FIGURE 2.2 – Illustration du module de la transformée de Fourier polaire \widehat{f}^{p2} d'une image f de la base COIL-100 [Nene 96]

2.3.1.2 Obtention des invariances

Translation L'invariance en translation est obtenue en redéfinissant le centre de l'image (à l'aide d'une méthode *ad hoc*) pour définir l'origine de l'image polaire. On peut aussi faire l'hypothèse que l'objet d'intérêt est au centre de l'image. Dans ce cas, l'origine de l'image polaire est le centre de l'image originale.

Rotation La rotation de l'image originale produit une translation de l'image polaire, ce qui induit un déphasage dans la transformée de Fourier polaire.

Démonstration. Soit l'image g correspondant à l'image f ayant subi une rotation d'angle β , c'est-à-dire $g(r, \theta) = f(r, \theta - \beta)$ alors

$$\begin{aligned}
 \widehat{g}^{p2}(u_r, u_\theta) &= \int_0^\infty \int_0^{2\pi} f(r, \theta - \beta) e^{-2i\pi(ru_r + \theta u_\theta)} dr d\theta \\
 &= \int_0^\infty \int_0^{2\pi} f(r, \varphi) e^{-2i\pi(ru_r + (\varphi + \beta)u_\theta)} dr d\varphi \\
 &= \int_0^\infty \int_0^{2\pi} f(r, \varphi) e^{-2i\pi(ru_r + \varphi u_\theta)} dr d\varphi e^{-2i\pi(\beta u_\theta)} \\
 &= \widehat{f}^{p2}(u_r, u_\theta) e^{-2i\pi(\beta u_\theta)}
 \end{aligned} \tag{2.14}$$

avec le changement de variable $\varphi = \theta - \beta$. □

L'invariance en rotation est obtenue en prenant le module des coefficients de la transformée de Fourier polaire.

Homothétie Une homothétie de l'image dans le domaine spatial se traduit par une homothétie de l'axe radial dans le domaine polaire. Une homothétie de rapport α dans le domaine polaire se traduit par une homothétie de rapport $1/\alpha$ dans le domaine fréquentiel.

Démonstration. Soit l'image g correspondant à l'image f ayant subi une homothétie de rapport α , c'est-à-dire $g(x_1, x_2) = f(\alpha x_1, \alpha x_2)$. Si on convertit ces images en images polaires, on obtient $g(r, \theta) = f(\alpha r, \theta)$ et

$$\begin{aligned}\widehat{g^{p2}}(u_r, u_\theta) &= \int_0^\infty \int_0^{2\pi} f(\alpha r, \theta) e^{-2i\pi(r u_r + \theta u_\theta)} dr d\theta \\ &= \int_0^\infty \int_0^{2\pi} f(\xi, \theta) e^{-2i\pi(\frac{\xi}{\alpha} u_r + \theta u_\theta)} \frac{1}{\alpha} d\xi d\theta \\ &= \frac{1}{\alpha} \widehat{f^{p2}}\left(\frac{u_r}{\alpha}, u_\theta\right)\end{aligned}\quad (2.15)$$

avec le changement de variable $\xi = \alpha r$. □

Dans l'objectif de limiter les effets des changements d'échelle, tous les coefficients sont alors divisés par $|\widehat{f^{p2}}(0, 0)|$, ce qui annule le coefficient $\frac{1}{\alpha}$ induit par le changement d'échelle. Zhang et Lu proposent dans [Zhang 02a] de normaliser ce coefficient par le volume de l'objet d'intérêt ou par le cercle dans lequel l'image polaire réside. Cependant, l'invariance aux changements d'échelle n'est pas obtenue car la variable fréquentielle u_r est multipliée par $\frac{1}{\alpha}$.

L'ensemble des descripteurs génériques DFG de Zhang et Lu ainsi obtenus s'écrit

$$DFG = \left\{ \frac{|\widehat{f^{p2}}(0, 0)|}{aire}, \frac{|\widehat{f^{p2}}(0, 1)|}{|\widehat{f^{p2}}(0, 0)|}, \dots, \frac{|\widehat{f^{p2}}(1, 1)|}{|\widehat{f^{p2}}(0, 0)|}, \dots, \frac{|\widehat{f^{p2}}(m, n)|}{|\widehat{f^{p2}}(0, 0)|} \right\} \quad (2.16)$$

où *aire* est la surface du cercle englobant l'objet d'intérêt et m et n sont respectivement le nombre maximum de fréquences radiales et le nombre de fréquences angulaires sélectionnées.

Les auteurs préconisent de s'intéresser plus particulièrement aux coefficients du centre de la transformée de Fourier polaire. En effet, on a une concentration de l'énergie dans les premiers coefficients. De plus, si les valeurs des pixels de l'image sont réels, seule la moitié des coefficients est utile pour décrire la forme en utilisant la symétrie hermitienne de la transformée de Fourier $2D$ (si $f \in \mathbb{R}$, alors $\widehat{f}(u, v) = \widehat{f}(-u, -v)$).

2.3.1.3 Discussion

Les descripteurs DFG sont définis pour les images en niveaux de gris. Si l'image à traiter ne contient qu'un seul objet, cela permet de le considérer dans son ensemble. Dans le cadre de la reconnaissance de formes, les objets dont le contour est identique, mais dont "l'intérieur" est différent, sont ainsi discriminés.

Ces descripteurs ont la propriété d'être invariants en translation et en rotation. Ces descripteurs sont rapides à calculer puisqu'ils ne mettent en oeuvre que la transformation de l'image dans le domaine polaire et une transformée de Fourier $2D$ rapide.

Toutefois, on peut noter que le passage de l'image dans le domaine polaire implique une interpolation des valeurs de l'image polaire. L'invariance aux changements d'échelle n'est pas obtenue, ce qui rend difficile la reconnaissance du même objet subissant une forte dilatation ou contraction. De plus, la définition du centre de gravité de l'objet d'intérêt reste problématique. La définition de ces descripteurs n'a pas de justification mathématique rigoureuse car cette méthode est *ad hoc*. Enfin, ces descripteurs de Fourier ne sont pas complets car seul le module est considéré.

Comme nous l'avons montré, la transformée \widehat{f}^{p2} ne permet pas de définir des descripteurs invariants en échelle. Nous avons étudié dans le chapitre précédent une transformée de Fourier permettant d'obtenir des invariants en translation, rotation et en échelle : la transformée de Fourier-Mellin.

2.3.2 Les descripteurs de Fourier-Mellin

Les descripteurs de Fourier-Mellin sont issus des transformées de Fourier-Mellin classique et analytique décrites dans la partie 1.1.2.1. Ces transformées de Fourier ont été construites pour obtenir facilement des descripteurs invariants en translation, rotation et échelle.

2.3.2.1 Les descripteurs de Fourier-Mellin classiques

Les descripteurs de Fourier-Mellin sont issus de la transformée de Fourier-Mellin définie dans la partie 1.1.2 :

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \hat{f}(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{-iv-1} e^{-ik\theta} d\theta dr \quad (2.17)$$

Translation Comme précédemment, l'invariance en translation est obtenue par un pré-traitement consistant à définir l'origine de l'image polaire.

Rotation L'invariance en rotation est obtenue en prenant le module des coefficients $\hat{f}(k, v)$ car une rotation d'angle β de l'image initiale induit un déphasage de la transformée de Fourier-Mellin de $e^{ik\beta}$ (cf. partie 1.1.2.2).

Homothétie Un changement d'échelle de rapport α de l'image, soit $f(\alpha r, \theta)$, induit la multiplication des coefficients de la transformée de Fourier-Mellin par α^{iv} .

Dans [Sheng 86a], un ensemble de descripteurs de Fourier-Mellin normalisé invariant en translation, en rotation, en échelle s'écrit

$$\Phi = \left\{ \frac{|\hat{f}(0, 1)|^2}{\sum_{k=0}^K |\hat{f}(k, 1)|^2}, \dots, \frac{|\hat{f}(1, 1)|^2}{\sum_{k=0}^K |\hat{f}(k, 1)|^2}, \frac{|\hat{f}(1, 2)|^2}{\sum_{k=0}^K |\hat{f}(k, 2)|^2}, \dots, \frac{|\hat{f}(K, V)|^2}{\sum_{k=0}^K |\hat{f}(k, V)|^2} \right\} \quad (2.18)$$

où K est la fréquence angulaire maximale et V la fréquence radiale maximale choisis par l'utilisateur.

Cet ensemble de descripteurs est de grande taille. De plus, dans [Sheng 86a], il est prouvé que cet ensemble contient des redondances d'informations. En effet, l'ensemble des fonctions noyaux $\{r^0, r^1, r^2, \dots, r^v\}$ de l'équation 2.17 n'est pas orthogonal. Pour éviter ces redondances, une sélection des descripteurs est réalisée.

À priori, la dimension de l'espace des caractéristiques n'est pas connue. La normalisation de l'équation 2.18 ne peut donc pas être réalisée et l'ensemble de descripteurs Φ n'est plus adapté. Un autre ensemble de descripteurs Φ' est alors défini par

$$\Phi' = \left\{ \frac{|\hat{f}(1,1)|^2}{|\hat{f}(0,1)|^2}, \dots, \frac{|\hat{f}(2,1)|^2}{|\hat{f}(0,1)|^2}, \frac{|\hat{f}(2,2)|^2}{|\hat{f}(0,2)|^2}, \dots, \frac{|\hat{f}(K,V)|^2}{|\hat{f}(0,V)|^2} \right\} \quad (2.19)$$

où K et V sont choisis par l'utilisateur. Dans [Sheng 86a], les auteurs préconisent de s'intéresser plus particulièrement aux descripteurs calculés avec $v = 2$ car plus v augmente, plus les descripteurs de Fourier-Mellin amènent de la redondance d'information. Il en résulte une perte d'efficacité pour la reconnaissance de formes [Sheng 86a].

Ces deux ensembles de descripteurs sont invariants en translation, rotation et en échelle. Par contre, la complétude de ces descripteurs n'est pas obtenue car le module seul est utilisé [Derrode 01].

2.3.2.2 Les descripteurs de Fourier-Mellin Analytiques

Commençons par rappeler la définition de la transformée de Fourier-Mellin analytique qui est donnée dans la partie 1.1.2.1. Soit f une fonction représentant une image en niveaux de gris, la transformée de Fourier-Mellin analytique s'écrit

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \hat{f}_\sigma(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r} \quad (2.20)$$

avec $\sigma > 0$.

Dans [Derrode 01], Derrode et Ghorbel définissent une autre famille de descripteurs invariante et complète, les descripteurs de Fourier-Mellin analytique, qui s'écrivent

$$I(k, v) = \hat{f}_\sigma(0, 0) \frac{-\sigma+iv}{\sigma} e^{-ik \arg(\hat{f}_\sigma(0,1))} \hat{f}_\sigma(k, v) \quad (2.21)$$

$\forall k, v \in \mathbb{Z} \times \mathbb{R}$.

L'ensemble des descripteurs de Fourier-Mellin analytiques Φ_a est alors

$$\Phi_a = \{I(-K, -V), \dots, I(1, 1), I(1, 2), \dots, I(K, V)\} \quad (2.22)$$

où K et V sont choisis par l'utilisateur. Dans [Derrode 01], il est spécifié que le choix de ces deux variables peut être guidé par le contenu fréquentiel de l'image et de l'application, mais qu'aucune règle générale ne peut être définie. Cependant, l'algorithme rapide défini dans [Derrode 01] correspond au calcul de la transformée de Fourier usuelle sur l'image log-polaire. Il est évident que dans ce cas les variables K et V sont bornées par la taille $M \times N$ de l'image dans le domaine log-polaire, c'est à dire $K \leq 2M + 1$ et $V \leq 2N + 1$.

2.3.2.3 Discussion

Les descripteurs de Fourier-Mellin, contrairement aux descripteurs DFG , respectent bien les invariances en translation, rotation et en échelle. Les descripteurs de Fourier-Mellin analytiques Φ_a sont des familles d'invariants complètes et convergentes. Leur définition mathématique est rigoureuse et une transformée de Fourier rapide peut être utilisée pour calculer la transformée de Fourier-Mellin.

Néanmoins, comme nous l'avons évoqué dans la partie 1.1.2, la transformée de Fourier-Mellin ne peut être calculée que par interpolation, ce qui la rend imparfaitement inversible en pratique. Le calcul de la transformée de Fourier-Mellin nécessite aussi d'estimer le centre de l'image. La qualité des invariants extraits dépend de cette estimation qui reste difficile en pratique car elle est sensible au bruit de l'image.

2.3.3 Les descripteurs de Fourier généralisés

Le premier ensemble de descripteurs de Fourier généralisés a été défini par Gauthier *et al.* [Gauthier 91] à partir de la M_2 -transformée. Plus récemment, Smach *et al.* ont proposé dans [Smach 08] un second ensemble de descripteurs issu de cette transformée. Ceux-ci sont invariants par l'action du groupe M_2 . Ces descripteurs sont idéaux d'un point de vue théorique car ils sont invariants en translation et rotation par construction.

2.3.3.1 Définitions

Les descripteurs de Fourier généralisés sont définis à partir de la M_2 -transformée (voir partie 1.1.4). Nous avons vu dans la partie 2.2.1 que les descripteurs P de type densité spectrale et R de type décalage de phase forment un ensemble complet de descripteurs de Fourier. Ces descripteurs, définis pour le contour des objets, peuvent être étendus aux images en niveaux de gris. Deux ensembles de descripteurs sont alors définis dans [Smach 08] :

Définition 2.3.3.1. *Les invariants $I_1^\xi(f)$ de type densité spectrale et les invariants de type décalage de phase $I_2^{\xi_1, \xi_2}(f)$ sont définis pour des fonctions $f \in L^2(\mathbb{R}^2)$ par*

$$I_1^\xi(f) = \int_0^{2\pi} \left| \widehat{f}(R_\theta(\xi)) \right|^2 d\theta \quad (2.23)$$

$$I_2^{\xi_1, \xi_2}(f) = \int_0^{2\pi} \widehat{f}(R_\theta(\xi_1 + \xi_2)) \widetilde{\widehat{f}}(R_\theta(\xi_1)) \widetilde{\widehat{f}}(R_\theta(\xi_2)) d\theta \quad (2.24)$$

où f est l'image, \widehat{f} est la transformée de Fourier usuelle, ξ , ξ_1 et ξ_2 sont des vecteurs définis dans le domaine fréquentiel et R_θ est un opérateur de rotation d'angle θ . Il est prouvé dans [Smach 08] la complétude faible de ces invariants.

En reprenant des notations similaires à celles utilisées pour les descripteurs génériques DFG , un vecteur de descripteurs utilisé dans [Smach 07] à valeurs réelles se note :

$$GFD1 = \left\{ I_1^{(0,0)^t}(f), \frac{I_1^{(0,1)^t}(f)}{I_1^{(0,0)^t}(f)}, \dots, \frac{I_1^{(0,m)^t}(f)}{I_1^{(0,0)^t}(f)} \right\} \quad (2.25)$$

où m est le nombre de descripteurs calculés.

Ces descripteurs sont invariants en translation, rotation et en réflexion. L'ensemble des vecteurs $\xi = (0, k)^t$ avec $k \in [1, N/2]$ (N étant le diamètre du cercle inscrit dans l'image de la transformée de Fourier) peuvent être pris pour calculer les invariants $I_1^\xi(f)$.

De la même manière, le vecteur de descripteurs $GFD2$ est construit à partir de la seconde famille d'invariants $I_2^{\xi_1, \xi_2}$:

$$GFD2 = \left\{ I_2^{(0,0)^t, (0,0)^t}(f), \frac{I_2^{(0,1)^t, (-o, -p)^t}(f)}{I_2^{(0,0)^t, (0,0)^t}(f)}, \dots, \frac{I_2^{(0,1)^t, (o,p)^t}(f)}{I_2^{(0,0)^t, (0,0)^t}(f)}, \dots, \frac{I_2^{(0,m)^t, (-o, -p)^t}(f)}{I_2^{(0,0)^t, (0,0)^t}(f)}, \dots, \frac{I_2^{(0,m)^t, (o,p)^t}(f)}{I_2^{(0,0)^t, (0,0)^t}(f)} \right\} \quad (2.26)$$

où m, o et p sont choisis par l'utilisateur sachant que la norme euclidienne du vecteur $(0, m)^t + (o, p)^t$ doit être inférieure à $N/2$.

Les $GFD2$ sont invariants en translation et en rotation. Les vecteurs ξ_1 et ξ_2 sont choisis par l'utilisateur. Le choix de cette paire de vecteurs est moins contraint que le choix de ξ , et il est donc plus délicat. En prenant les vecteurs ξ_1 de la forme $(0, k)^t$ ($k \in [1; N - 1]$), cet ensemble de descripteurs est complètement déterminé.

Cela laisse encore beaucoup de choix pour ξ_1 et ξ_2 . Cependant, nous pouvons réduire ces choix en privilégiant les basses fréquences si l'on veut se focaliser sur la forme de l'objet ou bien sur les hautes fréquences si l'on veut considérer la texture de la forme.

Nous pouvons aussi remarquer que pour un choix particulier de ξ_1 et ξ_2 pour $I_2^{\xi_1, \xi_2}$, on obtient les invariants I_1^ξ . En effet, si $\xi_1 = (0, r)^t$, $r \in [1; \frac{N}{2} - 1]$ et $\xi_2 = (0, 0)^t$, alors $I_2^{\xi_1, \xi_2}(f) = \hat{f}(0, 0) I_1^{\xi_1}(f)$.

Le calcul de ces deux ensembles de descripteurs nécessite d'intégrer les valeurs de la transformée de Fourier sur des cercles discrets. La méthode de calcul est donnée dans la partie suivante.

2.3.3.2 Méthode de calcul

Pour le premier ensemble d'invariants $I_1^\xi(f)$, les valeurs sur la grille discrète de la transformée de Fourier discrète $2D$ doivent être intégrées dans un cercle dont le rayon r est défini pour chaque descripteur [Smach 08].

Sur chaque cellule de la grille de la transformée de Fourier discrète, la valeur de la transformée de Fourier continue est considérée comme constante. L'intégration sur un cercle continu est alors approximée par une intégration des valeurs discrètes pondérées par les longueurs d'arcs traversant chaque cellule.

Le calcul de la longueur d'arc est illustré dans la figure 2.3. La longueur d'arc est égale au rayon du cercle d'intégration multiplié par l'angle ψ entre le point d'entrée du cercle dans le pixel et son point de sortie. Comme on peut le vérifier sur la figure 2.3, la longueur d'arc L est égale à

$$\begin{aligned} \psi &= \arcsin \left(\sqrt{1 - \frac{b^2}{r^2}} \right) - \arcsin \left(\frac{a}{r} \right) \\ L &= r\psi \end{aligned} \quad (2.27)$$

où r est le rayon du cercle d'intégration, a est l'abscisse du point de départ de l'arc, b est l'ordonnée du point de sortie de l'arc.

Concernant le second type d'invariants, la quantité $\widehat{f}(R_\theta(\xi_1 + \xi_2))\widehat{f}(R_\theta(\xi_1))\widehat{f}(R_\theta(\xi_2))$ est pondérée par le plus petit des trois arcs des trois pixels traversés par $R_\theta(\xi_1 + \xi_2)$, $R_\theta(\xi_1)$ et $R_\theta(\xi_2)$.

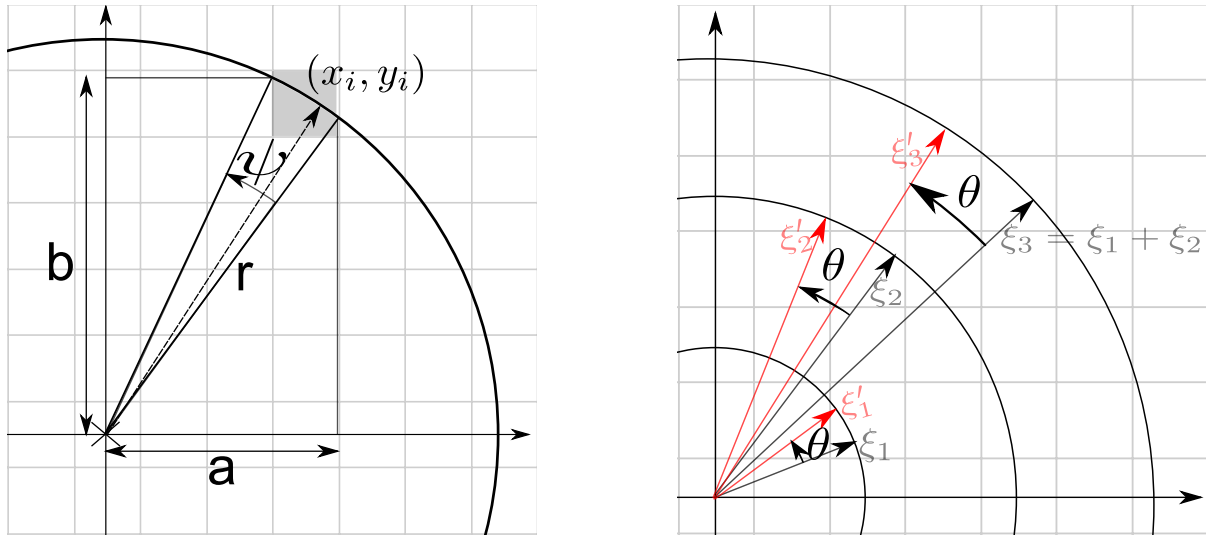


FIGURE 2.3 – Calcul de GFD1 (à gauche) et calcul de GFD2 (à droite)

2.3.3.3 Discussion

Ces descripteurs sont invariants en translation et en rotation et ne nécessitent pas d'estimer le centre de l'image comme pour les descripteurs de Fourier-Mellin. La complétude faible de ces descripteurs est prouvée dans [Smach 08]. Elle signifie que les descripteurs sont complets pour un large sous-ensemble de l'ensemble des images, ce qui est suffisant en général. Des expérimentations montrent que les taux de reconnaissance obtenus avec ces descripteurs sont très bons (voir chapitre 3). De plus, ces descripteurs peuvent être facilement calculés avec une transformée de Fourier rapide. Néanmoins, ces descripteurs ne sont pas invariants en échelle et le choix des vecteurs ξ_1 et ξ_2 est difficile en pratique.

2.3.4 Comparaison des descripteurs de Fourier

Il existe de nombreux descripteurs de Fourier dans la littérature. Cependant, l'utilisation d'un ensemble particulier de descripteurs dépend de l'application. Nous pouvons distinguer deux types de descripteurs de Fourier : ceux définis sur le contours des objets et ceux définis sur l'image toute entière.

Les descripteurs de Fourier 1D usuels ont pour inconvénients de nécessiter l'extraction du contour de l'objet et de ne pas prendre en compte l'intérieur de l'objet à reconnaître.

Les descripteurs de Fourier définis pour les images en niveaux de gris répondent à ces problématiques. Cependant, pour que ces méthodes soient efficaces, une hypothèse est faite sur le

contenu de l'image à traiter. Le fond ne doit pas être trop complexe car il fausserait la reconnaissance de formes. La transformée de Fourier $2D$ contient la totalité des fréquences de l'image et mélange donc les fréquences du fond et de l'objet d'intérêt. Cela nous permet de rappeler que la transformée de Fourier est une méthode globale. Ces méthodes sont d'autant plus efficaces que l'image contient un seul objet et que le fond est relativement uniforme.

La problématique principale de cette approche reste de définir des descripteurs invariants en translation, rotation et aux changements d'échelle. Les descripteurs de Fourier-Mellin respectent bien ces invariances mais dépendent du calcul du centre de gravité de l'objet ou de l'image. Cette estimation reste difficile et peut dégrader les résultats. Les descripteurs $GFD1$ et $GFD2$ peuvent être utilisés. Ces descripteurs ne nécessitent pas d'étape d'interpolation ou d'estimation de centre et ils sont définis de manière rigoureuse. C'est pourquoi, ces descripteurs ont retenu notre attention pour la suite de ce manuscrit.

Nous avons étudié dans cette partie les principaux descripteurs de Fourier dédiés aux contours et aux images en niveaux de gris. L'extension aux images à trois canaux (RVB par exemple) ou plus généralement à n -canaux de ces descripteurs est décrit dans la partie suivante.

2.4 Les descripteurs de Fourier pour les images couleur

Dans la littérature, la majorité des méthodes fréquentielles de reconnaissance d'images sont dédiées aux images en niveaux de gris. L'extension de ces méthodes aux images couleur n'est pas simple car les images ne sont plus définies comme des fonctions de \mathbb{R}^2 dans \mathbb{R} mais de \mathbb{R}^2 dans \mathbb{R}^3 . La valeur de chaque pixel n'est plus un scalaire mais un vecteur.

De nombreuses approches fréquentielles sont étendues à la couleur en utilisant des méthodes *ad hoc*. Certains auteurs se sont donnés l'objectif de définir des descripteurs de Fourier couleur plus rigoureusement. Dans cette partie, nous étudions les principales approches permettant de définir des descripteurs de Fourier couleur et nous définissons trois nouveaux ensembles de descripteurs.

2.4.1 Méthode $1D$ utilisant des signatures dépendantes de la couleur

L'approche retenue par Kunttu [Kunttu 05] est basée sur une fonction de surface. L'objet est décomposé en parcelles triangulaires (voir figure 2.4). La surface de chacune de ces parties est définie par

$$a(t) = \frac{|(x(t) - x_c)(y(t+1) - y_c) - (x(t+1) - x_c)(y(t) - y_c)|}{2} \quad (2.28)$$

avec $t \in [1, N-1]$, N le nombre de points du contour et $a(t)$ représente l'aire des triangles formés par 2 points du contours et le centre de l'objet (x_c, y_c) .

La couleur est associée à cette signature en considérant une statistique réelle $C(t)$ (par exemple la moyenne d'un des canaux RVB) de cette information pour chaque région t . La signature tenant compte de la couleur s'écrit

$$c_a(t) = a(t) + iC(t) \quad (2.29)$$

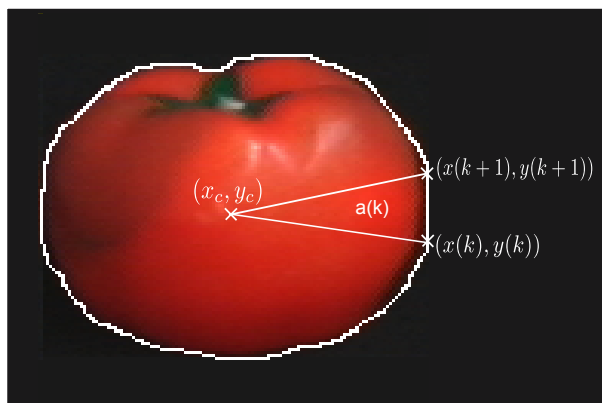


FIGURE 2.4 – Extraction de la signature $a(k)$ d'une image de la base COIL-100

La transformée de Fourier 1D discrète est alors calculée sur cette signature pour obtenir les coefficients $\hat{c}_a(u)$. L'invariance en translation est obtenue en définissant le centre de la forme (x_c, y_c) . Comme pour les descripteurs de Fourier 1D usuels, le module de ces coefficients est utilisé pour obtenir une invariance en rotation. Une normalisation par le coefficient $|\hat{c}_a(0)|$ (qui correspond à la valeur moyenne du signal) est réalisée pour obtenir une invariance aux changements d'échelle.

L'ensemble de descripteurs S_c est alors

$$S_c = \left\{ \frac{|\hat{c}_a(-N/2)|}{|\hat{c}_a(0)|}, \dots, \frac{|\hat{c}_a(N/2)|}{|\hat{c}_a(0)|} \right\} \quad (2.30)$$

Cette méthode apporte une information supplémentaire pertinente en reconnaissance de formes. L'inconvénient est que cette approche nécessite la définition du centre de la forme. De plus, l'information colorimétrique que peut contenir $C(t)$ est très limitée car elle n'est représentée que par un réel. La perte d'information couleur est donc importante.

Nous pouvons suggérer une extension de cette méthode en utilisant les quaternions, *i.e.*,

$$c'_a(t) = a(t) + iC_1(t) + jC_2(t) + kC_3(t) \quad (2.31)$$

2.4.2 Méthode marginale 2D

Cette méthode correspond à un traitement marginal des canaux couleur (par exemple *RVB*). Cela signifie que les trois canaux sont considérés comme indépendants et donc un ensemble de descripteurs est extrait pour chacun de ces canaux. Il en résulte trois vecteurs de descripteurs qui sont ensuite concaténés dans un seul et unique vecteur de description (voir figure 2.5). Cette méthode a pour inconvénient de traiter indépendamment des canaux couleur qui sont corrélés et donc de produire des descripteurs corrélés. La redondance d'information qui en découle peut faire baisser le taux de reconnaissance.

On retrouve cette approche dans [Torres 99] où les "eigenfaces", étendus à certaines parties du visages ("eigennose", "eigeneyes", etc.) sont calculés sur les trois canaux des images couleur de visages. Dans [Smach 08], les descripteurs de Fourier généralisés *GFD* sont extraits des trois

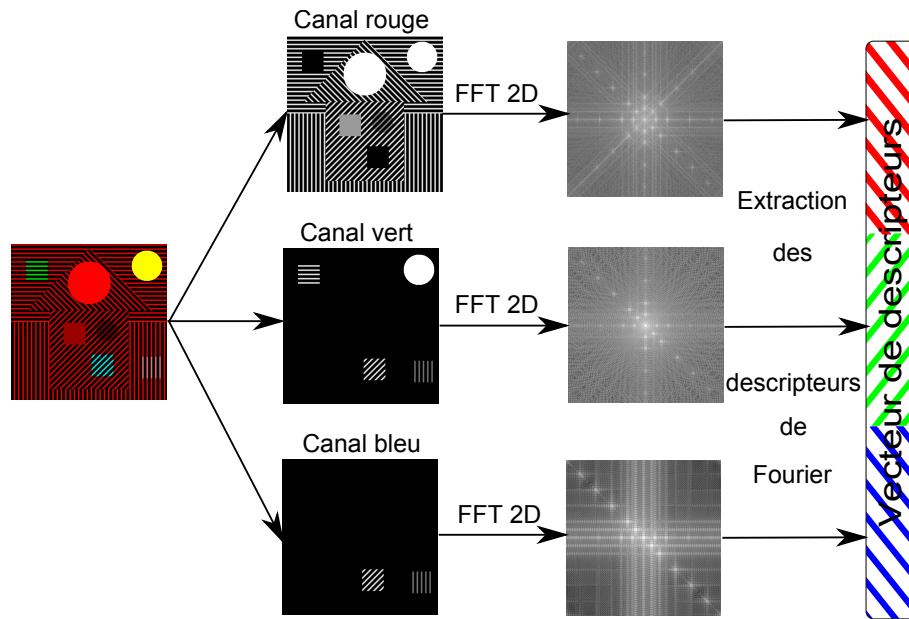


FIGURE 2.5 – Extraction de descripteurs de Fourier selon la méthode marginale

canaux *RVB* marginalement. Trois ensembles de descripteurs sont alors obtenus, nécessitant le calcul de trois transformées de Fourier 2D et triplant la taille du vecteur de descripteurs. Notons que cette méthode peut être appliquée à n'importe quelle approche destinée aux images en niveaux de gris comme [Zhang 02a, Derrode 01].

2.4.3 Méthode par projection de l'image sur un seul plan couleur

Une autre méthode pour gérer les images couleur revient à projeter l'image couleur sur un plan couleur. Cette approche est proposée par Kakuta *et al.* [Kakuta 97] qui, dans le cadre de traitements optiques, projettent l'image couleur sur un vecteur couleur ou sur un plan couleur. Une image couleur est alors définie comme

$$f(x, y) = r(x, y)\mathbf{e}_1 + v(x, y)\mathbf{e}_2 + b(x, y)\mathbf{e}_3 \quad (2.32)$$

où r, v, b sont les canaux rouge, vert et bleu de l'image et $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ les vecteurs de base de l'espace colorimétrique *RVB*.

L'image projetée sur un vecteur définit un signal de \mathbb{R}^2 dans \mathbb{R} :

$$f_{1D}(x, y) = \langle f(x, y), \mathbf{u} \rangle \quad (2.33)$$

où (x, y) sont les coordonnées spatiales, $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3$ un vecteur représentant une couleur et $\langle \cdot, \cdot \rangle$ est le produit scalaire usuel.

L'autre solution proposée par les auteurs consiste à projeter l'image sur un plan couleur formé par deux vecteurs orthonormaux \mathbf{u} et \mathbf{v} . Les deux projections sur ces vecteurs sont alors identifiées à un nombre complexe

$$f_{2D}(x, y) = \langle f(x, y), \mathbf{u} \rangle + i\langle f(x, y), \mathbf{v} \rangle \quad (2.34)$$

avec $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3$ et $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$ deux vecteurs orthonormaux. Le choix des vecteurs \mathbf{u} et \mathbf{v} sont laissés à l'utilisateur. D'un point de vue pratique, ce choix dépend de l'image couleur à traiter car sa projection sur le plan défini par ces deux vecteurs ne doit pas être nulle si l'on veut un résultat intéressant.

Les descripteurs de Fourier [Smach 08, Derrode 01, Zhang 02a] peuvent alors être extraits de l'image réelle f_{1D} ou de l'image complexe f_{2D} .

Cette approche met alors en avant le choix d'une direction d'analyse (le choix d'un plan ou d'un vecteur couleur). Le problème avec cette méthode est qu'elle implique une perte d'information colorimétrique car l'analyse n'a lieu que sur un seul vecteur ou un seul plan couleur.

2.4.4 Les moments de Fourier-Mellin quaternioniques

Récemment, des descripteurs quaternioniques couleur ont été proposés à partir de deux transformées de Fourier-Mellin quaternionique définies dans [Guo 11]. Soit une image couleur $f(r, \theta)$ exprimée en coordonnées polaires, deux transformées de Fourier-Mellin quaternioniques ont été définies

$$\begin{aligned}\hat{f}_{\sigma\mu_1}^R(k, v) &= \int_0^\infty \int_0^{2\pi} r^{v-1} e^{-\mu_1 k \theta} f(r, \theta) d\theta dr \\ \hat{f}_{\sigma\mu_1}^L(k, v) &= \int_0^\infty \int_0^{2\pi} r^{v-1} f(r, \theta) e^{-\mu_1 k \theta} d\theta dr\end{aligned}\quad (2.35)$$

où $\hat{f}_{\sigma\mu_1}^L$ (respectivement $\hat{f}_{\sigma\mu_1}^R$) correspond à la transformée de Fourier-Mellin quaternionique à gauche (respectivement droite), μ_1 un quaternion pur et unitaire et k, v sont des entiers positifs. Il est montré dans [Guo 11] que ces deux transformées sont similaires pour décrire les rotations et les changements d'échelle des images couleurs. L'algorithme de calcul de ces moments est

- Estimer le centre de l'image et déplacer l'origine des coordonnées à ce centre tel que

$$x_c = \frac{\sum_{x=1}^M \sum_{y=1}^N x |f(x, y)|}{\sum_{x=1}^M \sum_{y=1}^N |f(x, y)|}, \quad y_c = \frac{\sum_{x=1}^M \sum_{y=1}^N y |f(x, y)|}{\sum_{x=1}^M \sum_{y=1}^N |f(x, y)|}\quad (2.36)$$

où M, N sont la largeur et la hauteur de l'image et x_c, y_c sont les coordonnées du centre.

- Convertir l'image originale en image polaire
- Calculer une transformée de Fourier quaternionique 1D à gauche (ou à droite) sur chacune des lignes de l'image polaire 2.35
- Calculer les moments de Fourier-Mellin quaternioniques de l'équation 2.35 à l'aide d'une transformée de Mellin

L'invariance en translation est obtenue en définissant l'origine de l'image polaire comme le centre estimé de l'image originale. La rotation conduit à un déphasage de ces moments et

le module des coefficients est alors utilisé pour obtenir cette invariance. Enfin, l'invariance en échelle est obtenue en divisant chaque coefficient $\hat{f}_{\sigma_{\mu_1}}(k, v)$ par le coefficient $\hat{f}_{\sigma_{\mu_1}}(0, v)$ car un changement d'échelle de rapport α provoque la multiplication des coefficients par α^{-v} . L'ensemble des descripteurs invariants sont alors définis par

$$\phi_{\mu_1} = \left\{ \left| \frac{\hat{f}_{\sigma_{\mu_1}}(1, 1)}{\hat{f}_{\sigma_{\mu_1}}(0, 1)} \right|, \left| \frac{\hat{f}_{\sigma_{\mu_1}}(2, 1)}{\hat{f}_{\sigma_{\mu_1}}(0, 1)} \right|, \dots, \left| \frac{\hat{f}_{\sigma_{\mu_1}}(1, 2)}{\hat{f}_{\sigma_{\mu_1}}(0, 2)} \right|, \dots, \left| \frac{\hat{f}_{\sigma_{\mu_1}}(K, V)}{\hat{f}_{\sigma_{\mu_1}}(0, V)} \right| \right\} \quad (2.37)$$

Comme pour les descripteurs de Fourier-Mellin niveaux de gris, le nombre de descripteurs à extraire doit être choisi par l'utilisateur. Concernant ce choix, les auteurs ne donnent pas d'indications précises, de même que pour la direction d'analyse μ_1 .

2.4.5 De nouveaux descripteurs de Fourier couleur : GCFD1 et GCFD2

Dans [Smach 08], deux ensembles de descripteurs invariants en translation et en rotation ont été définis. Ces descripteurs dédiés aux images en niveaux de gris sont étendus à la couleur à l'aide de la méthode marginale décrite dans la partie 2.4.2. Les descripteurs ainsi extraits de chaque canal sont concaténés pour former un vecteur de descripteurs. Dans le but d'éviter ce traitement marginal et d'améliorer les résultats de Smach *et al.*, nous définissons de nouveaux descripteurs à partir de la transformée de Fourier Clifford définie par Batard *et al.* [Batard 10]. Cette approche originale fait partie de nos contributions.

Nous avons vu que cette transformée de Fourier Clifford couleur se décompose en une partie parallèle (projection de l'image sur un plan défini par un bivecteur B) et une partie orthogonale (projection de l'image sur le plan orthogonal à celui défini par B). Ces deux parties peuvent être identifiées respectivement à une image réelle et une image complexe. Les descripteurs de Smach *et al.* peuvent alors être extraits de ces deux images, fournissant alors deux ensembles de descripteurs et nécessitant deux transformées de Fourier 2D rapides.

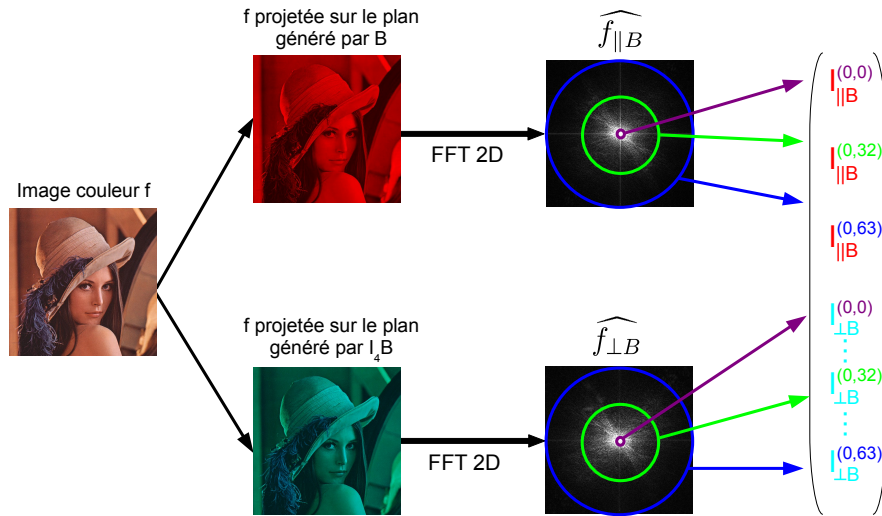
Nous proposons alors deux ensembles de descripteurs de Fourier généralisés couleur ($GCFD1$ et $GCFD2$). Ceux-ci se décomposent en deux sous-ensembles correspondant à la partie parallèle et orthogonale de la transformée de Fourier Clifford couleur ($GCFD1_{\parallel B}, GCFD1_{\perp B}$ et $GCFD2_{\parallel B}, GCFD2_{\perp B}$) qui sont définis comme

$$GCFD1_{\parallel B}(f) = \left\{ I_{\parallel B}^{(0,0)}(f) = \left| \widehat{f}_{\parallel B}(0, 0) \right|^2, \frac{I_{\parallel B}^{(0,1)}(f)}{I_{\parallel B}^{(0,0)}(f)}, \dots, \frac{I_{\parallel B}^{(0,m)}(f)}{I_{\parallel B}^{(0,0)}(f)} \right\} \quad (2.38)$$

où $I_{\parallel B}^{\xi}(f) = \int_0^{2\pi} \left| \widehat{f}_{\parallel B}(R_{\theta}(\xi)) \right|^2 d\theta$ et m est le nombre de descripteurs calculés. De manière similaire, $GCFD1_{\perp B}$ est défini grâce à $\widehat{f}_{\perp B}$. Finalement, le vecteur de descripteurs est de taille $2 \times m$:

$$GCFD1_B(f) = \{GCFD1_{\parallel B}(f), GCFD1_{\perp B}(f)\} \quad (2.39)$$

Le calcul de ces descripteurs est illustré dans la figure 2.6.

FIGURE 2.6 – Extraction des descripteurs $GCFD1$ d'une image couleur

Les descripteurs $GCFD2_B$ sont construits à partir de I^{ξ_1, ξ_2} . La principale différence avec les descripteurs de Smach *et al.* est que les descripteurs extraits de la partie orthogonale ne sont pas réels. En effet, les descripteurs de type “décalage de phase” sont réels pour des fonctions réelles [Smach 08], mais sont complexes pour des fonctions complexes. Or, l'image correspondant à la partie parallèle est réelle et celle correspondant à la partie orthogonale est complexe. Nous avons alors décidé de considérer le module des descripteurs $GCFD2_{\perp B}$ dans le but d'obtenir des descripteurs réels qui sont utilisés dans le chapitre 3.

Soulignons qu'il résulte de notre approche seulement deux ensembles de descripteurs au lieu de trois pour la méthode marginale. La taille du vecteur de descripteurs est réduit par rapport à la méthode marginale. Les descripteurs sont aussi moins redondants et donc plus pertinents qu'avec la méthode marginale. De plus, notre méthode apporte un gain en complexité algorithmique car la transformée de Fourier Clifford ne nécessite que deux transformées de Fourier (au lieu de trois pour la méthode marginale).

2.4.6 Un nouvel ensemble de descripteurs obtenu à partir de la transformée de Fourier Clifford reconstruite : les $GCFD3$

La transformée de Fourier Clifford peut être reconstruite tel un vecteur de $\mathbb{R}_{4,0}^1$ (voir partie 1.2.3.6). Un ensemble de descripteurs invariants en translation et en rotation peut alors être extrait en s'inspirant des descripteurs de type densité spectrale de Smach *et al.* [Smach 08]. En étudiant les propriétés de la transformée de Fourier Clifford couleur, nous définissons un nouvel ensemble de descripteurs de Fourier couleur à partir de la transformée de Fourier Clifford reconstruite.

La transformée de Fourier Clifford d'une fonction de \mathbb{R}^2 dans $\mathbb{R}_{4,0}^1$ est une fonction à valeurs dans $\mathbb{R}_{4,0}^1$ paramétrée par un bivecteur B . Alors, en remplaçant le module d'un complexe

par la magnitude d'un vecteur de $\mathbb{R}_{4,0}$ dans l'équation 2.23, l'ensemble de descripteurs réels $GCFD3_B(f)$ est obtenu

$$GCFD3_B(f) = \left\{ I_B^{(0,0)}(f) = |\widehat{f}_B(0,0)|^2, \frac{I_B^{(0,1)}(f)}{I_B^{(0,0)}(f)}, \dots, \frac{I_B^{(0,m)}(f)}{I_B^{(0,0)}(f)} \right\} \quad (2.40)$$

avec

$$I_B^\xi(f) = \int_0^{2\pi} |\widehat{f}_B(R_\theta(\xi))|^2 d\theta \quad (2.41)$$

où \widehat{f}_B la transformée de Fourier Clifford couleur calculée pour un bivecteur B .

Nous obtenons alors seulement un vecteur de descripteurs couleur et donc trois fois moins de descripteurs que pour la méthode marginale. Le nombre de transformées de Fourier rapides nécessaires au calcul est toujours de deux comme pour les descripteurs $GCFD1$ et $GCFD2$ et ils nécessitent en supplément la reconstruction de la transformée de Fourier Clifford couleur.

Nous pouvons aussi remarquer que les descripteurs $GCFD3_B(f)$ sont la somme des descripteurs $GCFD1_{\parallel B}(f)$ et des $GCFD1_{\perp B}(f)$. En effet, la transformée de Fourier Clifford couleur conserve la propriété d'isométrie de la transformée de Fourier usuelle (égalité de Parseval [Bracewell 86]). Cette propriété énonce que l'énergie totale d'un signal ne dépend pas de la représentation choisie (fréquentielle ou temporelle), *i.e.*,

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(\mathbf{x})|^2 d\mathbf{x} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\widehat{f}_B(\mathbf{u})|^2 d\mathbf{u} \quad (2.42)$$

D'après la définition de la transformée de Fourier Clifford couleur, on peut prouver que $|\widehat{f}_B|^2 = |\widehat{f}_{\parallel B}|^2 + |\widehat{f}_{\perp B}|^2$ et nous permet d'écrire

$$\begin{aligned} I_B^\xi(f) &= \int_0^{2\pi} |\widehat{f}_B(R_\theta(\xi))|^2 d\theta = \int_0^{2\pi} |\widehat{f}_{\parallel B}(R_\theta(\xi))|^2 + |\widehat{f}_{\perp B}(R_\theta(\xi))|^2 d\theta \\ &= \int_0^{2\pi} |\widehat{f}_{\parallel B}(R_\theta(\xi))|^2 d\theta + \int_0^{2\pi} |\widehat{f}_{\perp B}(R_\theta(\xi))|^2 d\theta \\ &= I_{\parallel B}^\xi + I_{\perp B}^\xi \end{aligned} \quad (2.43)$$

De plus, nous verrons dans la partie applicative que ces descripteurs sont invariants au choix du bivecteur.

Les descripteurs $GCFD3$ sont invariants en translation et en rotation. L'invariance en translation est évidente car, d'après l'équation 1.102, une translation dans le domaine spatial induit un déphasage dans le domaine spectral et en prenant la magnitude au carré, l'invariance en translation est obtenue. L'invariance en rotation est obtenue en intégrant les valeurs de la transformée de Fourier sur un cercle. En effet, nous avons vu dans la partie 1.2.3.3 que l'effet d'une rotation de l'image dans le domaine spectral de la transformée de Fourier Clifford s'écrit :

Démonstration. Soit $f, g \in L^2(\mathbb{R}^2, \mathbb{R}_{4,0}^1)$ tel que $g(x_1, x_2) = f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta))$ et θ un angle de rotation alors

$$\begin{aligned} \widehat{g}_B(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(x_1 u_1 + x_2 u_2)(B + I_4 B)} f(x_1 \cos(\theta) + x_2 \sin(\theta), -x_1 \sin(\theta) + x_2 \cos(\theta)) \\ &\quad e^{-\frac{1}{2}(x_1 u_1 + x_2 u_2)(B + I_4 B)} dx_1 dx_2 \end{aligned} \quad (2.44)$$

Par changements de variables,

$$\begin{aligned} x'_1 &= x_1 \cos(\theta) + x_2 \sin(\theta) \Rightarrow x_1 = x'_1 \cos(\theta) - x'_2 \sin(\theta) \\ x'_2 &= -x_1 \sin(\theta) + x_2 \cos(\theta) \Rightarrow x_2 = x'_1 \sin(\theta) + x'_2 \cos(\theta) \end{aligned} \quad (2.45)$$

On obtient l'équation suivante

$$\begin{aligned} \widehat{g}_B(u_1, u_2) &= \int_{\mathbb{R}^2} e^{\frac{1}{2}(u_1 x'_1 \cos(\theta) - u_1 x'_2 \sin(\theta) + u_2 x'_1 \sin(\theta) + u_2 x'_2 \cos(\theta))(B + I_4 B)} f(x'_1, x'_2) \\ &\quad e^{-\frac{1}{2}(u_1 x'_1 \cos(\theta) - u_1 x'_2 \sin(\theta) + u_2 x'_1 \sin(\theta) + u_2 x'_2 \cos(\theta))(B + I_4 B)} dx'_1 dx'_2 \\ &= \int_{\mathbb{R}^2} e^{-\frac{1}{2}([(u_1 \cos(\theta) + u_2 \sin(\theta))x'_1 + (-u_1 \sin(\theta) + u_2 \cos(\theta))x'_2]) (B + I_4 B)} f(x'_1, x'_2) \\ &\quad e^{-\frac{1}{2}([(u_1 \cos(\theta) + u_2 \sin(\theta))x'_1 + (-u_1 \sin(\theta) + u_2 \cos(\theta))x'_2]) (B + I_4 B)} dx'_1 dx'_2 \\ &= \widehat{f}_B(u_1 \cos(\theta) + u_2 \sin(\theta), -u_1 \sin(\theta) + u_2 \cos(\theta)) \end{aligned} \quad (2.46)$$

□

Les coordonnées fréquentielles subissent donc une rotation d'angle θ . En intégrant sur l'ensemble des valeurs de θ , l'invariance en rotation est obtenue.

2.4.7 Comparaison des descripteurs couleur

Le premier constat que nous pouvons faire est que toutes ces méthodes (à part la méthode marginale) font le choix d'une direction d'analyse. Les signatures "couleur" de [Kunttu 05] utilisent une statistique (moyenne d'un canal couleur par exemple) calculée sur l'information colorimétrique, provoquant une perte d'information. La méthode marginale est meilleure que cette approche précédente car, même si elle implique une redondance des informations colorimétriques, elle prend en compte l'ensemble de cette information et pas seulement une agrégation de celle-ci. La méthode de projection sur un plan couleur introduit la notion de plan d'analyse. L'utilisateur doit choisir la couleur sur laquelle il veut réaliser son analyse. Cette méthode est intéressante mais en ne considérant qu'un seul plan, toute l'information fréquentielle couleur n'est pas considérée.

Les approches que nous privilégions dans ce document sont basées sur les quaternions et les algèbres de Clifford. Elles correspondent à une généralisation à la couleur des méthodes classiques pour les images en niveaux de gris. Dans ce cas, toute l'information colorimétrique étant considérée, les taux de reconnaissance ne devraient en être que meilleurs. De plus, un gain en complexité algorithmique par rapport à la méthode marginale (deux transformées de Fourier rapides au lieu de trois) est obtenu.

Le nombre de descripteurs est aussi réduit (1 ensemble pour les *GCFD3* au lieu de 3 pour la méthode marginale). Par exemple, pour une image 128×128 , le nombre maximal de descripteurs est de $64 * 3$ pour les *GFD1* avec la méthode marginale, $64 * 2$ pour les *GCFD1* et seulement 64 pour les *GCFD3*. L'utilisateur peut choisir d'extraire moins de descripteurs mais cela engendre une perte d'information fréquentielle.

L'une des problématiques de ces méthodes reste le choix d'une direction d'analyse donnée par le quaternion pur et unitaire μ_1 ou le bivecteur B . Celui-ci a une influence sur les résultats obtenus pour certains descripteurs (voir la partie expérimentation).

2.5 Tableau récapitulatif

Dans le tableau 2.1, nous récapitulons les différents types de descripteurs de Fourier niveaux de gris et couleur. Il est spécifié pour chaque méthode si les descripteurs sont issus de signatures de formes, d'images en niveaux de gris ou d'images couleur, y figure aussi la taille, le coût algorithmique, les invariances de ces descripteurs. Les deux dernières colonnes précisent si une interpolation de l'image en coordonnées polaires ou log-polaires et si la méthode dépend de l'estimation du centre de l'image.

Descripteurs	Signature/Image	Taille	Coût algorithmique	Invariances	Interp.	Est. centre
FD'	Signature	$N_c/2$	1 FFT 1D	Tr., Rot., Éch.	×	✓
{P,R}	Signature	$2N_c$	1 FFT 1D	Tr., Rot.	×	×
DFG	Image ndg	$M_i N_i$	1 FFT 2D + interp. polaire	Tr., Rot.	✓	✓
Φ	Image ndg	K_1	1 FFT 1D + 1 trans. Mellin	Tr., Rot., Éch.	✓	✓
Φ'	Image ndg	K_1	1 FFT 1D + 1 trans. Mellin	Tr., Rot., Éch.	✓	✓
Φ_a	Image ndg	K_2	1 FFT 2D + 1 interp. log-polaire	Tr., Rot., Ech.	✓	✓
GFD1	Image ndg	$\min(M_i, N_i)/2$	1 FFT 2D + integr. cercle discret	Tr., Rot.	×	×
GFD2	Image ndg	K_3	1 FFT 2D + integr. cercle discret	Tr., Rot.	×	×
S_c	Signature couleur	N	1 FFT 1D	Tr., Rot., Éch.	×	✓
Méthode marg.	image couleur	3 ensembles	3 FFT 2D	*	*	*
Proj. couleur	image couleur	1 ensemble	1 FFT 2D	*	*	*
ϕ_{μ_1}	image couleur	K_1	2 FFT 1D + 1 trans. Mellin	Tr., Rot., Éch.	×	✓
GCFD ₁	image couleur	$\min(M_i, N_i)$	2 FFT 2D + integr. cercle discret	Tr., Rot.	×	×
GCFD ₂	image couleur	$2K_3$	2 FFT 2D + integr. cercle discret	Tr., Rot.	×	×
GCFD ₃	image couleur	$\min(M_i, N_i)/2$	2 FFT 2D + reconstruction + integr. cercle discret	Tr., Rot.	×	×

TABLE 2.1 – Tableau de synthèse des descripteurs de Fourier niveaux de gris et couleur

- N_c : nombre de points d'un contours
- M_i et N_i : hauteur et largeur d'une image
- Tr., Rot., Éch. : translation, rotation et changement d'échelle.
- K_1 : choix laissé à l'utilisateur. Dans la littérature, un choix typique est $V \leq 25$ et $K \leq 5$.
- K_2 : choix laissé à l'utilisateur. Cependant, si l'algorithme rapide de [Derrode 01] est choisi, le nombre de descripteurs est borné par l'échantillonnage de l'image log-polaire.
- K_3 : choix laissé à l'utilisateur. Il dépend du choix des vecteurs ξ_1 et ξ_2 .
- Interp. : interpolation de l'image
- Est. centre : estimation du centre de l'image
- * : dépend du type de descripteurs extraits
- ✓ : une étape requise pour la construction des descripteurs
- × : une étape qui n'est pas requise pour la construction des descripteurs

2.6 Conclusion

Dans la première partie de ce chapitre, pour chacune des transformées de Fourier définies dans le chapitre 1, un état de l'art original des définitions des descripteurs a été présenté. Notre attention s'est portée sur leurs différentes propriétés afin de montrer leurs invariances. Pour guider la construction de nouveaux descripteurs de Fourier, dans chacun des cas, les avantages et les inconvénients sont étudiés.

Les descripteurs de Fourier de Smach *et al.*, invariants en translation et en rotation par construction, ont permis la définition de deux nouveaux ensembles de descripteurs de Fourier Clifford couleur. Ces derniers sont construits à partir des parties parallèle et orthogonale de la transformée de Fourier Clifford couleur de Batard *et al.*.

Motivé par la réduction du nombre de descripteurs, nous avons défini un nouvel ensemble de descripteurs à partir de la transformée de Fourier Clifford reconstruite. Ainsi, leur nombre a été divisé par deux. La preuve des invariances ainsi que le rapprochement entre les *GCFD1* et les *GCFD3* a été montré.

Afin d'avoir une meilleure visibilité parmi tous ces descripteurs, un tableau de synthèse original mettant en avant leur taille, leur coût algorithmique et leurs invariances a été présenté.

Les trois ensembles de descripteurs de Fourier Clifford couleur doivent maintenant être mis en concurrence. Le chapitre suivant se consacrera donc, d'une part, à établir une comparaison. Celle-ci s'effectuera entre notre approche, la méthode marginale et les moments de Fourier-Mellin quaternioniques.

CHAPITRE 3

APPLICATIONS À LA RECONNAISSANCE D'IMAGES COULEUR

Applications à la reconnaissance d'images couleur

Sommaire

3.1	Évaluation des descripteurs de Fourier couleur	98
3.1.1	Les bases d'images couleur	98
3.1.2	Extraction des descripteurs	100
3.1.3	Classification	101
3.1.4	Comparaison des GFD , $GCFD$ et ϕ_{μ_1}	102
3.1.4.1	Sur la base COIL-100	102
3.1.4.2	Sur la base color FERET	104
3.1.4.3	Sur la base color FERET bruitée	105
3.1.5	Choix du bivecteur	105
3.1.5.1	Sur la base COIL-100	107
3.1.5.2	Sur la base color-FERET	107
3.1.5.3	Sur la base ALOI	108
3.1.6	Discussion	108
3.2	Une application de la transformée de Fourier : la corrélation de phase	110
3.2.1	La corrélation de phase pour les images en niveaux de gris	110
3.2.2	La corrélation de phase pour les images couleur	112
3.2.2.1	La corrélation de phase quaternionique couleur	112
3.2.2.2	Proposition d'une corrélation de phase Clifford couleur	114
3.2.3	Comparaison des corrélations de phase quaternionique et Clifford	117
3.2.3.1	Expérimentations sur des images synthétiques	117
3.2.3.2	Utilisation dans le cadre de la reconnaissance d'images	118
3.2.3.3	Expérimentations sur la base COIL-100	119
3.2.3.4	Expérimentations sur la base color FERET	120
3.2.4	Discussion	121
3.3	Conclusion	121

Dans le premier chapitre, un état de l'art des différentes transformées de Fourier en niveaux de gris et en couleur a été présenté. Suite à cette étude, la transformée de Fourier Clifford couleur est apparue comme la plus prometteuse. En effet, elle permet de traiter l'information couleur dans sa globalité, *i.e.* non marginalement. Ensuite, dans un deuxième chapitre, les principaux descripteurs de Fourier de la littérature sont abordés. Parmi ceux-ci, les descripteurs de Smach *et al.* ont permis de définir trois nouveaux ensembles de descripteurs : les $GCFD$. Ces descripteurs, construits à partir de la transformée de Fourier Clifford couleur, sont invariants en translation

et en rotation. Contrairement à une méthode marginale, ils contiennent moins de redondances fréquentielles, sont plus compacts et ont un coût algorithmique moindre.

Dans l'objectif de comparer nos ensembles de descripteurs couleur aux principaux de la littérature, une série d'expérimentations complexes est menée. Cette dernière comparera la méthode marginale, les moments de Fourier-Mellin quaternioniques et notre approche en utilisant des bivecteurs B prédéfinis en termes de taux de reconnaissance.

Les expérimentations ainsi menées mettront en évidence que l'ensemble des nouveaux descripteurs est sensible au choix du bivecteur B . Il apparaît donc primordial de définir le meilleur choix pour la direction d'analyse de la transformée de Fourier Clifford. Ainsi, une deuxième série d'expérimentations est menée afin de répondre à cette problématique.

Dans les deux premières mises en pratique de la transformée de Fourier Clifford, il apparaît que celles-ci montrent la pertinence de cette méthode géométrique pour la reconnaissance de formes dans le cadre couleur. Cette performance est-elle valable pour d'autres méthodes fréquentielles? Pour répondre à cette interrogation, nous proposons de définir une corrélation de phase couleur. La corrélation de phase est une méthode applicative dérivée de la transformée de Fourier, c'est pour cela qu'elle est abordée dans ce chapitre. Sa définition usuelle pour les images en niveaux de gris est rappelée. Nous observerons alors que le "Shift-Theorem" est nécessaire à la construction de la corrélation de phase. Ainsi, nous définirons celui de la transformée de Fourier Clifford couleur. Pour calculer directement cette corrélation de phase couleur, on devrait disposer d'une transformée de Fourier Clifford couleur inverse pour des fonctions multivectorielles générales. Deux méthodes sont proposées dans la fin de ce chapitre afin de lever ce verrou technique. La première se base sur le calcul d'une corrélation de phase classique sur les deux parties parallèle et orthogonale de la transformée de Fourier Clifford couleur. Alors que la seconde est définie sur la transformée de Fourier Clifford reconstruite.

Une nouvelle série d'expérimentations est alors réalisée dans l'objectif de comparer nos deux méthodes avec la corrélation de phase usuelle et quaternionique. La pertinence de ces méthodes est étudiée au travers des taux de reconnaissance et des scores de corrélation.

3.1 Évaluation des descripteurs de Fourier couleur

Pour évaluer les descripteurs de Fourier couleur $GCFD$ que nous avons définis dans la partie 2.4.5, nous mettons en place différentes expérimentations sur des bases d'images couleur standards. Nous discutons du choix du bivecteur B , paramètre de la transformée de Fourier Clifford couleur, et des performances respectives de ces descripteurs en termes de taux de reconnaissance. Dans ce qui suit, nous décrivons le protocole expérimental que nous avons mis en place pour tester notre approche.

3.1.1 Les bases d'images couleur

Les bases d'images utilisées dans cette thèse sont les bases COIL-100 [Nene 96], color FERET [Phillips 98] et ALOI [Geusebroek 05]. Pour vérifier la robustesse des descripteurs au bruit, une

version bruitée de la base color FERET est également utilisée.

- La base d’images **COIL-100 (Columbia Object Image Library)** est composée de 7200 images couleur de taille 128×128 de 100 objets différents. Chaque image est composée d’un fond noir et chaque objet est pris sous 72 angles de vue différents. Cette base, utilisée dans des travaux similaires [Smach 08], peut être qualifiée de “facile” car le seul contenu fréquentiel est celui de l’objet d’intérêt. Nous pouvons ajouter que les images qui la composent ont un contenu très géométrique.



FIGURE 3.1 – Les 100 objets de COIL-100 (à gauche) - une image de la base (à droite)

- La base d’images color **FERET** est composée d’images de portraits de 1408 personnes différentes, prises sous différents angles de vue. Dans nos expérimentations, un ensemble de 2992 images contenant 272 personnes représentées de manière égale par 11 images est sélectionné. La taille de ces images est réduite à 128×128 . Cette base d’images est plus difficile que la première à cause des changements d’illumination du fond.



FIGURE 3.2 – Quelques images de la base color FERET

- La base color **FERET bruitée** est dérivée de la base color FERET, mais un bruit gaussien est ajouté indépendamment à chaque plan couleur des images, l’écart type en est fixé à 0.23, qui est le niveau de bruit maximum utilisé dans [Smach 08].
- La base d’images **ALOI (Amsterdam Library of Object Images)** est composée de 1000 objets et chacun d’entre eux est pris sous 72 angles de vue différents. Pour nos expérimentations, nous avons limité le nombre d’objets aux 100 premiers, ce qui nous donne un total de 7200 images. Comme pour la base COIL-100, cette base de donnée peut être



FIGURE 3.3 – Quelques images de la base color FERET bruitée

qualifiée de simple car le fond de l'image est noir. La taille des images est réduite à 128×128 .



FIGURE 3.4 – Quelques images de la base ALOI

3.1.2 Extraction des descripteurs

Pour tous les ensembles de descripteurs de Fourier que nous considérons, nous avons fait les choix suivants :

Concernant les descripteurs de type *GFD* [Smach 08], 64 descripteurs sont extraits de chaque canal couleur :

- Pour les *GFD1*, l'intégration est faite sur des cercles de rayons dans $[1, 64]$.
- Pour les *GFD2*, le vecteur ξ_1 est fixé à $(0, 1)^t$ et les valeurs de ξ_2 sont également réparties dans le domaine polaire $[0, 2\pi] \times [1, 8]$, permettant ainsi de prendre en compte les basses fréquences de l'image.
- Calculés marginalement, $192 = 64 * 3$ descripteurs *GFD1* et *GFD2* sont obtenus.

Pour les *GCFD1* et *GCFD2*, la longueur des vecteurs de descripteurs sont de 64×2 (correspondant à ceux de la partie parallèle et de la partie orthogonale de la transformée de Fourier Clifford couleur).

Le nombre de descripteurs *GCFD3* extraits est de 64. Cela correspond à 64 valeurs de rayon r dans $I_B^{(0,r)^t}$

Dans les but de comparer nos descripteurs avec un autre ensemble de descripteurs couleur, nous avons extrait autant de descripteurs *GCFD* que de descripteurs de Fourier-Mellin quaternioniques. Le choix de 64 descripteurs correspond au choix de $K = [1, \dots, 8]$ et $V = [1, \dots, 8]$. Le choix de 128 descripteurs correspond à celui de $K = [1, \dots, 12]$ et $V = [1, \dots, 11]$. Dans ce cas, nous ne considérons pas les 4 derniers descripteurs (correspondant aux valeurs $k = 12$ et $v = [8, \dots, 11]$) car $11 * 12 = 132$.

Concernant les descripteurs de Fourier-Mellin quaternioniques, nous n'avons pas redéfini automatiquement l'origine des images polaires pour les bases COIL-100 et ALOI car les objets sont centrés. Cette origine se situe donc au centre des images dans le domaine spatial. Par contre, concernant la base color FERET, nous présentons les résultats obtenus avec et sans l'étape de redéfinition du centre proposée dans [Guo 11]. Le quaternion pur $\mu_1 = (i + j + k)/\sqrt{3}$ est choisi pour le calcul de ces descripteurs car il correspond à un choix classique dans la littérature.

3.1.3 Classification

L'étape de classification est réalisée en utilisant un algorithme de Séparateur Vaste Marge ("Support Vector Machine" *SVM*) standard [Vapnik 98] dans l'implémentation LIBSVM [Chang 01]. Les *GFD*, les *GCFD* et les ϕ_{μ_1} sont utilisés comme des vecteurs d'attributs du classifieur. Ceux-ci sont préalablement normalisés pour appartenir à l'intervalle $[-1; 1]$ pour ramener la dynamique à une même plage pour tous à l'aide de la formule suivante :

$$D' = \left(\frac{D - D_{min}}{D_{max} - D_{min}} * 2 \right) - 1 \quad (3.1)$$

où D est un descripteur et D' est le descripteur normalisé, D_{min} et D_{max} représentent les valeurs minimale et maximale de ce descripteur calculé sur toute la base d'image.

Dans nos tests, le noyau classique à base radiale (noyau gaussien) est utilisé :

$$K(\mathbf{x}, \mathbf{y}) = e^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}} \quad (3.2)$$

où \mathbf{x} et \mathbf{y} sont les vecteurs de descripteurs des objets, σ est la largeur de bande.

Les paramètres σ et C (le coefficient de pénalisation) ont été empiriquement déterminés par une 10-validation croisée [Duda 01] pour maximiser le taux de reconnaissance pour chaque expérimentation [Chang 01]. Pour cela, les taux de reconnaissance sont calculés avec un grand nombre de couple de paramètres (C, σ) ($C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ et $\sigma = 2^{-15}, 2^{-13}, \dots, 2^3$). Le couple (C, σ) qui est alors choisi est celui qui donne les meilleurs résultats. Cette méthode est connue sous le nom de "grid search" [Chang 01].

Les paramètres σ et C ne changent que très peu. Nous avons choisi les paramètres ($\sigma = 0.1, C = 97$) pour les expérimentations sur les *GFD1*, *GFD2*, *GCFD1*, *GCFD2*, les paramètres ($\sigma = 0.3, C = 110$) pour *GCFD3* et ($\sigma = 0.1, C = 110$) pour les expérimentations avec les moments de Fourier-Mellin quaternioniques.

3.1.4 Comparaison des *GFD*, *GCFD* et ϕ_{μ_1}

Les résultats des expérimentations sur chaque base d'images sont reportés dans les tableaux 3.1, 3.3, 3.5. Dans tous les cas, nous avons sélectionné les mêmes bivecteurs. Le premier bloc de résultats montre les taux de reconnaissance pour un bivecteur choisi par l'utilisateur ($B_r = \mathbf{r} \wedge \mathbf{e}_4, B_v = \mathbf{v} \wedge \mathbf{e}_4, B_b = \mathbf{b} \wedge \mathbf{e}_4, B_\mu = \boldsymbol{\mu} \wedge \mathbf{e}_4$, μ étant l'axe des niveaux de gris) ou aléatoire ($B_{rand} = rand \wedge \mathbf{e}_4$ avec $rand$ un vecteur unitaire du cube *RVB*). Dans ce dernier cas, nous avons calculé la moyenne et l'écart type du taux de reconnaissance pour 100 bivecteurs aléatoires.

Les deux dernières lignes concernent des résultats avec trois bivecteurs. La première, notée $B_r + B_v + B_b$, utilise comme vecteur de descripteurs la concaténation des descripteurs calculés à partir de B_r, B_v et B_b (des descripteurs de taille 384 pour les *GCFD*). Dans le cas des *GFD*, cela correspond à la méthode marginale.

La dernière ligne, notée $B_1 + B_2 + B_3$, montre les résultats obtenus en utilisant un algorithme de sélection automatique pour trouver les trois bivecteurs maximisant les taux de reconnaissance. L'algorithme dit "SFFS" (Sequential Floating Forward Selection) est un algorithme de sélection sous-optimal qui évite une recherche exhaustive dans l'espace des caractéristiques (pour plus d'informations voir [Jain 97]). Ici, nous l'avons modifié pour réaliser une sélection par "groupe" de descripteurs dans l'objectif de choisir les meilleurs bivecteurs en termes de taux de reconnaissance. C'est-à-dire qu'à chaque étape de notre algorithme de "SFFS" par groupe, nous ajoutons (ou retirons) tout un ensemble de descripteurs correspondant à un bivecteur B donné. Les groupes de descripteurs sélectionnés représentent alors les meilleurs choix de bivecteurs.

La taille des descripteurs est rappelée et le meilleur résultat pour chaque méthode est surligné. La couleur jaune met en évidence les meilleurs résultats pour un seul bivecteur et la couleur verte correspond aux résultats obtenus pour 3 bivecteurs (par concaténation ou par un algorithme *SFFS*).

3.1.4.1 Sur la base COIL-100

Les tableaux 3.1 et 3.2 contiennent les taux de reconnaissance de la base COIL-100. Cette base étant "facile" d'un point de vue classification, chaque ensemble de descripteur donne d'excellents résultats, très proches de 100%. Notons que de tels taux de classification valident le choix des descripteurs de Fourier pour ce genre d'images où la couleur du fond est homogène et similaire pour toutes les images. Les écarts types obtenus pour B_{rand} sont faibles et suggèrent la faible influence du choix du bivecteur pour cette base d'images.

Concernant les *GCFD3*, nous remarquons qu'ils sont invariants au choix du bivecteur. En effet, quelque soit le bivecteur choisi, le taux de reconnaissance est le même. Les descripteurs *GCFD3*, avec seulement 64 descripteurs, donnent un meilleur résultat que les *GFD1* pour n'importe quel choix de bivecteur et que des *GCFD1* avec le bivecteur B_μ . De plus, ce taux est très proche du meilleur taux de reconnaissance obtenu avec les *GFD2*. Par contre, les descripteurs *GCFD1* et *GCFD2*, composés de 128 descripteurs, restent meilleurs en général.

Les descripteurs de Fourier-Mellin quaternioniques sont meilleurs que les *GFD1* avec le même nombre de descripteurs mais sont moins performants que les *GCFD*.

COIL-100					
Bivecteurs	<i>GFD1</i> (64 desc.)	<i>GCFD1</i> (128 desc.)	<i>GFD2</i> (64 desc.)	<i>GCFD2</i> (128 desc.)	<i>GCFD3</i> (64 desc.)
$B_r = r \wedge e_4$	98.04	99.83	98.69	99.81	99.39
$B_v = v \wedge e_4$	98.06	99.56	99.39	99.85	
$B_b = b \wedge e_4$	96.90	99.86	94.03	99.85	
$B_\mu = \mu \wedge e_4$	98.49	99.25	99.40	99.47	
$B_{rand}(\times 100)$	98.42 ± 0.3	99.54 ± 0.3	98.88 ± 0.82	99.82 ± 0.1	
max.	98.87	99.89	99.47	99.96	

	192 desc.	384 desc.	192 desc.	384 desc.
$B_r + B_v + B_b$	99.9	99.92	99.89	99.87
$B_1 + B_2 + B_3$ (SFFS)	99.86	99.96	99.89	99.96

TABLE 3.1 – COIL-100 : Taux de reconnaissance en % avec *GFD1*, *GCFD1*, *GFD2*, *GCFD2* et *GCFD3*

COIL-100		
$\mu_1 = \frac{(i+j+k)}{\sqrt{3}}$	64 desc.	128 desc.
ϕ_{μ_1}	99.29	99.47

TABLE 3.2 – COIL-100 : Taux de reconnaissance en % avec ϕ_{μ_1} .

En dehors des choix de bivecteur tels B_r , B_v et B_b , toute couleur *RVB* normalisée peut être utilisée pour construire le bivecteur. Pour définir un meilleur choix de bivecteur pour le calcul des *GCFD1*, nous montrons dans la figure 3.5 les taux de reconnaissance obtenus avec 100 bivecteurs aléatoires. La couleur de chacun des points correspond à la couleur choisie pour la construction du bivecteur.

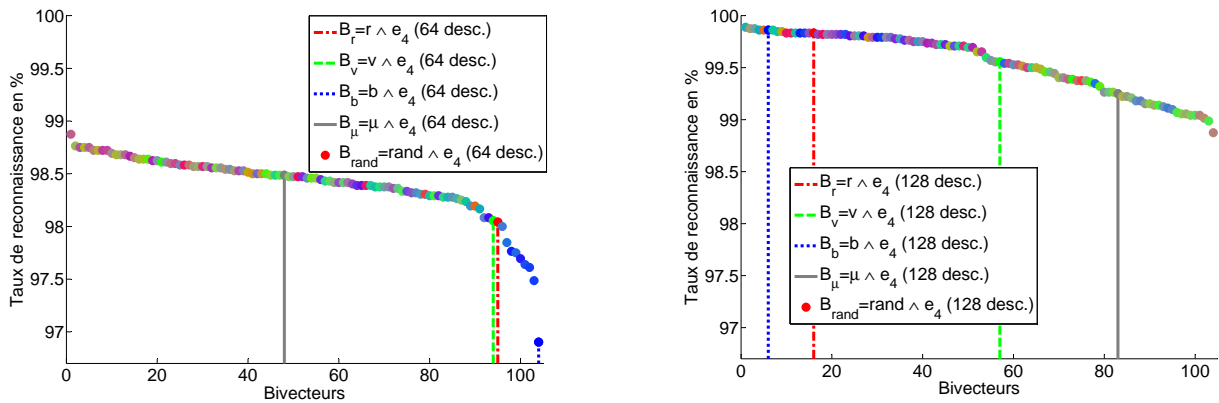


FIGURE 3.5 – COIL-100 : Taux de reconnaissance pour *GFD1* (à gauche) et *GCFD1* (à droite) avec différentes valeurs de B .

En considérant les *GFD1* seulement, on peut remarquer que le canal vert donne un meilleur taux de reconnaissance parmi les trois plans couleur mais reste inférieur à celui donné par μ . De plus, en choisissant un vecteur couleur aléatoire, on améliore quasi-systématiquement (dans

81% des cas) le taux de reconnaissance obtenu avec le bivecteur B_v (celui-ci donnant le meilleur résultat parmi B_r , B_v et B_b). Le choix d'un bivecteur construit avec l'une des trois couleurs marginales n'est donc pas un bon choix en général pour les $GFD1$.

En considérant les $GCFD1$, on double la taille des descripteurs et les taux de reconnaissance sont toujours supérieurs à ceux des $GFD1$. Nous pouvons en conclure que la partie orthogonale apporte beaucoup d'information colorimétrique. Précisons que pour B_μ , la partie orthogonale est égale au plan de chrominance, ce qui est analogue à l'approche de Ell et Sangwine [Ell 07]. Par ailleurs, il est important de remarquer qu'il n'existe que peu de bivecteurs choisis aléatoirement qui améliorent le résultat obtenu par les deux meilleures marginales (rouge et bleu).

3.1.4.2 Sur la base color FERET

Sur la base color FERET (voir tableau 3.3), les descripteurs $GCFD1$ sont plus performants que les $GFD1$ et $GFD2$ en termes de taux de reconnaissance pour n'importe quel choix de bivecteur. Les $GCFD2$, tout en fournissant des meilleurs résultats que les $GFD1$ et $GFD2$ ne sont pas plus performants que les $GCFD1$. Cela peut venir du choix des paramètres ξ_1 et ξ_2 , mais aucune de nos expérimentations pour trouver des valeurs pertinentes n'a conduit à une amélioration concernant ces paramètres.

Bivecteurs	color FERET				
	$GFD1(64 \text{ desc.})$	$GCFD1(128 \text{ desc.})$	$GFD2(64 \text{ desc.})$	$GCFD2(128 \text{ desc.})$	$GCFD3(64 \text{ desc.})$
B_r	76.70	87.90	77.31	84.42	75.60
B_v	73.66	79.65	77.37	80.01	
B_b	70.49	84.49	75.87	82.31	
B_μ	73.03	78.10	77.30	82.12	
$B_{rand}(\times 100)$	73.72 ± 1	85.34 ± 2.92	77.54 ± 0.69	84.50 ± 2.06	
max.	76.14	90.37	78.91	89.57	

	192 desc.	384 desc.	192 desc.	384 desc.
$B_r + B_v + B_b$	88.03	85.53	84.26	82.55
$B_1 + B_2 + B_3(\text{SFFS})$	85.46	93.15	82.89	90.07

TABLE 3.3 – Color FERET : Taux de reconnaissance en % avec $GFD1$, $GCFD1$, $GFD2$, $GCFD2$ et $GCFD3$

$\mu_1 = \frac{(i+j+k)}{\sqrt{3}}$	color FERET	
	64 desc.	128 desc.
ϕ_{μ_1} sans estimation du centre	56.52	58.92
ϕ_{μ_1} avec estimation du centre	38.53	39.34

TABLE 3.4 – color FERET : Taux de reconnaissance en % avec ϕ_{μ_1} .

Dans tous les cas, ces résultats montrent clairement l'apport de l'utilisation de la transformée de Fourier Clifford en indexation d'images. La concaténation ($B_r + B_g + B_b$) des bivecteurs améliore les résultats des GFD mais pas ceux des $GCFD$. Cela est probablement dû au meilleur encodage de l'information couleur effectué par la transformée de Fourier Clifford. Les descripteurs $GCFD$ obtenus avec les bivecteurs B_r , B_v et B_b sont très corrélés. C'est pourquoi le choix

des trois meilleurs bivecteurs aléatoires, différents de B_r , B_v et B_b améliore les résultats. Dans l'expérimentation avec des choix de bivecteurs aléatoires, l'écart type s des taux de reconnaissance calculés à partir des $GCFD1$ ($s = 2.92$) et $GCFD2$ ($s = 2.06$) est un peu plus important en comparaison avec ceux des $GFD1$ ($s = 1$) et $GFD2$ ($s = 0.69$). Cela montre clairement que la qualité des $GCFD$ dépend du choix du bivecteur.

Pour les descripteurs $GCFD3$, le taux de reconnaissance est proche du meilleur résultat obtenu avec les $GFD1$. Le résultat est clairement moins bon qu'avec les $GCFD1$ et $GCFD2$ car les informations fréquentielles couleur du fond et de la forme se retrouvent mélangées. L'utilisation des $GCFD3$ pour des images où l'objet d'intérêt n'est pas segmenté n'est donc pas un bon choix.

Les descripteurs de Fourier-Mellin quaternioniques donnent de mauvais résultats (voir tableau 3.4). En effet, comme nous l'avons évoqué dans le chapitre précédent, ces descripteurs dépendent d'une étape d'estimation automatique du centre de l'image. Or, dans la base color FERET, ce centre est difficile à estimer car les images qui la composent n'ont pas un arrière plan homogène. Le centre estimé ne correspond alors pas forcément au centre de l'objet d'intérêt de ces images : le visage des personnes photographiées. Dans le tableau 3.4, la première ligne correspond aux résultats obtenus sans estimation automatique du centre (l'origine de l'image polaire est le centre de l'image). La seconde concerne les résultats obtenus avec l'estimation automatique du centre de l'image décrite dans [Guo 11]. Nous pouvons remarquer que l'estimation automatique du centre dégrade encore plus les résultats.

Pour étudier l'influence du choix du bivecteur sur les $GCFD1$, les résultats d'une expérimentation avec des bivecteurs aléatoires sont détaillés dans la figure 3.6. On remarque que les meilleurs choix ont une dominante bleue. Comme le bivecteur est unique pour une base d'images donnée, il doit être choisi soit en fonction de connaissances *a priori* de la base, soit selon une méthode de recherche empirique. Dans la partie 3.1.5 des expérimentations sont réalisées pour définir un bon choix de bivecteur.

3.1.4.3 Sur la base color FERET bruitée

L'influence du bruit sur les taux de reconnaissance est illustré dans le tableau 3.5. Les descripteurs $GFD1$ sont les plus sensibles au bruit et, comme prévu (voir [Smach 08]), les $GFD2$ sont les plus robustes. Nous faisons le même constat entre les $GCFD1$ et les $GCFD2$.

Là encore, les $GCFD3$ donnent de bien meilleurs résultats que les $GFD1$. Cependant, ils sont moins robustes au bruit que les $GFD2$ et $GCFD2$.

Les taux obtenus avec les ϕ_{μ_1} n'ont que peu diminué montrant une résistance relative au bruit, mais ils restent bas par rapport aux GFD et $GCFD$.

3.1.5 Choix du bivecteur

Dans cette partie, le choix du bivecteur est discuté. Pour essayer de trouver le meilleur choix de bivecteur, nous avons effectué une nouvelle expérimentation en choisissant 100 nouveaux bivecteurs, notés B_T , dans l'espace TSL (Teinte-Saturation-Luminance) par valeur croissante de

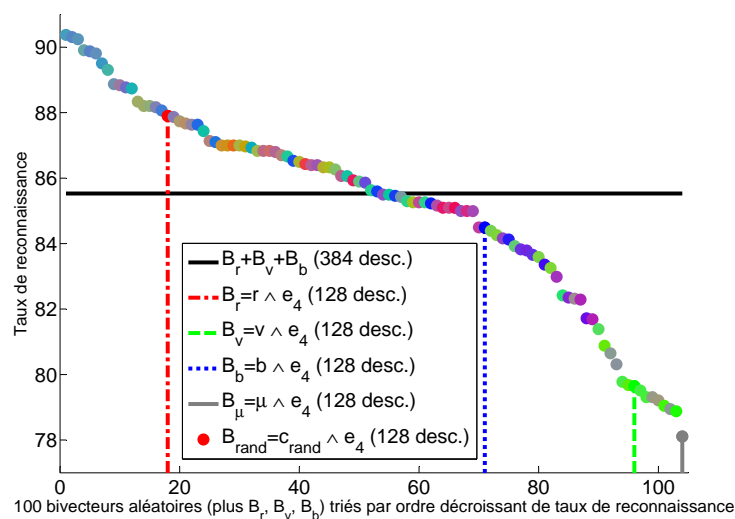


FIGURE 3.6 – color FERET : Taux de reconnaissance avec GCFD1 pour 100 bivecteurs aléatoires

Bivecteurs	color FERET bruitée				
	GFD1(64 desc.)	GCFD1(128 desc.)	GFD2(64 desc.)	GCFD2(128 desc.)	GCFD3 (64 desc.)
B_r	45.32	71.05	73.46	83.49	61.66
B_v	46.83	61.99	75.26	78.64	
B_b	48.49	73.46	74.77	81.78	
B_μ	55.28	62.03	77.34	80.98	
$B_{rand}(\times 100)$	54.23 ± 1.75	69.64 ± 3.21	76.59 ± 0.74	82.56 ± 1.80	
max.	57.55	77.27	78.41	87.00	

	192 desc.	384 desc.	192 desc.	384 desc.
$B_r + B_v + B_b$	73.16	72.16	83.25	81.12
$B_1 + B_2 + B_3(\text{SFFS})$	71.52	80.62	83.36	88.24

TABLE 3.5 – Color FERET bruitée : Taux de reconnaissance en % avec GFD1, GCFD1, GFD2, GCFD2 et GCFD3

$\mu_1 = \frac{(i+j+k)}{\sqrt{3}}$	color FERET bruitée	
	64 desc.	128 desc.
ϕ_{μ_1} sans estimation du centre	54.78	55.55
ϕ_{μ_1} avec estimation du centre	40.17	38.73

TABLE 3.6 – color FERET bruitée : Taux de reconnaissance en % avec ϕ_{μ_1} .

teinte ($T \in [0, 1]$) et pour $S = L = 1$. Cette expérimentation a été menée sur les bases COIL-100, color FERET et ALOI.

3.1.5.1 Sur la base COIL-100

La figure 3.7 contient les taux de reconnaissance obtenus avec les descripteurs $GCFD1$ sur la base COIL-100. Elle ne laisse pas apparaître de différences significatives ($\pm 0.2\%$) suivant la teinte sélectionnée. Sur cette même figure, nous avons ajouté les résultats pour le cas où les descripteurs sont issus des parties parallèle et orthogonale de trois bivecteurs. Dans le cas $B_r + B_v + B_b$, le taux de reconnaissance est très satisfaisant (99.92%) mais ne dépasse pas le meilleur résultat obtenu avec un bivecteur. La meilleure performance absolue (99.96%) est donnée par la sélection de 3 meilleurs bivecteurs parmi 100 avec la méthode “SFFS”. Cependant, dans ce cas, les descripteurs sont au nombre de 384.

Parmi les meilleurs bivecteurs, on retrouve les bivecteurs B_r et B_b et l’un des plus mauvais est B_v . Nous pouvons alors nous demander si les bivecteurs B_r et B_b donnent aussi des bons résultats pour une autre base et si les résultats avec B_v sont toujours aussi mauvais.

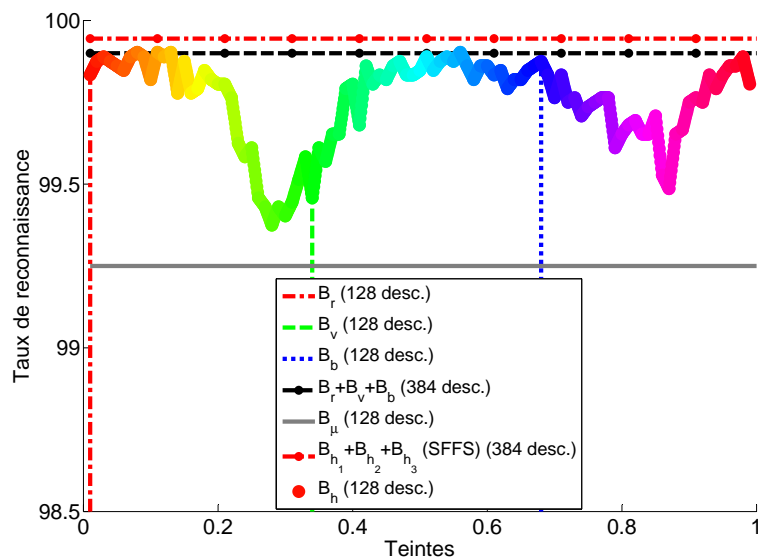


FIGURE 3.7 – COIL-100 : Taux de reconnaissance pour $GCFD1$ avec différentes valeurs de B_T , puis avec 3 bivecteurs (cas marginal et par SFFS)

3.1.5.2 Sur la base color-FERET

Comme pour la base de donnée COIL-100, nous avons testé les descripteurs $GCFD1$ sur la base color FERET en prenant 100 bivecteurs dans l’espace TS_L par valeur croissante de teinte ($T \in [0, 1]$) et pour $S = L = 1$. Les résultats apparaissent dans la figure 3.8.

Il est évident que les courbes des figures 3.7 et 3.8 ont la même allure générale. En effet, dans ces deux cas, le bivecteur B_v et les teintes de vert en général ne sont pas de bons choix. Par contre,

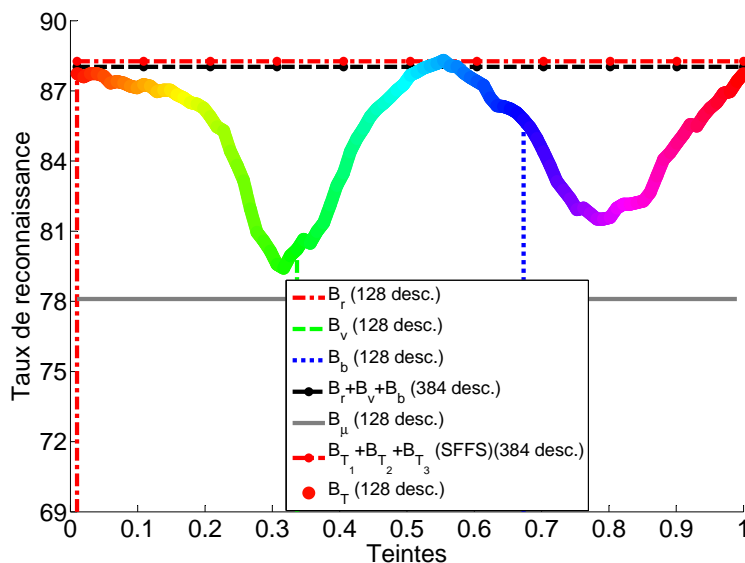


FIGURE 3.8 – color FERET : Taux de reconnaissance pour GCFD1 avec différentes valeurs de B_T , puis avec 3 bivecteurs (cas marginal et par SFFS)

les teintes de rouge et de bleu semblent être les meilleurs candidats pour ces bases d'images très différentes. Pour s'en convaincre, nous avons réalisé la même expérimentation sur la base ALOI.

3.1.5.3 Sur la base ALOI

Les résultats concernant la base ALOI sont présentés dans la figure 3.9. Là encore, la courbe obtenue est semblable à celles obtenues précédemment. Les meilleurs bivecteurs semblent se rapprocher du bivecteur B_b . Les teintes vertes restent des mauvais choix. Nous pouvons penser que cela est dû à la nature même des images naturelles. En effet, nous savons que ce type d'images contient une forte proportion de vert. Une grande majorité des caméras et appareils photos sont équipés de filtres de Bayer [Bayer 76] qui contiennent deux fois plus de capteurs d'intensité de vert que de bleu et de rouge. La raison de ce choix est la maximisation de la netteté dans la luminance, qui est principalement composée d'information verte. Choisir le bivecteur B_v revient à projeter l'image sur le vecteur vert dans la partie parallèle et sur le plan rouge-bleu dans la partie orthogonale. Or il semblerait que la partie orthogonale joue un rôle important. C'est pourquoi le choix B_v semble être un mauvais choix et B_b (respectivement B_r) un bon car ce dernier implique une projection sur le plan rouge-vert (respectivement vert-bleu) dans la partie orthogonale.

3.1.6 Discussion

Nous avons défini de nouveaux descripteurs de Fourier couleur à partir de la transformée de Fourier-Clifford couleur [Batard 10] et des descripteurs de Fourier généralisés [Smach 08]. Les meilleurs taux de reconnaissance obtenus avec les descripteurs GCFD1 et GCFD2 sont meilleurs que ceux de l'approche marginale. Leur complexité algorithmique est plus faible et la taille du

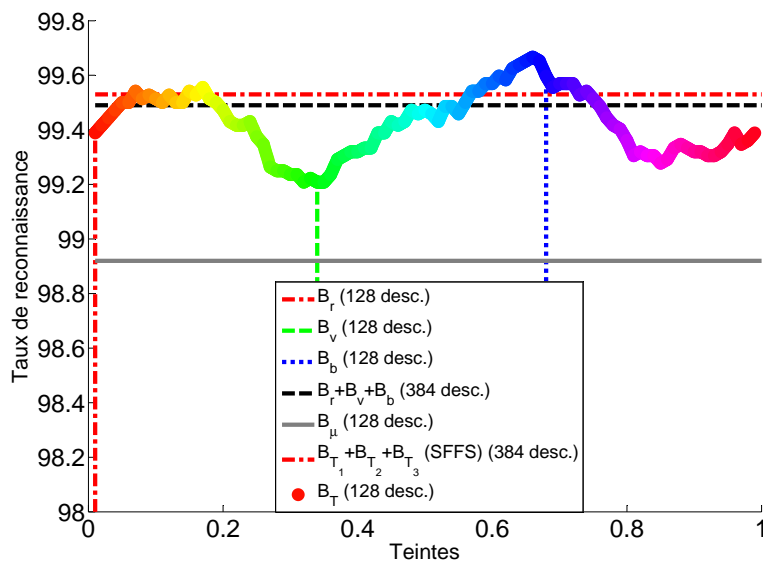


FIGURE 3.9 – ALOI : Taux de reconnaissance pour $GCFD1$ avec différentes valeurs de B_T , puis avec 3 bivecteurs (cas marginal et par SFSS)

vecteur de descripteurs est une fois et demie plus petite. Nous avons aussi montré que les descripteurs $GCFD2$ sont robustes au bruit.

Les descripteurs de Fourier $GCFD3$ donnent aussi de bons résultats. Même s'ils sont moins bons que ceux obtenus avec la méthode marginale, les $GCFD3$ permettent d'obtenir des taux supérieurs aux $GFD1$ avec le même nombre de descripteurs. De plus, avec un nombre égal de descripteurs, ils sont meilleurs que les descripteurs de Fourier-Mellin quaternioniques. Notons que les $GCFD3$ sont composés de trois fois moins de descripteurs que la méthode marginale et que ces descripteurs sont invariants au choix du bivecteur.

Les descripteurs de Fourier-Mellin quaternioniques ϕ_{μ_1} donnent des résultats comparables à ceux des $GCFD$ pour la base COIL-100. Par contre, ces descripteurs ne sont pas pertinents pour la base color FERET. En effet, les objets (visages) ne sont pas forcément centrés. De plus, les angles de vue sont très différents d'une image à l'autre. Il est donc difficile de définir l'origine de l'image polaire, ce qui pénalise beaucoup ces descripteurs.

Nous avons aussi mené des expérimentations pour tenter de définir un bon choix de bivecteur. Nous avons montré que pour les trois bases COIL-100, color FERET et ALOI, les bivecteurs B_c où c est une couleur de teinte bleu et rouge sont de bons choix. En général, quand c est de teinte verte, les résultats obtenus sont plus mauvais. Nous pensons que ce résultat est valable pour les images naturelles ou les objets manufacturés. Nous proposons alors à l'utilisateur d'utiliser un bivecteur de son choix si il a un *a priori* sur les images à classer. Sinon, nous lui conseillons de choisir un bivecteur B_c où c est une couleur de teinte bleu et rouge

3.2 Une application de la transformée de Fourier : la corrélation de phase

Dans la littérature, des méthodes dérivées de la transformée de Fourier sont largement utilisées dans le cadre de l'analyse et du traitement des images. Parmi celles-ci, la corrélation de phase est une méthode usuelle qui est utilisée dans un grand nombre d'applications comme la reconnaissance d'images [Reddy 96], la stabilisation d'images [Erturk 03], l'analyse de flots optique [Ebling 05], *etc.* La problématique de la corrélation de phase est d'estimer la meilleure translation entre deux images. Pour cela, l'outil le plus indiqué est la transformée de Fourier car elle est étroitement liée à l'action de la translation sur le domaine de définition de l'image (voir la partie 1.1.3). Nous savons que la translation d'une image implique un déphasage de la transformée de Fourier qui dépend des paramètres de translation. Le but est alors de récupérer le déphasage entre les transformées de Fourier de deux images pour pouvoir ensuite en extraire la translation. Définie pour les images en niveaux de gris, son extension à la couleur est problématique car l'information couleur est vectorielle et non scalaire.

Comme elle est basée sur une application directe du "Shift-Theorem" de la transformée de Fourier, sa définition dépend de la transformée de Fourier choisie. Il est théoriquement possible de calculer une corrélation de phase pour l'ensemble des transformées de Fourier définies dans les parties 1.1 et 1.2. Cependant, dans le cas de la transformée de Fourier-Mellin, c'est la rotation qui est estimée car la translation est supprimée en pré-traitement et une rotation de l'image implique une translation dans le domaine log-polaire.

Dans cette partie, nous nous intéressons à la définition de la corrélation de phase pour les images en niveaux de gris dans un premier temps. Puis, son extension aux images couleur à l'aide de la transformée de Fourier quaternionique et de la transformée de Fourier Clifford couleur est étudiée. Ces corrélations de phase sont ensuite testées dans une série d'expérimentations.

3.2.1 La corrélation de phase pour les images en niveaux de gris

La corrélation de phase est basée sur le "Shift-Theorem" de la transformée de Fourier. Rappelons ce théorème qui est défini dans la partie 1.1.1.2 :

Soit g , la version translatée de l'image f tel que $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{\Delta})$ où $\mathbf{\Delta}$ est un vecteur de translation, leurs transformées de Fourier \hat{f} et \hat{g} vérifient

$$\hat{g}(\mathbf{u}) = \hat{f}(\mathbf{u}) e^{-2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle} \quad (3.3)$$

où $\langle \cdot \rangle$ est le produit scalaire. Le déphasage entre \hat{f} et \hat{g} est obtenu à l'aide du "cross-power spectrum" qui s'écrit

$$\begin{aligned} R(\mathbf{u}) &= \frac{\hat{f}(\mathbf{u})\overline{\hat{g}(\mathbf{u})}}{|\hat{f}(\mathbf{u})\overline{\hat{g}(\mathbf{u})|}} = \frac{\hat{f}(\mathbf{u})\overline{\hat{f}(\mathbf{u})} e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle}}{|\hat{f}(\mathbf{u})\overline{\hat{f}(\mathbf{u})} e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle|}} \\ &= \frac{|\hat{f}(\mathbf{u})|^2 e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle}}{||\hat{f}(\mathbf{u})|^2 e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle|}} = \frac{e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle}}{|e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle|}} = e^{2i\pi\langle\mathbf{u},\mathbf{\Delta}\rangle} \end{aligned} \quad (3.4)$$

où l'opérateur $\bar{\cdot}$ est le conjugué complexe.

La corrélation de phase est alors idéalement, dans le cas où g est le translaté de f :

$$\begin{aligned}
 r(\mathbf{x}) &= \check{R}(\mathbf{u}) \\
 &= \int_{-\infty}^{+\infty} e^{2i\pi\langle\mathbf{u},\Delta\rangle} e^{2i\pi\langle\mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \\
 &= \int_{-\infty}^{+\infty} e^{2i\pi(\langle\mathbf{u},\mathbf{x}\rangle+\langle\mathbf{u},\Delta\rangle)} d\mathbf{u} \\
 &= \int_{-\infty}^{+\infty} e^{2i\pi\langle\mathbf{u},\mathbf{x}+\Delta\rangle} d\mathbf{u} \\
 &= \delta(\mathbf{x} + \Delta)
 \end{aligned} \tag{3.5}$$

où δ est la distribution de Dirac et \check{R} est la transformée de Fourier inverse de R . La meilleure translation estimée et le score de corrélation ρ sont donnés par $\Delta = -\operatorname{argmax}_{\mathbf{x}}(|r(\mathbf{x})|)$ et $\rho = \max_{\mathbf{x}}(|r(\mathbf{x})|)$. Le coefficient ρ est égal à 1 quand g est la version translatée de f et peut être utilisé comme un score de similarité entre les images. Notons que ρ est invariant seulement en translation. D'autres invariances peuvent être obtenues en passant dans le domaine log-polaire [Reddy 96].

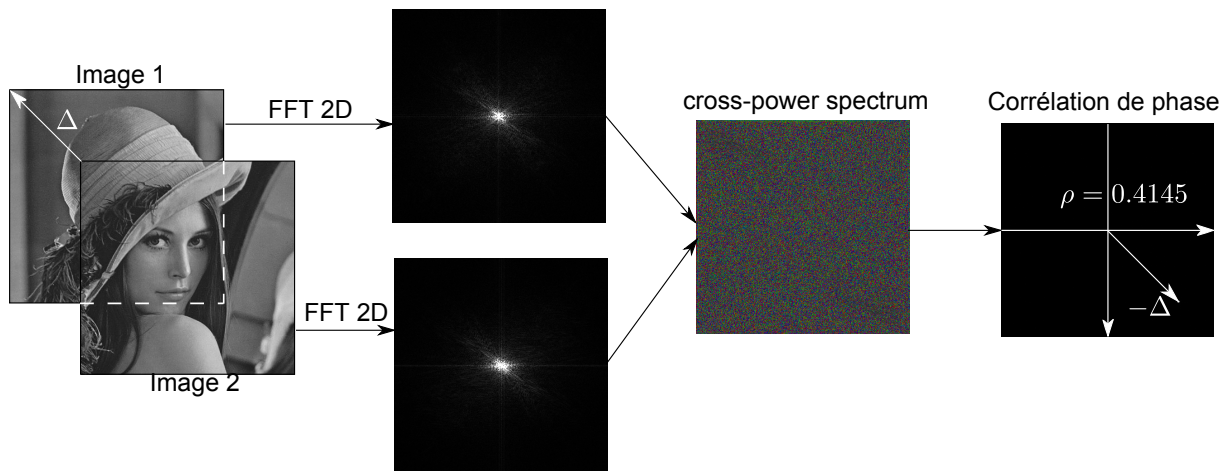


FIGURE 3.10 – La corrélation de phase pour les images en niveaux de gris

La figure 3.10 illustre la corrélation de phase pour deux images en niveaux de gris de Lenna. L'image 2 représente l'image 1 qui a été translatée d'un vecteur Δ . L'angle du "cross-power spectrum" est affiché dans un espace colorimétrique TSL . La corrélation de phase fournit un score de corrélation de 0.4145 et un vecteur de translation égal à $-\Delta$.

La corrélation de phase est une méthode bien définie pour les images en niveaux de gris et elle a été largement utilisée dans de nombreuses applications. Son extension à la couleur est problématique. En effet, une solution pour obtenir une corrélation de phase couleur à l'aide la transformée de Fourier 2D classique est de réaliser un traitement marginal des canaux RVB . Trois scores sont alors obtenus, nécessitant une agrégation pour obtenir un score unique. Une solution alternative donnée par Katuta *et al.* [Kakuta 97] correspond à une projection de l'image couleur sur un vecteur ou un plan couleur avant de calculer la corrélation de phase.

En utilisant une transformée de Fourier pour les images couleur (quaternionique ou Clifford), il est possible de définir rigoureusement une corrélation de phase couleur. La partie suivante traite de l'extension de la corrélation de phase aux images couleur à l'aide de la transformée de Fourier quaternionique et de la transformée de Fourier Clifford.

3.2.2 La corrélation de phase pour les images couleur

Pour réaliser une corrélation de phase couleur, la transformée de Fourier quaternionique ou la transformée de Fourier Clifford peuvent être utilisées. Sangwine *et al.* proposent dans [Sangwine 01] une corrélation de phase à partir de la transformée quaternionique à gauche et à droite. Concernant la transformée de Fourier Clifford, nous proposons deux méthodes pour calculer une corrélation de phase pour les images couleur.

3.2.2.1 La corrélation de phase quaternionique couleur

La première définition de la corrélation de phase quaternionique est donnée dans [Sangwine 01]. Pour définir une corrélation de phase couleur, Sangwine *et al.* se sont basés sur la définition classique de la corrélation de phase. En effet, celle-ci correspond à un calcul de la corrélation croisée entre images à l'exception de l'étape de normalisation du "cross-power spectrum". La première étape du calcul de la corrélation de phase quaternionique correspond alors au calcul de la corrélation croisée couleur quaternionique.

Pour cela, l'extension de la corrélation croisée classique aux fonctions complexes doit être considérée. Elle s'écrit

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \overline{g(\tau - t)} d\tau \quad (3.6)$$

Il est bien connu qu'une corrélation croisée de deux images dans le domaine spatial revient à faire une multiplication dans le domaine de Fourier. Cette méthode est alors utilisée pour calculer la corrélation croisée avec la transformée de Fourier quaternionique. La corrélation croisée pour les images couleur dans le domaine fréquentiel avec la transformée de Fourier quaternionique s'écrit

$$\begin{aligned} C_{\mu_1}^{\pm R}(\mathbf{u}) &= \mathcal{F}_{\mu_1}^{\pm R} [f(\mathbf{x}) * g(\mathbf{x})] \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(\mathbf{x}) \overline{g(\mathbf{x} - \mathbf{x}_a)} d\mathbf{x} \right\} e^{\mp \mu_1 \langle \mathbf{u}, \mathbf{x}_a \rangle} d\mathbf{x}_a \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \left\{ \int_{-\infty}^{\infty} \overline{g(\mathbf{x} - \mathbf{x}_a)} e^{\mp \mu_1 \langle \mathbf{u}, \mathbf{x}_a \rangle} d\mathbf{x}_a \right\} d\mathbf{x} \\ &\quad \text{en posant } \mathbf{x}' = \mathbf{x} - \mathbf{x}_a \text{ et } \mathbf{x} = \mathbf{x}' + \mathbf{x}_a, \text{ on obtient} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \left\{ \int_{-\infty}^{\infty} \overline{g(\mathbf{x}')} e^{\mp \mu_1 \langle \mathbf{u}, \mathbf{x} - \mathbf{x}' \rangle} d\mathbf{x}' \right\} d\mathbf{x} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \left\{ \int_{-\infty}^{\infty} \overline{g(\mathbf{x}')} e^{\pm \mu_1 \langle \mathbf{u}, \mathbf{x}' \rangle} d\mathbf{x}' \right\} e^{\mp \mu_1 \langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \left\{ \overline{\int_{-\infty}^{\infty} e^{\mp\mu_1 \langle \mathbf{u}, \mathbf{x}' \rangle} g(\mathbf{x}') d\mathbf{x}'} \right\} e^{\mp\mu_1 \langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \left\{ \widehat{g}_{\parallel\mu_1}^{\pm L}(\mathbf{u}) + \widehat{g}_{\perp\mu_1}^{\pm L}(\mathbf{u}) \right\} e^{\mp\mu_1 \langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \widehat{g}_{\parallel\mu_1}^{\pm L}(\mathbf{u}) e^{\mp\mu_1 \langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} + \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}) \widehat{g}_{\perp\mu_1}^{\pm L}(\mathbf{u}) e^{\mp\mu_1 \langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \\
 &= \widehat{f}_{\mu_1}^{\pm R}(\mathbf{u}) \widehat{g}_{\parallel\mu_1}^{\pm L}(\mathbf{u}) + \widehat{f}_{\mu_1}^{\mp R}(\mathbf{u}) \widehat{g}_{\perp\mu_1}^{\pm L}(\mathbf{u})
 \end{aligned} \tag{3.7}$$

où μ_1 est un quaternion pur et unitaire, $\mathcal{F}_{\mu_1}^{\pm R}$ est l'opérateur de transformée de Fourier quaternionique à droite. Dans [Moxey 02], plusieurs formes alternatives de la corrélation sont données en utilisant la transformée de Fourier quaternionique à gauche et à droite. L'étape suivante consiste à normaliser la corrélation dans le domaine fréquentiel calculée ci-dessus pour retrouver un "cross-power spectrum" :

$$R_{\mu_1}^{\pm R}(u) = \frac{C_{\mu_1}^{\pm R}(u)}{|C_{\mu_1}^{\pm R}(u)|} \tag{3.8}$$

Enfin, la corrélation de phase couleur quaternionique est calculée en prenant la transformée de Fourier quaternionique inverse à droite du "cross-power spectrum" :

$$r_{\mu_1}(u) = \mathcal{F}^{\mp R} \{ R_{\mu_1}^{\pm R}(u) \} \tag{3.9}$$

Il en résulte une "image" dont les pixels sont des quaternions. Le score de corrélation ρ_{μ}^R en est alors la valeur maximale des magnitudes de ces quaternions. La figure 3.11 illustre la corré-

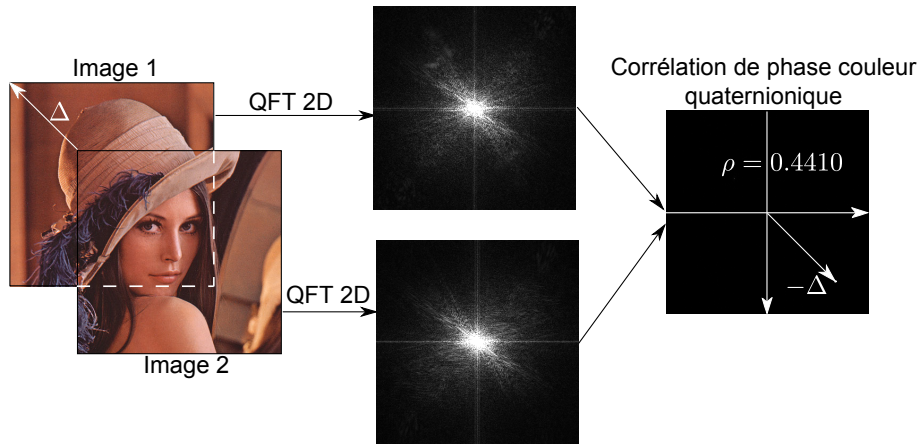


FIGURE 3.11 – La corrélation de phase couleur avec la transformée de Fourier quaternionique

lation de phase quaternionique de Sangwine *et al.*, calculée à l'aide de la bibliothèque *QTFM* [Sangwine 05]. L'image 2 est la version tradlatée par un vecteur Δ de l'image 1. Le score de corrélation obtenu est de 0.4410 et un vecteur de translation égal à $-\Delta$. Nous pouvons alors remarquer que le score obtenu est plus élevé que la corrélation de phase en niveaux de gris (= 0.4145).

De manière surprenante, la corrélation de phase quaternionique calculée avec la librairie *QTFM* [Sangwine 05] donne un résultat inférieur à 1 (0.88) si l'on calcule la corrélation de phase quaternionique entre l'image de Lenna et elle-même. Cela peut être dû à des problèmes numériques pendant le calcul des transformées de Fourier.

3.2.2.2 Proposition d'une corrélation de phase Clifford couleur

Pour définir une corrélation de phase à l'aide de la transformée de Fourier Clifford couleur, rappelons tout d'abord le "Shift-Theorem" de cette transformée que nous avons défini dans le premier chapitre. Pour une image g , égale à une image f translatée, il s'écrit

$$\widehat{g}_B(\mathbf{u}) = e^{\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle B + I_4 B \rangle} \widehat{f}_B(\mathbf{u}) e^{-\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle B + I_4 B \rangle} \quad (3.10)$$

où Δ contient les paramètres de la translation et B un bivecteur unitaire.

Des équations plus simples peuvent être obtenues en utilisant la décomposition de f comme la somme de sa projection parallèle $f_{\parallel B}$ sur le plan défini à partir de B et de sa projection orthogonale $f_{\perp B}$ sur le plan défini par $I_4 B$. Selon cette décomposition, l'équation 3.10 donne deux équations.

Soit $f, g \in L^2(\mathbb{R}^2, \mathbb{R}_{4,0}^1)$ tel que $g(\mathbf{x}) = f(\mathbf{x} + \Delta)$ et B un bivecteur unitaire dans $\mathbb{R}_{4,0}$ et

$$\widehat{g}_{\parallel B}(\mathbf{u}) = e^{\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle B \rangle} \widehat{f}_{\parallel B}(\mathbf{u}) e^{-\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle B \rangle} = \widehat{f}_{\parallel B}(\mathbf{u}) e^{-\langle \mathbf{u}, \Delta \rangle \langle B \rangle} \quad (3.11)$$

$$\widehat{g}_{\perp B}(\mathbf{u}) = e^{\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle I_4 B \rangle} \widehat{f}_{\perp B}(\mathbf{u}) e^{-\frac{\langle \mathbf{u}, \Delta \rangle}{2} \langle I_4 B \rangle} = \widehat{f}_{\perp B}(\mathbf{u}) e^{-\langle \mathbf{u}, \Delta \rangle \langle I_4 B \rangle} \quad (3.12)$$

Pour calculer une corrélation de phase similaire aux équations du "cross-power spectrum" (équation 3.4) et de la corrélation de phase (équation 3.5), on devrait disposer d'une transformée de Fourier Clifford couleur inverse pour des fonctions multivectorielles générales. Malheureusement, celle-ci reste à définir car la transformée de Fourier Clifford couleur n'est définie que pour des fonctions vectorielles.

Une de nos nouvelles approches consiste à utiliser les décompositions de \widehat{f}_B et \widehat{g}_B par rapport à B . En considérant les équations 3.11 et 3.12, la corrélation de phase s'appuie maintenant sur la détection de deux pics de Dirac à une position identique dans les parties parallèle et orthogonale,

$$\begin{aligned} R_{\parallel B}(\mathbf{u}) &= \frac{\widehat{f}_{\parallel B}(\mathbf{u}) \overline{\widehat{g}_{\parallel B}(\mathbf{u})}}{|\widehat{f}_{\parallel B}(\mathbf{u}) \overline{\widehat{g}_{\parallel B}(\mathbf{u})}|} = e^{\langle \mathbf{u}, \Delta \rangle \langle B \rangle} \rightarrow r_{\parallel B}(\mathbf{x}) = \delta_{-\Delta}(\mathbf{x}) \\ R_{\perp B}(\mathbf{u}) &= \frac{\widehat{f}_{\perp B}(\mathbf{u}) \overline{\widehat{g}_{\perp B}(\mathbf{u})}}{|\widehat{f}_{\perp B}(\mathbf{u}) \overline{\widehat{g}_{\perp B}(\mathbf{u})}|} = e^{\langle \mathbf{u}, \Delta \rangle \langle I_4 B \rangle} \rightarrow r_{\perp B}(\mathbf{x}) = \delta_{-\Delta}(\mathbf{x}) \end{aligned} \quad (3.13)$$

Les scores de corrélation $\rho_{\parallel B}$ et $\rho_{\perp B}$ sont extraits de $r_{\parallel B}(\mathbf{x})$ et $r_{\perp B}(\mathbf{x})$ avec

$$\rho_{\parallel B} = \max_{\mathbf{x}} (|r_{\parallel B}(\mathbf{x})|) \text{ et } \rho_{\perp B} = \max_{\mathbf{x}} (|r_{\perp B}(\mathbf{x})|) \quad (3.14)$$

En pratique, le score final est obtenu par agrégation de $\rho_{\parallel B}$ et $\rho_{\perp B}$. De nombreux choix d'agrégation sont possibles (minimum, maximum, etc.). Le processus complet est illustré dans la figure 3.12.

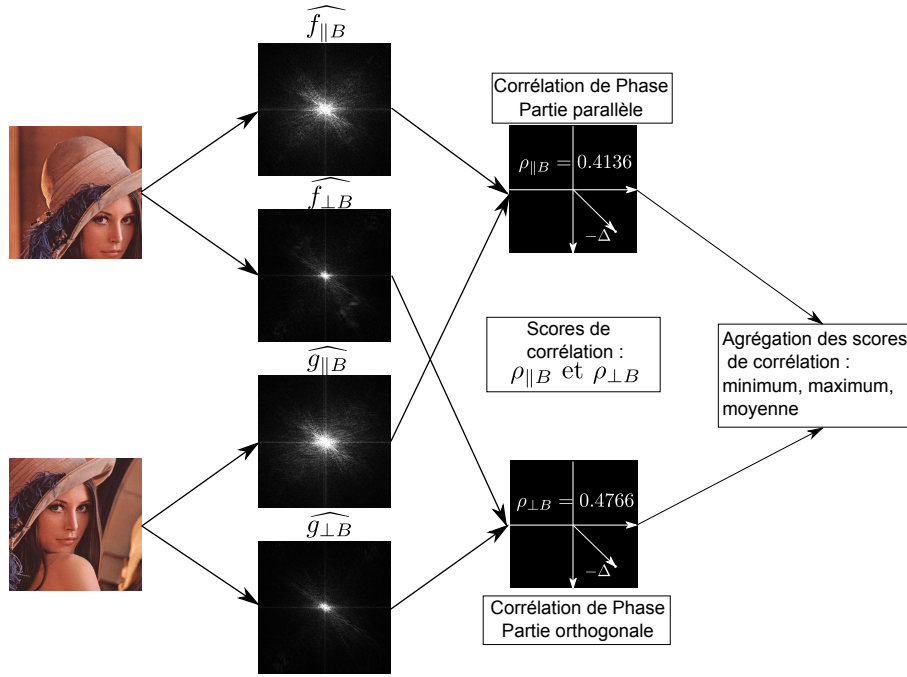


FIGURE 3.12 – L’image de similarité calculée à partir de l’agrégation de deux scores avec un bivecteur B

La figure 3.12 illustre la première corrélation de phase couleur que nous proposons avec la transformée de Fourier Clifford. Les scores de corrélation obtenus sont 0.4136 pour la partie parallèle et 0.4766 pour la partie orthogonale. Les vecteurs de translation extraits de $r_{\parallel B}$ et $r_{\perp B}$ sont égaux à $-\Delta$ dans le cas où les deux images f et g sont les versions translattées l’une de l’autre. Dans le cas où les deux images sont très différentes, les vecteurs de translation peuvent être différents. On peut alors considérer que le vecteur de translation est celui qui est associé au maximum des deux coefficients $\rho_{\parallel B}$ et $\rho_{\perp B}$. Dans nos applications, seuls les scores de corrélations sont considérés.

Une autre méthode que nous proposons consiste à choisir l’unique rotation dans \mathbb{R}^4 qui laisse invariant le plan engendré par les vecteurs $\widehat{f}_B(\mathbf{u})$ et $\widehat{g}_B(\mathbf{u})$. Cette rotation est caractérisée par le spineur $\tau_B(\mathbf{u})$ tel que $g_B^\sharp(\mathbf{u}) = \tau_B(\mathbf{u})f_B^\sharp(\mathbf{u})\tau_B^{-1}(\mathbf{u})$ donné par la formule

$$\tau_B(\mathbf{u}) = \exp \left[\frac{\theta(\mathbf{u})}{2} \frac{g_B^\sharp(\mathbf{u}) \wedge f_B^\sharp(\mathbf{u})}{|g_B^\sharp(\mathbf{u}) \wedge f_B^\sharp(\mathbf{u})|} \right] = \frac{1 + g_B^\sharp(\mathbf{u})f_B^\sharp(\mathbf{u})}{\sqrt{2(1 + g_B^\sharp(\mathbf{u}) \cdot f_B^\sharp(\mathbf{u}))}} \quad (3.15)$$

où $f_B^\sharp(\mathbf{u}) = \widehat{f}_B(\mathbf{u})/|\widehat{f}_B(\mathbf{u})|$ et $g_B^\sharp(\mathbf{u}) = \widehat{g}_B(\mathbf{u})/|\widehat{g}_B(\mathbf{u})|$. Cette rotation est classée dans les rotations simples dans [Lounesto 01]. Cette approche s’écarte des conditions du “Shift-Theorem” en relaxant la contrainte sur le type de rotation entre les vecteurs \widehat{f}_B et \widehat{g}_B . C’est-à-dire que l’on ne se situe plus dans le cas des rotations isoclines. Ici encore, la transformée de Fourier Clifford inverse de $\tau_B(\mathbf{u})$ n’est pas disponible. Cependant, il est possible de négliger la partie bivectorielle $\left(\frac{g_B^\sharp(\mathbf{u}) \wedge f_B^\sharp(\mathbf{u})}{|g_B^\sharp(\mathbf{u}) \wedge f_B^\sharp(\mathbf{u})|} \right)$ du spineur, qui diffère pour chaque \mathbf{u} , en le transformant en un

bivecteur identifiable à un imaginaire complexe i . Dans ce cas, le plan de rotation est négligé et seul l'angle de rotation est considéré.

Un nouveau score de corrélation ρ_B peut alors être construit à partir de $\theta(\mathbf{u})$

$$R_B(\mathbf{u}) = e^{i\theta(\mathbf{u})} \implies \rho_B = \max_{\mathbf{x}} (|r_B(\mathbf{x})|) \quad (3.16)$$

où $r_B(\mathbf{x}) = \check{R}_B(\mathbf{u})$.

Une formulation alternative de ce critère est donnée par le cosinus de l'angle entre \widehat{f}_B et \widehat{g}_B (noté $\angle \widehat{f}_B(\mathbf{u}) \widehat{g}_B(\mathbf{u})$)

$$\Re(R_B(\mathbf{u})) = \cos\left(\angle \widehat{f}_B(\mathbf{u}) \widehat{g}_B(\mathbf{u})\right) = \frac{\widehat{f}_B(\mathbf{u}) * \widetilde{\widehat{g}_B(\mathbf{u})}}{|\widehat{f}_B(\mathbf{u})| |\widehat{g}_B(\mathbf{u})|} \quad (3.17)$$

où \sim et $*$ correspondent à l'opérateur de réversion et au produit scalaire entre multivecteurs [Hestenes 84] et \Re à la partie réelle d'un complexe. Notons qu'avec cette formulation, l'image résultante contient deux pics de corrélation de même amplitude (voir figure 3.13). Le score de corrélation ρ_B est alors égal à

$$\rho_B = \rho_B^1 + \rho_B^2 = 2 \times \max_{\mathbf{x}} (|\Re(\widetilde{R_B(\mathbf{u})})|) \quad (3.18)$$

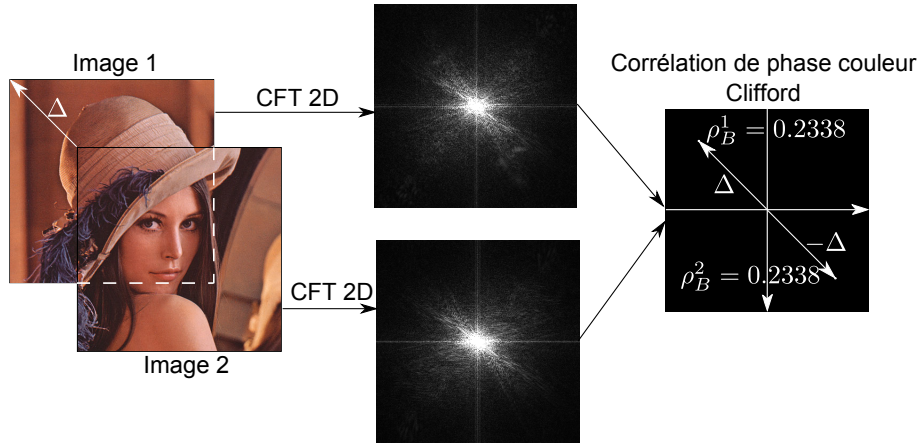


FIGURE 3.13 – L'image de similarité calculée à partir du cosinus entre multivecteurs

La figure 3.13 illustre la seconde corrélation de phase couleur que nous proposons à partir du cosinus entre multivecteurs. La corrélation de phase fournit un score de corrélation de 0.4677 et deux vecteurs de translation égaux à $-\Delta$ et Δ . En effet, la transformée de Fourier inverse d'une fonction cosinus est une image qui contient deux pics de dirac symétriques et de même amplitude. Le score obtenu est supérieur à celui de la corrélation de phase classique et à celui de la corrélation de phase quaternionique.

Nous avons défini dans ce chapitre deux nouvelles méthodes pour calculer une corrélation de phase pour les images couleur. Elles dépendent d'un bivecteur B_c lié à la transformée de Fourier, où c est une couleur choisie. La première dépend de deux scores de corrélation, $\rho_{\parallel B_c}$ et $\rho_{\perp B_c}$,

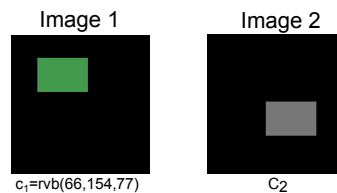
donnés par la partie parallèle et orthogonale de la transformée de Fourier Clifford et requiert une étape d'agrégation pour donner un score final. La seconde, notée ρ_{B_c} , ne requiert pas d'étape d'agrégation. Dans la partie suivante, nous nous intéressons à l'évaluation des corrélations de phase que nous proposons en termes de scores et de taux de reconnaissance. Ces résultats sont comparés avec ceux obtenus avec la corrélation de phase quaternionique et en niveaux de gris.

3.2.3 Comparaison des corrélations de phase quaternionique et Clifford

Trois bases de tests sont utilisées pour comparer les corrélations de phase classique, quaternionique et Clifford. Dans ce dernier cas, le choix du bivecteur est discuté.

3.2.3.1 Expérimentations sur des images synthétiques

Deux images simples contenant un même rectangle sur un fond noir sont considérées. Le second est translaté et sa couleur \mathbf{c}_2 change pour chaque expérimentation. Les bivecteurs B_μ , B_r et $B_{\mathbf{c}_1}$, où $\mathbf{c}_1 = rvb(66, 154, 77)$ est la couleur du premier rectangle, sont utilisés pour calculer les scores de corrélations du tableau de la figure 3.14.



B		\mathbf{c}_2				μ
		\mathbf{c}_1	$rvb(66, 0, 0)$	$rvb(0, 154, 0)$	$rvb(0, 0, 77)$	
B_μ	$\rho_{\parallel B}$	1	1	1	1	1
	$\rho_{\perp B}$	1	1	1	1	-
	ρ_B	1	0.43	0.84	0.50	0.93
B_r	$\rho_{\parallel B}$	1	1	-	-	1
	$\rho_{\perp B}$	1	-	1	1	1
	ρ_B	1	0.36	0.93	0.93	0.96
$B_{\mathbf{c}_1}$	$\rho_{\parallel B}$	1	1	1	1	1
	$\rho_{\perp B}$	-	-	-	-	-
	ρ_B	1	0.36	0.84	0.42	0.93

FIGURE 3.14 – Scores de corrélation entre les images 1 et 2 pour des choix différents de \mathbf{c}_2 et B . Le symbole '-' décrit une valeur qui ne peut pas être calculée parce que la partie parallèle ou orthogonale est nulle.

Dans le cas $\mathbf{c}_2 = rvb(0, 0, 77)$, remarquons que $\rho_{\parallel B_\mu} = \rho_{\perp B_\mu} = 1$. Cela signifie donc que les rectangles de couleur \mathbf{c}_1 et \mathbf{c}_2 sont égaux. Or, ils ne sont pas de même couleur. Ce résultat n'est donc pas correct car deux rectangles qui ont des couleurs différentes doivent être considérés comme différents. Au contraire, les scores de corrélation ρ_{B_c} sont inférieurs à 1 et dépendent de la quantité de couleur que les deux rectangles ont en commun. La même remarque peut être faite en prenant \mathbf{c}_2 égal à $rvb(0, 154, 0)$ ou $rvb(0, 0, 77)$. Le comportement pour $\mathbf{c}_2 = \mu$ est différent car les ρ_{B_c} sont toujours élevés quelque soit le bivecteur choisi. Ce résultat n'est pas surprenant car ils dépendent fortement de la forme sans prendre en compte l'information couleur.

La même expérimentation en utilisant la corrélation de phase quaternionique donne des scores toujours égaux à 1 quelle que soit la couleur des rectangles. Cela montre une insensibilité de cette corrélation aux changements de couleur dans le cas d'images simples et monochromatiques.

Cette première expérimentation montre clairement que notre approche, avec le score ρ_{B_c} , discrimine les rectangles qui ont une couleur différente. Dans un contexte plus général de reconnaissance de formes couleur, nous pouvons penser cette méthode permettra de mieux séparer les objets de même couleur mais de forme différente, contrairement aux autres méthodes (en niveaux de gris, quaternionique, ou en considérant les scores $\rho_{\parallel B}$ et $\rho_{\perp B}$).

3.2.3.2 Utilisation dans le cadre de la reconnaissance d'images

La corrélation de phase peut être assimilée à une mesure de similarité entre images. À ce titre, elle n'est pas assimilable à un vecteur de caractéristiques extraits d'une image. L'utilisation directe d'un classifieur SVM ne peut donc pas être faite comme dans la partie dédiée aux descripteurs.

Une première façon de contourner ce problème est de considérer le score de corrélation dans un classifieur de type 1-plus similaire [Duda 01]. Dans ce cas, l'algorithme SVM n'est pas utilisé. Une seconde approche est proposée dans [Chen 09]. Les auteurs proposent de remplacer l'évaluation du noyau par une fonction de similarité. Ainsi, la distance entre les vecteurs de caractéristiques dans l'espace de représentation est remplacée par la valeur de similarité. Dans ce cadre, la matrice de similarité S (matrice de Gram) est utilisée pour cet algorithme :

$$S_{i,j} = Sim(I_i, I_j) \quad (3.19)$$

où I_i et I_j sont deux images d'indices i et j . La matrice de similarité devrait être symétrique et définie positive pour être utilisée dans un algorithme SVM.

La propriété de symétrie pour la matrice contenant les corrélations de phase est évidente dans le cas où un seul bivecteur est choisi pour calculer toutes les corrélations de phase. En effet, le calcul de corrélation de phase est symétrique. Par contre, dans le cas où l'on choisit un bivecteur par image requête, la matrice n'est plus symétrique. Une matrice de similarité S peut facilement être symétrisée :

$$S_{sym} = \frac{1}{2}(S + S^t) \quad (3.20)$$

où S^t correspond à la matrice S transposée.

La propriété de matrice définie positive est moins évidente. Pour qu'une matrice symétrique réelle soit définie positive, toutes ses valeurs propres doivent être strictement positives. En pratique, on observe quelques valeurs propres négatives concernant les matrices de Gram.

Quand certaines valeurs propres ne sont pas positives, on parle alors de matrice de similarité indéfinie. Il est tout à fait possible d'ignorer le fait que la matrice de similarité soit indéfinie [Chen 09]. Cette méthode a l'avantage d'être simple mais elle a pour inconvénient de ne pas assurer la minimisation du risque fonctionnel car elle peut amener à des solutions correspondant à des points scelle.

D'autres auteurs proposent des méthodes pour forcer la positivité des valeurs propres. Elles s'appuient pour la plupart sur les valeurs propres de la matrice d'origine.

Une matrice de similarité S symétrique admet une décomposition en valeurs propres qui s'écrit

$$S = U^t \Lambda U \tag{3.21}$$

où U est une matrice orthogonale et Λ est une matrice dont la diagonale contient les valeurs propres : $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ avec n le nombre d'images. La matrice S est donc énorme (de taille $n \times n$), ce qui implique un coût calculatoire très élevé.

Une autre méthode revient à ramener les valeurs propres négatives de la matrice de similarité à 0 en faisant l'hypothèse que les valeurs propres négatives correspondent à du bruit [Wu 05].

Une troisième méthode consiste à changer le signe des valeurs propres négatives en prenant la valeur absolue des valeurs propres.

Il est aussi possible de ramener la plus petite valeur propre à 0 en additionnant la valeur absolue de celle-ci à toutes les valeurs propres.

Enfin, la matrice de similarité SS^t peut être considérée comme une matrice de Gram. En effet, cela revient à passer au carré ses valeurs propres.

Ces solutions sont détaillées dans [Chen 09].

Dans la partie suivante, les résultats obtenus en utilisant directement la matrice de similarité comme noyau et la matrice de similarité SS^t sont présentés. Cette dernière a été choisie car elle ne nécessite pas de calculer explicitement les valeurs propres des matrices de similarité et a un coût faible.

3.2.3.3 Expérimentations sur la base COIL-100

Les taux de reconnaissance obtenus avec les bivecteurs B_r , B_v , B_b et B_{c_i} sont donnés dans le tableau 3.7. La couleur c_i est la couleur dominante de chaque image et correspond au bivecteur choisi pour chaque image requête.

Ce tableau montre clairement que les scores de corrélations ρ_B calculés à partir de la transformée de Fourier Clifford couleur donnent les meilleurs résultats pour cette base de donnée. Cela nous confirme qu'elle est la plus performante pour les objets couleur. De plus, le critère ρ_B semble insensible au choix du bivecteur. On peut aussi voir que la plupart des méthodes de corrélation de phase couleur sont plus performantes que la corrélation de phase classique pour les images en niveaux de gris. Les différents choix de bivecteurs n'améliorent pas significativement les résultats, néanmoins, les taux de reconnaissance sont très bons (même s'ils sont inférieurs aux *GCFD*).

ρ	ρ_{μ}^R		$\rho_{\parallel B}$	$\rho_{\perp B}$	$\rho_{\parallel, \perp B}^{min.}$	$\rho_{\parallel, \perp B}^{max.}$	$\rho_{\parallel, \perp B}^{moy.}$	ρ_B
94.96	96.85	B_r	95.38	96.33	96.46	96.14	96.58	97.50
		B_v	95.29	96.79	96.68	95.90	96.68	97.49
		B_b	95.08	96.58	96.60	95.89	96.51	97.49
		B_{c_i}	95.33	83.92	82.60	96.08	95.50	97.53

TABLE 3.7 – COIL-100 : Taux de reconnaissance en % avec la corrélation de phase pour les images couleur et la matrice de Gram S

ρ	ρ_{μ}^R		$\rho_{\parallel B}$	$\rho_{\perp B}$	$\rho_{\parallel, \perp B}^{min.}$	$\rho_{\parallel, \perp B}^{max.}$	$\rho_{\parallel, \perp B}^{moy.}$	ρ_B
94.53	96.19	B_r	95.19	95.68	95,58	95.68	95.68	97.24
		B_v	95.18	96.26	95.92	95.65	95.86	97.22
		B_b	94.46	96.21	95.79	95.56	95.72	97.18
		B_{c_i}	95.61	94.46	94.40	96.24	95.51	97.30

TABLE 3.8 – COIL-100 : Taux de reconnaissance en % avec la corrélation de phase pour les images couleur et la matrice de Gram SS^t

La table 3.8 contient les taux de reconnaissance en considérant le noyau SS^t . Des résultats similaires sont obtenus malgré une légère détérioration des taux de reconnaissance.

3.2.3.4 Expérimentations sur la base color FERET

Le critère ρ_B n'est plus le meilleur mais reste très stable. Cette relativement faible performance est due aux confusions induites par les différentes couleurs qui constituent l'arrière plan. Nous nous sommes alors intéressés à une autre stratégie, nous choisissons un bivecteur B_{c_i} par image requête. Il correspond à la couleur dominante c_i et sépare l'arrière plan du premier plan. Ce choix donne le meilleur taux de reconnaissance mais nécessite de recalculer toutes les transformées de Fourier des images d'apprentissage pour chaque image requête avec le bivecteur B_{c_i} .

Cette stratégie donne le meilleur résultat parmi toutes les méthodes de corrélation phase. Le taux de reconnaissance (78.38%) est même supérieur à ceux obtenus avec les $GFD1$ et $GFD2$ (voir tableau 3.3 de la partie 3.1.4.2).

ρ	ρ_{μ}^R		$\rho_{\parallel B}$	$\rho_{\perp B}$	$\rho_{\parallel, \perp B}^{min.}$	$\rho_{\parallel, \perp B}^{max.}$	$\rho_{\parallel, \perp B}^{moy.}$	ρ_B
65.61	68.98	B_r	66.51	66.00	67.64	68.05	69.15	66.74
		B_v	66.00	67.91	67.31	66.95	67.51	66.41
		B_b	65.57	67.41	68.01	66.95	67.71	66.81
		B_{c_i}	66.34	74.50	76.84	71.52	78.38	66.94

TABLE 3.9 – color FERET : Taux de reconnaissance en % avec la corrélation de phase pour les images couleur et la matrice de Gram S

La table 3.10 contient les taux de reconnaissance de la base color FERET en considérant le noyau SS^t . La plupart des taux de reconnaissance ont augmenté de 1%.

ρ	ρ_{μ}^R		$\rho_{\parallel B}$	$\rho_{\perp B}$	$\rho_{\parallel, \perp B}^{min.}$	$\rho_{\parallel, \perp B}^{max.}$	$\rho_{\parallel, \perp B}^{moy.}$	ρ_B
66.84	67.44	B_r	67.51	67.25	68.28	68.82	69.25	67.95
		B_v	67.14	68.58	67.98	67.78	67.75	67.51
		B_b	66.84	68.12	68.25	67.91	68.28	67.68
		B_{c_i}	67.11	76.74	78.38	73.50	78.07	68.15

TABLE 3.10 – color FERET : Taux de reconnaissance en % avec la corrélation de phase pour les images couleur et la matrice de Gram SS^t

3.2.4 Discussion

Dans cette partie, la corrélation de phase classique et la corrélation de phase couleur quaternionique ont été rappelées. Nous avons aussi défini deux nouvelles corrélations de phase couleur à partir de la transformée de Fourier Clifford couleur. Les corrélations de phase couleur (quaternionique et Clifford) donnent clairement de meilleurs résultats que la corrélation de phase classique (pour les images en niveaux de gris).

Les constructions de la corrélation de phase quaternionique et à partir de la transformée de Fourier clifford sont relativement différentes. La corrélation de phase quaternionique est calculée en prenant la transformée de Fourier quaternionique de la corrélation. Ceci est possible car le produit de deux quaternions est un quaternion. Dans le cas de la corrélation de phase Clifford, c'est un spineur qui devrait être extrait par le "cross-power spectrum", soit l'exponentielle d'un bivecteur. Malheureusement, cette quantité n'est pas de nature vectorielle et il n'existe pas encore de définition de la transformée de Fourier Clifford couleur pour les multivecteurs quelconques. Il n'est donc pas possible actuellement de calculer directement la corrélation de phase couleur à l'aide de la transformée de Fourier Clifford couleur. La première de nos contributions consiste à traiter séparément les parties parallèle et orthogonale de la transformée de Fourier Clifford couleur. La seconde est basée sur la nature de la corrélation de phase et plus précisément celle du "cross-power spectrum". Celui-ci revient à récupérer l'angle entre les deux vecteurs de \mathbb{R}^4 , \hat{f} et \hat{g} , correspondants aux valeurs de la transformée de Fourier Clifford couleur de deux images. Il existe plusieurs types de rotations entre deux vecteurs de \mathbb{R}^4 . Nous nous sommes restreints aux rotations simples (voir partie 1.2.3.2) qui agissent sur un seul plan.

Pour les expérimentations sur la base COIL-100, les scores de corrélation ρ_B donnent les meilleurs résultats. Par contre, pour la base color FERET, les corrélations de phase quaternioniques ρ_μ^R et Clifford reconstruites ρ_B ne sont plus les meilleures méthodes. En effet, la couleur du fond de l'image change suivant les images. En choisissant un bivecteur pour chaque image de test (séparant le fond de la forme) et en considérant séparément les parties parallèles et orthogonales, les meilleurs taux de reconnaissance sont obtenus.

Nous pouvons en conclure que les deux corrélations que nous proposons ont deux applications différentes :

- la corrélation de phase ρ_B est la meilleure quand seul l'objet d'intérêt est présent dans l'image. Cette méthode est donc préférable pour la base COIL-100.
- l'intérêt de considérer un bivecteur B_{c_i} par image requête est de séparer l'objet d'intérêt du fond dans les images. Cette méthode est donc mieux adaptée à la base color FERET.

3.3 Conclusion

Dans la première partie de ce chapitre, les descripteurs *GCFD* ont été comparés par rapport à la méthode marginale et aux moments de Fourier-Mellin quaternioniques. Sur trois bases de données différentes, en termes de taux de reconnaissance, les *GCFD* apparaissent comme les descripteurs les plus pertinents. En effet, dans le cadre de l'expérimentation sur la base *colorFERET*, la construction de manière non *ad hoc* des descripteurs *GCFD* permet l'obtention de bons taux de reconnaissance contrairement aux moments de Fourier-Mellin quaternioniques. De plus, un

gain en taille de descripteurs et en complexité sont observés.

Par cette série d'expérimentations, nous avons montré l'apport des transformées de Fourier Clifford couleur et quaternionique pour la reconnaissance de formes. Toutefois, dans la transformée de Fourier Clifford couleur, la direction d'analyse reste un paramètre primordial. À l'aide d'expérimentations sur le choix du bivecteur pour les *GCFD1*, nous avons montré que certains choix de bivecteur étaient plus pertinents que d'autres. Par exemple, dans le cas d'images naturelles ou d'objets manufacturés, le choix d'une direction d'analyse de teinte rouge ou bleu est un bon choix.

Afin de montrer la pertinence de l'utilisation de la transformée de Fourier Clifford couleur pour la reconnaissance de formes, une autre méthode fréquentielle est détaillée : la corrélation de phase. Pour étendre cette méthode à la couleur, nous avons défini le "Shift-Theorem" de la transformée de Fourier Clifford couleur. Deux approches différentes par leurs constructions ont été implémentées. La première se base sur la fusion de deux scores de corrélations issus des parties parallèle et orthogonale de la transformée de Fourier Clifford couleur. La seconde obtient le score de corrélation directement sur la transformée de Fourier Clifford reconstruite.

Une fois encore, les expérimentations montrent une amélioration des taux de reconnaissance avec les transformées de Fourier couleur.

Conclusion générale et perspectives

Les travaux décrits dans cette thèse portent sur la définition et la mise en œuvre de descripteurs de Fourier et de corrélations de phase couleur dans le cadre original des algèbres de Clifford. Nos descripteurs reposent sur deux développements récents de méthodes, pourtant classiques, qui sont d'une part la transformée de Fourier étendue à la couleur [Batard 10] et d'autre part les descripteurs généralisés de Smach *et al.* [Smach 08]. Ces deux avancées ont comme socle commun la théorie de représentation des groupes [Vilenkin 68]. Nos contributions tirent le meilleur parti de l'information colorimétrique grâce à la modélisation de la géométrie de l'espace couleur. Cette approche, qui élimine la nécessité d'un traitement *ad hoc*, met en avant la notion fondamentale de direction d'analyse qui est commune aux méthodes fréquentielles reposant sur les algèbres de Clifford et aux méthodes quaternioniques.

Nous avons, dans le premier chapitre, fait le lien entre les deux approches en réécrivant la transformée quaternionique à droite de Sangwine et Ell [Sangwine 00] dans le formalisme de la transformée Clifford de Batard *et al.* [Batard 10]. Ces deux approches sont elles-mêmes replacées dans un état de l'art original des transformées de Fourier. Le panorama que nous dressons débute par les séries de Fourier et s'achève par les transformées pour les images couleur en passant par celles pour les images en niveaux de gris. L'accent est mis sur leurs propriétés au regard des transformations usuelles des images (translation, rotation et homothétie). Notons que ces propriétés sont données par construction dans le cas des transformées de Fourier Clifford et M_2 qui s'appuient sur la transformée de Fourier abstraite. En termes d'action de groupe, la transformée de Fourier classique correspond à l'action du groupe des translations $(\mathbb{R}^2, +)$ sur \mathbb{R}^2 au travers de l'action du groupe $Spin(2)$ sur $\mathbb{R}_{2,0}^1$. De même, la transformée de Fourier Clifford et la M_2 -transformée sont définies à l'aide d'actions de groupes.

La complexité mathématique requise pour la définition et la compréhension de ces transformées ne se traduit pas par une plus grande complexité algorithmique puisqu'elles sont implémentables efficacement à partir de la classique transformée de Fourier rapide. En particulier, nous donnons un algorithme constant en espace et quasi-linéaire en temps pour la transformée de Fourier Clifford couleur. Ce résultat est obtenu en plongeant l'image dans $\mathbb{R}_{4,0}$ et en utilisant une partie de $Spin(4)$, lequel est isomorphe à $Spin(3) \times Spin(3)$: la transformée est ainsi décomposée en deux parties, dites parallèles et orthogonales, par projection sur deux plans orthogonaux, sur lesquels des rotations indépendantes agissent.

Sur la base de ces transformées, divers descripteurs ont été développés pour les images en niveaux de gris [Smach 08, Derrode 01]. L'étude que nous en avons faite dans le deuxième chapitre nous a conduit à privilégier les descripteurs généralisés de Smach *et al.* [Smach 08] qui sont invariants en translation et rotation sans le recours à une méthode *ad hoc*. Associés à la transformée de Fourier Clifford couleur, ils nous ont permis de définir trois nouveaux descripteurs couleur ($GCFD1$, $GCFD2$ et $GCFD3$). Les deux premiers correspondent à l'application des descripteurs généralisés sur les parties parallèles et orthogonales de la transformée de Fourier Clifford couleur. Les descripteurs $GCFD3$ réutilisent le principe des $GCFD1$ en appliquant les

descripteurs généralisés directement à la transformée de Fourier Clifford reconstruite dans $\mathbb{R}_{4,0}$. L'intérêt de cette approche est de pouvoir calculer les deux parties de la transformée par deux transformées rapides et d'obtenir un descripteur indépendant du choix du bivecteur d'analyse, qui de plus est remarquablement compact. Les mérites relatifs des descripteurs sont résumés à la fin du chapitre dans le tableau 2.1. Soulignons la compacité des *GCFD* relativement aux autres approches couleur, qu'elles soient marginales ou quaternioniques. La raison majeure en est l'absence de redondance entre les parties parallèle et orthogonale.

Dans le chapitre final, l'efficacité en termes de taux de bonne classification de l'approche marginale et des descripteurs de Fourier-Mellin quaternioniques est comparée à celle de nos descripteurs couleur. Sur les trois bases de données retenues, les *GCFD* obtiennent les meilleurs résultats. Un autre point que nous avons étudié est celui du choix de la direction d'analyse, et nous avons montré que certains choix de bivecteur étaient plus pertinents que d'autres. Ainsi, dans le cas d'images naturelles ou d'objets manufacturés, une direction d'analyse de teinte rouge ou bleu est un bon choix, qui donne des taux de classification proches de l'optimum. Le bivecteur correspondant à une de ces deux couleurs peut donc être prise comme paramètre par défaut des descripteurs. Remarquons qu'un meilleur choix peut être fait si on dispose *a priori* d'informations sur les images à traiter : c'est typiquement le cas si le fond est de teinte connue.

La corrélation de phase est la dernière méthode fréquentielle proposée. Elle est naturellement fondée sur la transformée de Fourier Clifford couleur et utilise le "Shift-Theorem" qui lui est associé et que nous avons démontré. Deux approches différentes ont été implémentées à l'aide d'algorithmes efficaces. La première se base sur la fusion de deux scores de corrélations issus des parties parallèle et orthogonale de la transformée de Fourier Clifford couleur. La seconde obtient le score de corrélation directement sur la transformée de Fourier Clifford reconstruite. Cette fois encore, les expérimentations menées montrent les performances de l'approche.

Nos travaux reposent sur des modèles mathématiques génériques pour l'étude des images couleur, et plus généralement multidimensionnelles. Ainsi, la transformée de Fourier de Batard *et al.* [Batard 10] est définie pour des images à quatre canaux. Nous n'utilisons que la version couleur de cette transformée qui en est la restriction à trois canaux. Une suite logique, et assez immédiate, sera d'écrire une implémentation rapide de cette transformée générale puis de reprendre la méthodologie utilisée précédemment pour obtenir des descripteurs, et une corrélation de phase, pour des images à quatre canaux. Des images couleurs enrichies d'un canal infrarouge en sont des exemples évidents, particulièrement pertinents pour des applications de suivi d'objets dans des séquences vidéo. La corrélation de phase ou les descripteurs sont, dans ce contexte, directement utilisables pour définir la vraisemblance utilisée dans un filtre particulière [Isard 98].

D'autres applications de la transformée de Fourier Clifford à la reconnaissance de formes sont envisageables sans développements théoriques supplémentaires. Par exemple, le même schéma que celui utilisé pour les *GCFD* nous permettra de construire des descripteurs de Fourier-Mellin Clifford couleur. Nous y gagnerons l'invariance en échelle mais avec les difficultés d'implémentation de la transformée de Fourier-Mellin (voir le chapitre 1.1.2).

Pour conclure les perspectives, et en quittant le cadre de la reconnaissance des formes pour celui du traitement d'images, la dernière application de la transformée de Fourier Clifford que

nous proposons dans ce document est la segmentation d'images couleur que nous développerons sur la base de la congruence de phase de Kovési [Kovési 03]. En effet, nous sommes persuadés, au regard des images reconstruites uniquement avec la phase couleur (voir la figure 1.12) que nous avons obtenues au chapitre 1.2, de la pertinence de la phase couleur pour la détection de contours et de coins.

Les quaternions

Dans cette annexe, nous explorons la définition ainsi que les principales propriétés des quaternions pour apporter des compléments aux notions abordées dans le manuscrit.

Définition

Les quaternions ont été définis par Sir William Hamilton en 1843 dans l'objectif était de généraliser les nombres complexes dans l'espace à trois dimensions. La difficulté était alors de définir une algèbre avec des triplets de nombres réels et, plus particulièrement, la multiplication entre ces triplets. La solution consistait alors à utiliser une dimension supplémentaire, soit des quadruplets de nombres réels. Ainsi, l'ensemble des quaternions a été créé et il est noté \mathbb{H} en référence à leur créateur.

L'espace vectoriel \mathbb{H} des quaternions est un espace vectoriel de dimension 4 sur le corps des réels ayant pour base le quadruplet $(1, i, j, k)$. Ainsi, tout quaternion q s'écrit de manière unique

$$q = a + ib + jc + kd \quad (\text{A.1})$$

avec $a, b, c, d \in \mathbb{R}$.

Le sous-espace vectoriel engendré par l'élément 1 est identifiable à \mathbb{R} . On dénote ainsi $\mathcal{S}(q)$ la partie scalaire (ou réelle) de q soit $\mathcal{S}(q) = a$. Dans le cas où $\mathcal{S}(q) = q$ alors q est dit réel.

Le sous-espace vectoriel engendré par les vecteurs (i, j, k) est identifiable à \mathbb{R}^3 . On appelle $\mathcal{V}(q)$ la partie vectorielle (ou imaginaire) d'un quaternion q soit $\mathcal{V}(q) = ib + jc + kd$. Dans le cas où $\mathcal{V}(q) = q$ alors q est dit pur. Le sous-espace vectoriel des quaternions purs est noté \mathbb{H}_0 .

Pour un quaternion quelconque $q = a + ib + jc + kd$, on définit

- le conjugué \bar{q} de q : $\bar{q} = a - ib - jc - kd$;
- l'inverse q^{-1} de q : $q^{-1} = \frac{\bar{q}}{|q|^2}$ si $q \neq 0$;
- la norme (ou module) de q est définie par $|q| = \sqrt{a^2 + b^2 + c^2 + d^2} = \sqrt{q\bar{q}}$;
- si $|q| = 1$, alors le quaternion q est qualifié d'unitaire. L'ensemble des quaternions unitaires est noté \mathbb{H}_1 . Si $q \in \mathbb{H}_1$, alors $q^{-1} = \bar{q}$.

Nous pouvons aussi remarquer qu'une sous-algèbre de $\mathbb{R}_{3,0}$ est isomorphe aux quaternions (voir l'annexe B) en réalisant l'identification suivante

$$1 \leftrightarrow 1 \quad e_1 e_2 \leftrightarrow i \quad e_2 e_3 \leftrightarrow j \quad e_1 e_3 \leftrightarrow k \quad (\text{A.2})$$

. Un quaternion peut alors être vu comme la somme d'un scalaire et d'un bivecteur.

La multiplication de Hamilton entre quaternions

La multiplication entre deux quaternions est une opération interne, associative et distributive par rapport à l'addition, qui vérifie des règles spécifiques sur les éléments de la base :

$$\begin{aligned}
 ii &= jj = kk = -1 \\
 ij &= -ji = k \\
 jk &= -kj = i \\
 ki &= -ik = j
 \end{aligned} \tag{A.3}$$

On déduit de ces formules $ijk = jki = kji = -1$. Ces équations nous montrent clairement que la multiplication de deux quaternions n'est pas commutative. Plus précisément, pour deux quaternions $q_1 = a_1 + b_1i + c_1j + d_1k = \mathcal{S}(q_1) + \mathcal{V}(q_1)$ et $q_2 = a_2 + b_2i + c_2j + d_2k = \mathcal{S}(q_2) + \mathcal{V}(q_2)$:

$$\begin{aligned}
 q_1q_2 &= a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)i + \\
 &\quad (a_1c_2 + c_1a_2 + d_1b_2 - b_1d_2)j + (a_1d_2 + d_1a_2 + b_1c_2 - c_1b_2)k \\
 &= (\mathcal{S}(q_1)\mathcal{S}(q_2) - (\mathcal{V}(q_1) \cdot \mathcal{V}(q_2))) + (\mathcal{S}(q_1)\mathcal{V}(q_2) + \mathcal{S}(q_2)\mathcal{V}(q_1) + \mathcal{V}(q_1) \times \mathcal{V}(q_2))
 \end{aligned} \tag{A.4}$$

où \cdot représente le produit scalaire et \times le produit vectoriel (“cross-product”) entre vecteurs.

Soit $q_1 = \mathcal{V}(q_1)$ et $q_2 = \mathcal{V}(q_2)$ deux quaternions purs et orthogonaux entre eux, le produit entre ces quaternions est égal à $\mathcal{V}(q_1') \times \mathcal{V}(q_2')$. Le produit vectoriel étant anti-commutatif, la multiplication de deux quaternions purs et orthogonaux entre eux l'est également ($q_1q_2 = -q_2q_1$).

De même, pour deux quaternions $q_1 = \mathcal{V}(q_1)$ et $q_2 = \mathcal{V}(q_2)$ purs et $q_1 \times q_2 = 0$ (leurs parties vectorielles sont parallèles), le produit de ces quaternions est réduit au produit scalaire $-(\mathcal{V}(q_1) \cdot \mathcal{V}(q_2))$, qui est commutatif. Le produit de ces deux quaternions est donc commutatif ($q_1q_2 = q_2q_1$).

Représentations polaires

La formule d'Euler pour les exponentielles complexes peut être généralisée pour les nombres hypercomplexes :

Soit $q \in \mathbb{H}$,

$$q = |q|e^{\mu\theta} = |q|(\cos(\theta) + \mu\sin(\theta)) \tag{A.5}$$

avec μ un quaternion unitaire et pur appelé l'axe et $0 \leq \theta \leq \pi$ appelé l'angle.

Les calculs de l'axe et de l'angle sont donnés par

$$\begin{aligned}
 \mu &= \frac{\mathcal{V}(q)}{|\mathcal{V}(q)|} \\
 \theta &= \operatorname{atan} \left(\frac{|\mathcal{V}(q)|}{|\mathcal{S}(q)|} \right)
 \end{aligned} \tag{A.6}$$

Cette formulation est utilisée dans [Eli 07] pour visualiser la transformée de Fourier quaternionique couleur.

On retrouve dans [Bülow 99] la définition d'une autre forme polaire qui comporte un module $|q|$ et trois phases ϕ , ψ et θ .

$$q = |q|e^{i\phi}e^{k\psi}e^{j\theta} \quad (\text{A.7})$$

avec $\phi \in [-\pi, \pi]$, $\psi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ et $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$. Le triplet (ϕ, θ, ψ) est appelé la phase angulaire de q .

Cette forme polaire est utilisée dans [Bülow 99] pour introduire la phase quaternionique locale d'un signal.

La décomposition symplectique

La forme de Cayley-Dickson [Dickson 19] d'un quaternion permet de définir récursivement un quaternion comme un nombre complexe généralisé, dont les parties réelle et imaginaire sont elles-mêmes des nombres complexes. La forme de Cayley-Dickson d'un quaternion $q = a + bi + cj + dk$ est

$$q = A + Bj \quad (\text{A.8})$$

où $A = a + bi$ et $B = c + di$.

L'idée que l'on retrouve dans [Ell 07] est de généraliser cette forme en prenant deux quaternions μ_1 et μ_2 purs, unitaires et orthogonaux entre eux pour représenter un quaternion. On parle alors de forme symplectique :

$$q = A' + B'\mu_2 \quad (\text{A.9})$$

où $A' = a' + b'\mu_1$ et $B' = c' + d'\mu_1$, *i.e.*

$$q = (a' + b'\mu_1) + (c' + d'\mu_1)\mu_2 \quad (\text{A.10})$$

La partie A' est appelée la partie simplexe et B' la partie perplexe. Chacune de ces parties peut être identifiée à un nombre complexe. En développant cette expression, on obtient

$$q = [a' + b'\mu_1] + [c'\mu_2 + d'\mu_1\mu_2] \quad (\text{A.11})$$

où $\mu_1\mu_2$ est orthogonal à μ_1 et μ_2 . Une image quaternionique peut être représentée selon cette décomposition. Elle est donnée pour une image $f \in \mathbb{H}_0$ par

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})\mu_2 \quad (\text{A.12})$$

où $f_1(\mathbf{x})$ (respectivement $f_2(\mathbf{x})\mu_2$) correspond à une projection de l'image $f(\mathbf{x})$ dans le plan $(1, \mu_1)$ (respectivement $(\mu_2, \mu_1\mu_2)$).

Quaternions et rotations

Les rotations en 3 dimensions

Soit un quaternion q unitaire $\in \mathbb{H}_1$, alors on peut écrire q sous une forme exponentielle hypercomplexe :

$$q = a + ib + jc + kd = e^{\mu\frac{\alpha}{2}} = \cos\left(\frac{\alpha}{2}\right) + \mu\sin\left(\frac{\alpha}{2}\right) \quad (\text{A.13})$$

avec μ un quaternion unitaire et pur.

Soit v un quaternion pur (par exemple un vecteur $3D$) et s un quaternion unitaire, alors l'équation

$$v' = sv s^{-1} \quad (\text{A.14})$$

correspond à la rotation du quaternion v par un angle α autour de l'axe μ . La rotation se fait dans le sens horaire dans le plan dont la normale est le vecteur μ . Cette opération est aussi appelée la conjugaison par s .

En utilisant la propriété d'anti-involution de la conjugaison des quaternions, la composition de rotations générées par deux quaternions unitaires s_1 et s_2 s'écrit

$$s_1(s_2 v s_2^{-1})s_1^{-1} = s_1(s_2 v \overline{s_2})\overline{s_1} = s_1 s_2 v (\overline{s_1 s_2}) = s_1 s_2 v (s_1 s_2)^{-1} \quad (\text{A.15})$$

La rotation inverse s'écrit avec la conjugaison par s_1^{-1}

$$s_1^{-1}(s_1 v s_1^{-1})s_1 = v \quad (\text{A.16})$$

La multiplication de quaternions n'est pas commutative, les rotations générées par les quaternions ne commutent pas non plus :

Soit deux quaternions unitaires $s_1 = \cos(\alpha_1) + \mu_1 \sin(\alpha_1)$ et $s_2 = \cos(\alpha_2) + \mu_2 \sin(\alpha_2)$,

$$\begin{aligned} s_1 s_2 &= (\cos(\alpha_1) + \mu_1 \sin(\alpha_1)) (\cos(\alpha_2) + \mu_2 \sin(\alpha_2)) \\ &= (\cos(\alpha_1)\cos(\alpha_2) + \mu_1 \sin(\alpha_1)\cos(\alpha_2) + \mu_2 \sin(\alpha_2)\cos(\alpha_1) + \mu_1 \mu_2 \sin(\alpha_1)\sin(\alpha_2)) \\ &\neq (\cos(\alpha_1)\cos(\alpha_2) + \mu_1 \sin(\alpha_1)\cos(\alpha_2) + \mu_2 \sin(\alpha_2)\cos(\alpha_1) + \mu_2 \mu_1 \sin(\alpha_1)\sin(\alpha_2)) \\ &\neq s_2 s_1 \end{aligned} \quad (\text{A.17})$$

Les rotations en 4 dimensions

Une complète description des rotations dans l'espace \mathbb{R}^4 est donnée dans [Lounesto 01]. Les rotations sont de trois types :

- Les rotations simples qui tournent uniquement dans un plan.
- Les rotations isoclines qui tournent dans deux plans orthogonaux d'un même angle. Une rotation isocline a une infinité de plans de rotation.
- Les rotations générales ont deux plans invariants qui sont complètement orthogonaux avec différents angles de rotations.

Selon la décomposition symplectique, un quaternion q peut être décomposé tel que

$$q = (q_1 + q_2 \mu_1) + (q_3 + q_4 \mu_1) \mu_2 \quad (\text{A.18})$$

Cette décomposition nous permettra de décrire les rotations de \mathbb{R}^4 qui peuvent être appliquées aux quaternions.

Les rotations isoclines

Elles sont représentées par la multiplication à gauche d'un quaternion unitaire s et d'un quaternion quelconque q , *i.e.*,

$$\begin{aligned} q' = sq &= \left(\cos\left(\frac{\alpha}{2}\right) + \mu_1 \sin\left(\frac{\alpha}{2}\right) \right) ((q_1 + \mu_1 q_2) + (q_3 + \mu_1 q_4) \mu_2) \\ &= \left(\cos\left(\frac{\alpha}{2}\right) + \mu_1 \sin\left(\frac{\alpha}{2}\right) \right) (q_1 + \mu_1 q_2) + \left(\cos\left(\frac{\alpha}{2}\right) + \mu_1 \sin\left(\frac{\alpha}{2}\right) \right) (q_3 + \mu_1 q_4) \mu_2 \end{aligned}$$

Commençons alors par analyser les rotations qui sont réalisées sur la projection de l'image dans le plan $(1, \mu_1)$:

$$\left(\cos\left(\frac{\alpha}{2}\right) + \mu_1 \sin\left(\frac{\alpha}{2}\right) \right) (q_1 + \mu_1 q_2) = \left(q_1 \cos\left(\frac{\alpha}{2}\right) - q_2 \sin\left(\frac{\alpha}{2}\right) \right) + \mu_1 \left(q_1 \sin\left(\frac{\alpha}{2}\right) + q_2 \cos\left(\frac{\alpha}{2}\right) \right)$$

où μ est un quaternion unitaire et pur. Nous pouvons alors remarquer que le quaternion q subit une rotation anti-horaire d'angle $\frac{\alpha}{2}$ dans le plan $(1, \mu_1)$.

De même, les rotations qui sont réalisées dans le plan $(\mu_1, \mu_1 \mu_2)$ s'écrivent

$$\left(\cos\left(\frac{\alpha}{2}\right) + \mu_1 \sin\left(\frac{\alpha}{2}\right) \right) (q_3 + \mu_1 q_4) \mu_2 = \left(\left(q_3 \cos\left(\frac{\alpha}{2}\right) - q_4 \sin\left(\frac{\alpha}{2}\right) \right) + \mu_1 \mu_2 \left(q_3 \sin\left(\frac{\alpha}{2}\right) + q_4 \cos\left(\frac{\alpha}{2}\right) \right) \right)$$

Les rotations qui sont réalisées dans le plan $(\mu_1, \mu_1 \mu_2)$ sont les mêmes que dans le plan $(1, \mu_1)$. On peut donc en conclure que les rotations effectuées sont isoclines.

Les rotations simples

Les rotations simples de \mathbb{R}^4 s'expriment dans l'algèbre des quaternions comme

$$\begin{aligned} q' &= sqs^{-1} \\ &= \left(\cos\left(\frac{\alpha}{2}\right) + \mu \sin\left(\frac{\alpha}{2}\right) \right) ((q_1 + \mu_1 q_2) + (q_3 + \mu_1 q_4) \mu_2) \left(\cos\left(\frac{\alpha}{2}\right) - \mu \sin\left(\frac{\alpha}{2}\right) \right) \end{aligned}$$

avec q un quaternion quelconque et s un quaternion unitaire pur.

En réalisant le même type de calcul que pour les rotations isoclines, les rotations de la projection de l'image dans le plan $(1, \mu_1)$ sont :

$$s(q_1 + \mu_1 q_2) s^{-1} = (q_1 + \mu_1 q_2)$$

Nous pouvons remarquer que les rotations simples n'agissent pas dans le plan $(1, \mu_1)$.

Dans le plan $(\mu_1, \mu_1 \mu_2)$, les rotations qui agissent s'écrivent

$$s((q_3 + \mu_1 q_4) \mu_2) s^{-1} = (\cos(\alpha)q_3 - \sin(\alpha)q_4) \mu_2 + (\sin(\alpha)q_3 - \cos(\alpha)q_4) \mu_1 \mu_2$$

On s'aperçoit alors facilement qu'une rotation d'angle α agit dans le plan $(\mu_1, \mu_1 \mu_2)$.

La rotation sqs^{-1} tourne alors uniquement dans un plan.

Les rotations générales

Soit $q \in \mathbb{H}$ un quaternion et $s_1, s_2 \in \mathbb{H}_1$, alors l'équation

$$q' = s_1 q s_2^{-1} \tag{A.19}$$

est une rotation de \mathbb{R}^4 . Toutes les rotations dans $SO(4)$ peuvent être représentées, et il y a deux éléments (s_1, s_2) et $(-s_1, -s_2)$ dans $S^3 \times S^3$ qui représentent la même rotation qui est $s_1 q s_2 = (-s_1) q (-s_2)^{-1}$ [Lounesto 01].

Cela nous montre que $SO(4)$ est isomorphe à $S^3 \times S^3$. Une rotation en dimension 4 peut être représentée par une paire de quaternions unitaires. Cette rotation a donc 6 paramètres.

Ces rotations ont deux plans invariants qui sont complètement orthogonaux. Les plans invariants peuvent tourner arbitrairement.

Les algèbres de Clifford

Dans cette annexe, nous rappelons quelques définitions concernant les algèbres de Clifford ainsi que leurs propriétés et leurs opérations qui sont utilisées dans ce manuscrit. Pour plus de détails, voir [Hestenes 84].

Définitions

Une **algèbre de Clifford** $Cl(E, Q)$ est une algèbre associative unitaire qui est engendrée par un espace vectoriel E muni d'une forme quadratique Q . Cette algèbre géométrique contient des éléments connus comme les scalaires ou les vecteurs, mais aussi d'autres éléments qui sont définis à l'aide du **produit géométrique**. À partir du produit géométrique, deux nouveaux produits peuvent être définis :

Pour deux vecteurs \mathbf{u} et \mathbf{v} de \mathbb{R}^n , la quantité scalaire est appelée le produit intérieur :

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}(\mathbf{u}\mathbf{v} + \mathbf{v}\mathbf{u}) = B(\mathbf{u}, \mathbf{v}) = \mathbf{v} \cdot \mathbf{u} \quad (\text{B.1})$$

avec $B(\mathbf{u}, \mathbf{v})$ la forme bilinéaire de E .

La quantité qui n'est pas scalaire

$$\mathbf{u} \wedge \mathbf{v} = \frac{1}{2}(\mathbf{u}\mathbf{v} - \mathbf{v}\mathbf{u}) = -\mathbf{v} \wedge \mathbf{u} \quad (\text{B.2})$$

est appelée le produit extérieur.

Par conséquent, le produit géométrique se décompose en deux parties symétrique et antisymétrique et s'écrit

$$\begin{aligned} \mathbf{u}\mathbf{v} &= B(\mathbf{u}, \mathbf{v}) + \mathbf{u} \wedge \mathbf{v} \\ \mathbf{u}^2 &= B(\mathbf{u}, \mathbf{u}) = Q(\mathbf{u}) \end{aligned} \quad (\text{B.3})$$

Nous pouvons alors constater que $\mathbf{u} \wedge \mathbf{v}$ n'est plus un vecteur mais un **bivecteur**. Un élément général de l'algèbre de Clifford est appelé **multivecteur**.

Si l'espace vectoriel E est de dimension n et si $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ est une base de E , alors **la base de l'algèbre de Clifford** $Cl(E, Q)$ est définie par

$$\{\mathbf{e}_{i_1}\mathbf{e}_{i_2}\dots\mathbf{e}_{i_k} \mid 1 \leq i_1 < i_2 < \dots < i_k \leq n \text{ et } 0 \leq k \leq n\} \quad (\text{B.4})$$

La dimension totale de $Cl(E, Q)$ est alors la somme de toutes les combinaisons possibles des vecteurs de base soit 2^n .

On notera $\mathbb{R}_{p,q}$ (avec $n = p + q$) l'algèbre de Clifford de l'espace vectoriel \mathbb{R}^n muni de la forme quadratique $Q(x_1, x_2, \dots, x_n) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_n^2$. Remarquons que si q est égal à

0 alors la **forme quadratique usuelle** (euclidienne) est obtenue.

Dans le cas de l'algèbre de $\mathbb{R}_{3,0}$, on trouve alors des **scalaires**, des **vecteurs** (e_1, e_2, e_3), des **bivecteurs** (e_1e_2, e_1e_3, e_2e_3) et un **trivecteur** ($e_1e_2e_3$ aussi appelé **pseudo-scalaire** dans le cas de $\mathbb{R}_{3,0}$).

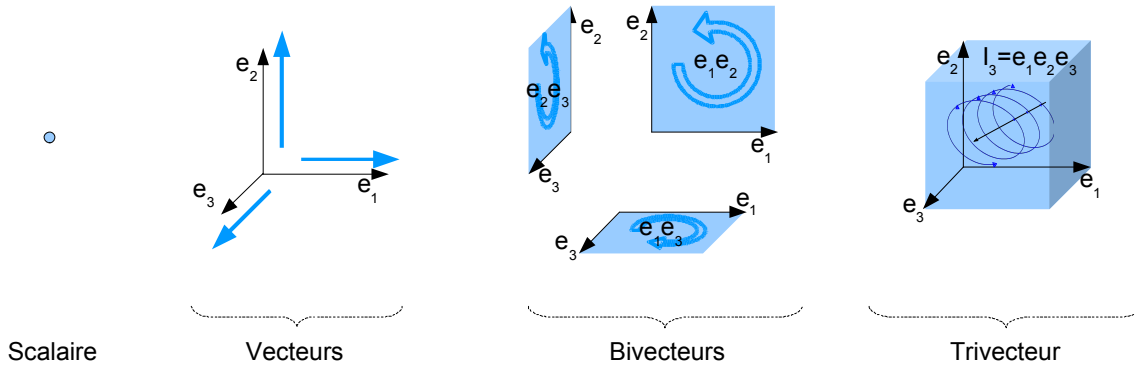


FIGURE B.1 – Illustration des éléments de $\mathbb{R}_{3,0}$

Un multivecteur M est décomposable en une somme d'éléments de **grades** différents :

$$M = \langle M \rangle_0 + \langle M \rangle_1 + \cdots + \langle M \rangle_n = \sum_r \langle M \rangle_r \quad (\text{B.5})$$

La quantité $\langle M \rangle_r$ est appelée la partie **r-vectorielle** de M . Plus spécifiquement, on appelle partie scalaire (respectivement vectorielle, bivectorielle, etc.) un 0-vecteur (respectivement 1-vecteur, 2-vecteur, etc.). On notera $\mathbb{R}_{p,q}^k$ le sous-espace de dimension $n + q$ des k -vecteurs de $\mathbb{R}_{p,q}$. Par exemple, $\mathbb{R}_{p,q}^1$ correspond au sous-espace des vecteurs de $\mathbb{R}_{p,q}$.

Pour un entier $r > 0$, un multivecteur M_r est appelé un **r-blade** ou un **simple r-vecteur** si et seulement si il peut être factorisé comme le produit de r vecteurs qui anti-commutent soit :

$$M_r = \mathbf{m}_1 \mathbf{m}_2 \cdots \mathbf{m}_r \quad (\text{B.6})$$

où $\mathbf{m}_j \mathbf{m}_k = -\mathbf{m}_k \mathbf{m}_j$ pour $j, k = 1, 2, \dots, r$ et $j \neq k$.

Si les vecteurs \mathbf{m}_j et \mathbf{m}_k anti-commutent, c'est qu'ils sont orthogonaux. Le produit géométrique $\mathbf{m}_j \mathbf{m}_k$ est alors égale à $\mathbf{m}_j \wedge \mathbf{m}_k$.

Pour deux vecteurs \mathbf{u} et \mathbf{v} orthogonaux entre eux, le cas particulier des 2-blades s'écrit $\mathbf{uv} = \mathbf{u} \wedge \mathbf{v}$.

Les différents produits

Il existe dans les algèbres de Clifford plusieurs produits dont voici les quatre principaux : le produit géométrique, le produit intérieur \cdot , le produit extérieur \wedge et le produit scalaire $*$.

- Le produit géométrique entre deux multivecteurs A et B s'écrit AB . Ce produit n'est pas commutatif mais il est associatif et distributif par rapport à l'addition, c'est à dire :

$$\begin{aligned}(AB)C &= A(BC) \\ A(B + C) &= AB + AC \\ (B + C)A &= BA + CA\end{aligned}\tag{B.7}$$

pour trois multivecteurs A , B et C .

- Le produit intérieur \cdot entre un r -blade A_r et un s -blade B_s s'écrit

$$A_r \cdot B_s \equiv \langle A_r, B_s \rangle_{|r-s|}, \text{ si } r, s > 0\tag{B.8}$$

Pour deux multivecteurs A et B , il s'écrit

$$A \cdot B = \sum_r \sum_s \langle A \rangle_r \cdot \langle B \rangle_s\tag{B.9}$$

- Le produit extérieur \wedge entre un r -blade A_r et un s -blade B_s s'écrit

$$A_r \wedge B_s \equiv \langle A_r, B_s \rangle_{r+s}, \text{ si } r, s > 0\tag{B.10}$$

Pour deux multivecteurs A et B , il s'écrit

$$A \wedge B = \sum_r \sum_s \langle A \rangle_r \wedge \langle B \rangle_s\tag{B.11}$$

- Le produit scalaire $*$ est défini par

$$A * B = \langle AB \rangle_0\tag{B.12}$$

Opérations

Les principales opérations des algèbres de Clifford sont :

- La reversion d'un produit de vecteurs :

$$(\mathbf{v}_1 \widetilde{\mathbf{v}_2 \cdots \mathbf{v}_r}) = \mathbf{v}_r \cdots \mathbf{v}_2 \mathbf{v}_1\tag{B.13}$$

- L'inverse d'un r -blade A_r non nul :

$$A_r^{-1} = |A_r|^{-2} \widetilde{A}_r\tag{B.14}$$

- Le dual d'un r -blade A_r :

$$A_r^* = I A_r\tag{B.15}$$

où I est le pseudo-scalaire de l'algèbre considéré.

- Le conjugué d'un r -blade A_r :

$$\overline{A}_r = (-1)^r (\widetilde{A}_r)\tag{B.16}$$

- La norme (aussi appelée magnitude) d'un multivecteur A :

$$|A|^2 = \widetilde{A} * A = \sum_r |\langle A_r \rangle|^2\tag{B.17}$$

- Le cosinus entre deux multivecteurs A et B :

$$\cos \angle AB = \frac{\widetilde{A} * B}{|A||B|}\tag{B.18}$$

Projections

Dans [Hestenes 84], une décomposition d'un vecteur \mathbf{v} par rapport à un r -blade B non nul est donnée par

$$\begin{aligned}\mathbf{v} &= \mathbf{v}BB^{-1} = (\mathbf{v} \cdot B + \mathbf{v} \wedge B)B^{-1} \\ &= \mathbf{v}_{\parallel B} + \mathbf{v}_{\perp B}\end{aligned}\quad (\text{B.19})$$

Les quantités $\mathbf{v}_{\parallel B}$ et $\mathbf{v}_{\perp B}$ représentent les projections parallèles et orthogonales de \mathbf{v} par rapport à B :

$$\begin{aligned}\mathbf{v}_{\parallel B} &= P_B(\mathbf{v}) \equiv (\mathbf{v} \cdot B)B^{-1} \\ \mathbf{v}_{\perp B} &= P_B^\perp(\mathbf{v}) \equiv (\mathbf{v} \wedge B)B^{-1}\end{aligned}\quad (\text{B.20})$$

Les algèbres de Clifford et les quaternions

Soit $(\mathbf{e}_1, \mathbf{e}_2)$ une base orthonormée de \mathbb{R}^2 . L'algèbre de Clifford $\mathbb{R}_{0,2}$ est une algèbre de dimension 4 qui vérifie l'égalité $\mathbf{e}_1^2 = \mathbf{e}_2^2 = \mathbf{e}_1\mathbf{e}_2 = -1$. De plus, ces éléments anti-commutent tous car ils sont orthogonaux. On peut alors vérifier que l'on réalise un isomorphisme entre $\mathbb{R}_{0,2}$ et \mathbb{H} en réalisant l'identification suivante

$$1 \leftrightarrow 1 \quad \mathbf{e}_1 \leftrightarrow i \quad \mathbf{e}_2 \leftrightarrow j \quad \mathbf{e}_1\mathbf{e}_2 \leftrightarrow k \quad (\text{B.21})$$

De même, on peut prouver qu'une sous-algèbre de $\mathbb{R}_{3,0}$ est isomorphe aux quaternions. En effet, cette algèbre est de dimension 8 et admet pour base $(1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1\mathbf{e}_2, \mathbf{e}_1\mathbf{e}_3, \mathbf{e}_2\mathbf{e}_3, \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3)$. Les éléments pairs de cette algèbre vérifient l'égalité $(\mathbf{e}_1\mathbf{e}_2)^2 = (\mathbf{e}_2\mathbf{e}_3)^2 = (\mathbf{e}_1\mathbf{e}_3)^2 = -1$. Cette sous-algèbre de $\mathbb{R}_{3,0}$ est isomorphe à \mathbb{H} en réalisant l'identification suivante

$$1 \leftrightarrow 1 \quad \mathbf{e}_1\mathbf{e}_2 \leftrightarrow i \quad \mathbf{e}_2\mathbf{e}_3 \leftrightarrow j \quad \mathbf{e}_1\mathbf{e}_3 \leftrightarrow k \quad (\text{B.22})$$

Enfin, l'espace vectoriel \mathbb{H} des quaternions peut être identifié à l'espace vectoriel \mathbb{R}^4 sous l'identification :

$$\mathbf{e}_1 \leftrightarrow i \quad \mathbf{e}_2 \leftrightarrow j \quad \mathbf{e}_3 \leftrightarrow k \quad \mathbf{e}_4 \leftrightarrow 1 \quad (\text{B.23})$$

La sous-algèbre constituée des vecteurs de $\mathbb{R}_{4,0}$ est alors isomorphe à \mathbb{H} .

Caractérisation des rotations de \mathbb{R}^2 , \mathbb{R}^3 et \mathbb{R}^4

Pour réaliser des rotations des vecteurs de $\mathbb{R}_{n,0}$, le groupe $Spin(n)$ doit être considéré. Celui-ci est défini par

$$Spin(n) = \{s \in \mathbb{R}_{n,0}^{(0)}, N(s) = 1, s^{-1}\mathbf{v}s \in \mathbb{R}_{n,0}^1, \forall \mathbf{v} \in \mathbb{R}_{n,0}^1\} \quad (\text{B.24})$$

où $N(s) = s\bar{s}$. Dans [Lounesto 01], il est spécifié que tout spineur de $spin(3)$ et $spin(4)$ peut s'écrire comme l'exponentielle d'un bivecteur. Plus généralement, tout spineur s peut s'écrire comme l'exponentielle d'un bivecteur (voir [Helgason 01, Batard 09]),

$$s = exp(B) = \sum_{i \geq 0} \frac{1}{i!} B^i \quad (\text{B.25})$$

pour un bivecteur B .

Le groupe $Spin(n)$ est étroitement lié au groupe $SO(n)$ (le groupe spécial orthogonal de dimension n). En effet, l'action de $Spin(n)$ sur les vecteurs de $\mathbb{R}_{n,0}$ correspond à l'action de $SO(n)$ sur \mathbb{R}^n [Batard 09].

Les rotations de \mathbb{R}^2

Il est bien connu que les nombres complexes unitaires $z \in \mathbb{C}$, $|z| = 1$ forment le cercle unité $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$. On peut alors réécrire les nombres complexes unitaires sous la forme d'une exponentielle complexe $e^{i\theta} = \cos(\theta) + i \sin(\theta)$. Une rotation anti-horaire dans le plan complexe \mathbb{C} par un angle θ peut être représentée par la multiplication :

$$(x + iy) \mapsto (x + iy)e^{i\theta} = (x + iy)(\cos(\theta) + i \sin(\theta)) \quad (\text{B.26})$$

En plongeant successivement dans \mathbb{R}^2 , puis \mathbb{R}^2 dans les vecteurs de $\mathbb{R}_{2,0}$, un nombre complexe peut être vu comme un vecteur de cette algèbre. De plus, en identifiant l'imaginaire pur i à e_1e_2 , on peut réécrire la rotation de l'équation B.26 comme

$$\begin{aligned} (xe_1 + ye_2) \mapsto & \left(\cos\left(-\frac{\theta}{2}\right) + e_1e_2 \sin\left(-\frac{\theta}{2}\right) \right) (xe_1 + ye_2) \left(\cos\left(\frac{\theta}{2}\right) + e_1e_2 \sin\left(\frac{\theta}{2}\right) \right) \\ & = e^{-e_1e_2 \frac{\theta}{2}} (xe_1 + ye_2) e^{e_1e_2 \frac{\theta}{2}} \end{aligned} \quad (\text{B.27})$$

avec $\cos(\frac{\theta}{2}) + e_1e_2 \sin(\frac{\theta}{2}) \in Spin(2) = \{s \in \mathbb{R}_{2,0}^{(0)}, N(s) = 1, s^{-1}\mathbf{v}s \in \mathbb{R}_{2,0}^1, \forall \mathbf{v} \in \mathbb{R}_{2,0}^1\}$. L'action de $Spin(2)$ sur $\mathbb{R}_{2,0}^1$ correspond alors à l'action de S^1 sur \mathbb{C} .

Les rotations de \mathbb{R}^3

Dans l'annexe dédiée aux quaternions, nous avons vu que la multiplication d'un quaternion pur $v \in \mathbb{H}_0$ par un quaternion unitaire $q \in \mathbb{H}_1$ réalise une rotation dans \mathbb{R}^3 . Pour coder des rotations de \mathbb{R}^3 à l'aide des algèbres de Clifford, il faut s'intéresser aux éléments de $Spin(3)$. On peut vérifier que $Spin(3)$ est le groupe

$$Spin(3) = \{a1 + be_1e_2 + ce_1e_3 + de_3e_1, a^2 + b^2 + c^2 + d^2 = 1\} \quad (\text{B.28})$$

et qu'il est isomorphe à \mathbb{H}_1 .

Les rotations de \mathbb{R}^4

Les rotations réalisées par $Spin(4)$ sont de trois types :

- Les rotations simples sont l'exponentielle de bivecteurs simples qui sont le produit extérieur de deux vecteurs. Ces rotations tournent uniquement dans un plan.
- Les rotations isoclines sont l'exponentielle de bivecteurs simples multipliés par un des éléments $(1 \pm I_4)/2$. Une rotation isocline a une infinité de plans de rotation.

- Les rotations générales ont deux plans invariants qui sont complètement orthogonaux avec différents angles de rotations.

Les rotations générales

Les éléments du groupe $Spin(4)$ sont 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 peut alors noter

$$i = e_1e_2 \quad j = e_2e_3 \quad k = e_1e_3 \quad i' = e_4e_3 \quad j' = e_4e_1 \quad k' = e_4e_2 \quad I_4 = e_1e_2e_3e_4 \quad (\text{B.29})$$

Les relations suivantes peuvent alors être vérifiées

$$ij = k \quad jk = i \quad ki = j \quad i^2 = j^2 = k^2 = -1 \quad iI_4 = I_4i = i' \quad jI_4 = I_4j = j' \quad kI_4 = I_4k = k' \quad (\text{B.30})$$

Si \mathbf{v} est un vecteur de $\mathbb{R}_{4,0}$ et $s \in Spin(4)$, alors $s\mathbf{v}s^{-1}$ est un vecteur de $\mathbb{R}_{4,0}$.

On peut en déduire que tout élément $s \in Spin(4)$ s'écrit sous la forme

$$s = (u + I_4v) = (a1 + be_1e_2 + ce_2e_3 + de_3e_1) + I_4(a'1 + b'e_1e_2 + c'e_2e_3 + d'e_3e_1) \quad (\text{B.31})$$

où I_4 est le pseudo-scalaire de $\mathbb{R}_{4,0}$ et u, v sont des quaternions tels que $u\bar{u} + v\bar{v} = 1$ et $u\bar{v} + v\bar{u} = 0$. L'application qui associe $(u + v, u - v)$ à $u + I_4v$ est un isomorphisme, ce qui implique que $Spin(4)$ est isomorphe à $Spin(3) \times Spin(3)$. Sachant que $Spin(3)$ est isomorphe à \mathbb{H}_1 , on peut alors prouver que $Spin(4)$ est isomorphe à $\mathbb{H}_1 \times \mathbb{H}_1$.

Les rotations simples

Pour deux vecteurs unitaires \mathbf{v}_1 et \mathbf{v}_2 de \mathbb{R}^4 , La rotation simple qui envoie \mathbf{v}_1 sur \mathbf{v}_2 utilise l'élément pair s de norme 1 :

$$s = \frac{(1 + \mathbf{v}_2\mathbf{v}_1)}{\sqrt{2(1 + \mathbf{v}_2 \cdot \mathbf{v}_1)}} \quad (\text{B.32})$$

qui satisfait $s\mathbf{v}_1s^{-1} = \mathbf{v}_2$. La rotation simple qui lui est alors associée s'écrit

$$\mathbf{v} \mapsto s\mathbf{v}s^{-1} \quad (\text{B.33})$$

Nous pouvons exprimer s sous la forme d'une exponentielle d'un bivecteur :

$$s = \exp \left[\frac{\theta}{2} \frac{\mathbf{v}_2 \wedge \mathbf{v}_1}{|\mathbf{v}_2 \wedge \mathbf{v}_1|} \right] \quad (\text{B.34})$$

Démonstration. Nous savons que

$$\begin{aligned} |\mathbf{v}_2 \cdot \mathbf{v}_1| &= \cos(\theta) = 2\cos^2(\theta/2) - 1 \\ |\mathbf{v}_2 \wedge \mathbf{v}_1| &= \sin(\theta) = 2\cos(\theta/2)\sin(\theta/2) \end{aligned} \quad (\text{B.35})$$

alors

$$\begin{aligned}
\frac{1 + \mathbf{v}_2 \mathbf{v}_1}{\sqrt{2(1 + \mathbf{v}_2 \cdot \mathbf{v}_1)}} &= \frac{1 + \mathbf{v}_2 \mathbf{v}_1}{\sqrt{4\cos^2(\theta/2)}} \\
&= \frac{1}{2\cos(\theta/2)} + \frac{(\mathbf{v}_2 \cdot \mathbf{v}_1)}{2\cos(\theta/2)} + \frac{(\mathbf{v}_2 \wedge \mathbf{v}_1)}{2\cos(\theta/2)} \\
&= \frac{1}{2\cos(\theta/2)} + \frac{2\cos^2(\theta/2) - 1}{2\cos(\theta/2)} + \frac{(\mathbf{v}_2 \wedge \mathbf{v}_1) |\mathbf{v}_2 \wedge \mathbf{v}_1|}{2\cos(\theta/2) |\mathbf{v}_2 \wedge \mathbf{v}_1|} \\
&= \frac{1}{2\cos(\theta/2)} - \frac{1}{2\cos(\theta/2)} + \cos(\theta/2) + \frac{(\mathbf{v}_2 \wedge \mathbf{v}_1) 2\cos(\theta/2) \sin(\theta/2)}{2\cos(\theta/2) |\mathbf{v}_2 \wedge \mathbf{v}_1|} \\
&= \cos(\theta/2) + \frac{(\mathbf{v}_2 \wedge \mathbf{v}_1)}{|\mathbf{v}_2 \wedge \mathbf{v}_1|} \sin(\theta/2) \\
&= \exp \left[\frac{\theta}{2} \frac{\mathbf{v}_2 \wedge \mathbf{v}_1}{|\mathbf{v}_2 \wedge \mathbf{v}_1|} \right] \tag{B.36}
\end{aligned}$$

□

Les rotations isoclines

Nous avons vu que les rotations isoclines sont l'exponentielle de bivecteurs simples multipliés par un des éléments $(1 \pm I_4)/2$. Les éléments s de $Spin(4)$ qui réalisent des rotations isoclines peuvent alors s'écrire

$$s = e^{\frac{\theta}{2}(1+I_4)B} = e^{\frac{\theta}{2}(B+I_4B)} \tag{B.37}$$

Une rotation isocline de \mathbb{R}^4 s'exprime lors dans les algèbres de Clifford comme

$$s \mathbf{v} s^{-1} = e^{\frac{\theta}{2}(B+I_4B)} \mathbf{v} e^{-\frac{\theta}{2}(B+I_4B)} \tag{B.38}$$

Algorithmes

Dans cette annexe, nous donnons l'implémentation MATLAB® de la transformée de Fourier Clifford couleur ainsi que de la transformée inverse.

La première fonction permet de calculer la transformée de Fourier Clifford couleur d'une image à trois canaux (par exemple *RVB*) donnée en paramètre. La direction d'analyse $B_c = \mathbf{c} \wedge e_4$ est définie par l'utilisateur, qui donne en paramètre de cette fonction un vecteur couleur \mathbf{c} . Les parties parallèle et orthogonale de la transformée ainsi que la transformée de Fourier Clifford reconstruite sont données en sortie.

La seconde fonction correspond à la transformée de Fourier Clifford inverse. Elle prend en paramètre une matrice de vecteurs de \mathbb{R}^4 , correspondant à la transformée de Fourier Clifford d'une image, ainsi qu'un vecteur couleur similaire à celui utilisé pour le calcul de la transformée de Fourier Clifford couleur. Une image couleur est alors retournée par cette fonction.

Nous donnons également dans cette annexe l'implémentation MATLAB® de la construction de la "maison de Fourier" couleur utilisée dans ce manuscrit, *i.e.*,

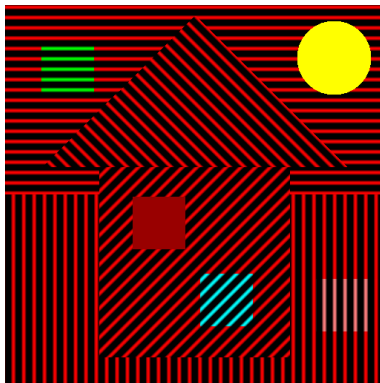


FIGURE C.1 – La "maison de Fourier" couleur

Algorithme 4 Implémentation MATLAB® de la transformée Fourier Clifford couleur

Tache : Calculer la transformée de Fourier Clifford couleur d'une image couleur avec un bivecteur $B_c = c \wedge e_4$.

Paramètres :

- *im* : une image couleur
- *c* : une couleur

Sorties :

- *CFT* : la transformée de Fourier Clifford couleur reconstruite
- *CFT_par* : la partie parallèle de la transformée de Fourier Clifford couleur
- *CFT_orth* : la partie orthogonale de la transformée de Fourier Clifford couleur

Algorithme :

```

I=double(I)/255;
C=C/norm(C);
I_v = reshape(I,size(I,1)*size(I,2),size(I,3));

if(sum(abs(C-ones(3,1)/sqrt(3)))~=0)
    mu=ones(3,1)/sqrt(3);
else
    mu=[1;0;0];
end

% Computation of the orthonormal basis {C,e4,vn,wn}

v=cross(C,cross(mu,C));
w=cross(v,C);

vn=v/norm(v);
wn=w/norm(w);

% Projection of the color image on the orthonormal basis

Real_par_Proj=reshape(I_v*C,size(I,1),size(I,2));
Imag_par_Proj=zeros(size(Real_par_Proj));
par_im = Real_par_Proj+Imag_par_Proj*1i;

Real_orth_Proj=reshape(I_v*vn,size(I,1),size(I,2));
Imag_orth_Proj=reshape(I_v*wn,size(I,1),size(I,2));
orth_im = Real_orth_Proj+Imag_orth_Proj*1i;

% Computation of two FFT2s on the two projected Image

CFT_par = fft2(par_im);
CFT_orth = fft2(orth_im);

if nargout == 3 % Reconstruction of the Clifford Fourier Transform
    alpha=real(CFT_par);
    beta=imag(CFT_par);
    gamma=real(CFT_orth);
    rho=imag(CFT_orth);
    CFT_recons(:,:,1)=alpha*C(1)+gamma*vn(1)+rho*wn(1);
    CFT_recons(:,:,2)=alpha*C(2)+gamma*vn(2)+rho*wn(2);
    CFT_recons(:,:,3)=alpha*C(3)+gamma*vn(3)+rho*wn(3);
    CFT_recons(:,:,4)=beta;
end

```

Algorithme 6 Implémentation MATLAB® de la “maison de Fourier” couleur

Tache : Construire une “maison de Fourier” couleur

Paramètres :

- *height* : hauteur de l'image
- *width* : largeur de l'image
- *colors* : une matrice de taille 3*4 qui contient les couleurs RVB de la partie supérieure et de la partie inférieure du fond, du mur, du toit.
- *sin_parameters* : une matrice de taille 3*4 qui contient les paramètres (amplitude, fréquence, phase, orientation) des sinus composants la partie supérieure et la partie inférieure du fond, le mur, le toit

Sorties :

- *fourier_house* : une image contenant la maison de Fourier couleur

Algorithme :

```
[background_up]=plot_rectangle(height,width,floor(height/2)-1,width-1,...
                               [1,1],colors(:,1),sin_parameters(:,1));
[background_down]=plot_rectangle(height,width,floor(height/2)-1,width-1,...
                                 [floor(height/2)+1,1],colors(:,2),sin_parameters(:,2));
background=background_up+background_down;

width_wall=floor(width/2);
height_wall=floor(height/2);
whg_c=floor(width/4);
whg_l=floor(height/2)-floor(width/4);
[im_wall,mask_wall]=plot_rectangle(height,width,height_wall,width_wall,...
                                   [whg_l,whg_c],colors(:,3),sin_parameters(:,3));

fourier_house=add_elt(background,im_wall,mask_wall);

div_width=10;
div_height=2.5;
[l,c]=find(mask_wall);
base_roof_l=min(l)-1;
roof_height=floor(height/div_height);
pos=[base_roof_l base_roof_l base_roof_l-round(whg_l/1.1);...
     floor(width/div_width) width-floor(width/div_width) floor(width/2+1)];
[im_roof,mask_roof]=plot_triangle(height,width,pos,colors(:,4),sin_parameters(:,4));
fourier_house=add_elt(fourier_house,im_roof,mask_roof);
```

Algorithme 7 Script de construction d'une "maison de Fourier" couleur

Tache : Construit une "maison de Fourier" couleur

Algorithme :

```

width=512;
height=512;

%%% Première étape : construire la maison de Fourier couleur

% % couleurs R,G,B
colors=[1 0 0; % fond : partie supérieur
        1 0 0; % fond : partie inférieure
        1 0 0; % mur
        1 0 0]';% toit

% % paramètres des sinus : amplitude, fréquence, phase, orientation
sin_parameters=[1 0.45 0 0; % fond : partie supérieur
                1 0.45 2.5 pi/2; % fond : partie inférieure
                1 0.45 0 pi/4; % mur
                1 0.45 0 3*pi/4]'; % toit
fourier_house=basic_fourier_house(height,width,colors,sin_parameters);

%%% seconde étape : ajouter des éléments

Col=[1;0;0];
sin_parameters=[0 0 0 0];
[im_circle,mask_circle]=plot_circle(height,width,[(height/4) (width/2)],70,Col,...
                                     sin_parameters);
colored_mask=reshape(repmat(mask_circle(:),1,3).*repmat(Col',...
                                     size(mask_circle(:),1),1),height,width,3);
fourier_house=add_elt(fourier_house,colored_mask,mask_circle);

Col=[0;1;0];
sin_parameters=[1 0.45 0 0];
[square_im,mask]=plot_rectangle(height,width,70,70,[50,50],Col,sin_parameters);
fourier_house=add_elt(fourier_house,square_im,mask);

Col=[0;1;1];
sin_parameters=[1 0.45 0 pi/4];
[square_im,mask]=plot_rectangle(height,width,70,70,[364,264],Col,sin_parameters);
fourier_house=add_elt(fourier_house,square_im,mask);

Col=[0.6;0;0];
sin_parameters=[0 0 0 0];
[square_im,mask]=plot_rectangle(height,width,70,70,[260,173],Col,sin_parameters);
colored_mask=reshape(repmat(mask(:),1,3).*repmat(Col',size(mask(:),1),1),...
                                     height,width,3);
fourier_house=add_elt(fourier_house,colored_mask,mask);

imshow(fourier_house);
imwrite(fourier_house, '\my_house.png');

```

Publications

Chapitre de livre

1. *Color Object Recognition Based On Clifford Fourier Transform*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, Guide to Geometric Algebra in practice. Springer.

Conférences internationales avec comité de lecture

1. *New geometric Fourier descriptors for color image recognition*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, International Conference on Image Processing (ICIP 2010), 26-29 septembre 2010, Hong-Kong (Chine).
2. *A phase correlation for color images using Clifford algebra*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, Applied Geometric Algebras in Computer Science and Engineering (AGACSE 2010), 14-16 juin 2010, Amsterdam (Pays-Bas).

Conférences nationales avec comité de lecture

1. *De nouveaux descripteurs de Fourier géométriques pour l'analyse d'images couleur*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, Reconnaissance des Formes et Intelligence Artificielle (RFIA 2010), 19-22 janvier 2010, Caen (France).
2. *Une corrélation de phase couleur pour la reconnaissance de formes*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, XXIIIe Colloque GRETSI - Traitement du Signal et des Images, 5-8 septembre 2011, Bordeaux (France).
3. *Un nouvel ensemble de descripteurs de Fourier Clifford pour les images couleur : les GCFD3*, José Mennesson, Christophe Saint-Jean, Laurent Mascarilla, Accepté pour publication dans Reconnaissance des Formes et Intelligence Artificielle (RFIA 2012), 24-27 janvier 2012, Lyon (France).

Table des figures

1.1	<i>Illustration du log-module de la transformée de Fourier à chaque étape de la construction de “la maison de Fourier”</i>	23
1.2	<i>Reconstruction de l’image de la “maison de Fourier” en conservant uniquement le module (en haut) et la phase (en bas)</i>	24
1.3	<i>Reconstruction de l’image de Lenna en conservant uniquement le module (en haut) et la phase (en bas)</i>	25
1.4	<i>Passage d’une image en coordonnées polaires</i>	27
1.5	<i>Transformée de Fourier-Mellin analytique d’une image</i>	28
1.6	<i>Transformée de Fourier-Mellin analytique d’une image ayant subi une rotation d’angle $\frac{\pi}{2}$ et une homothétie de rapport 0.75</i>	29
1.7	<i>Partie simplexe (à gauche) et partie perplexe (à droite) d’une image selon le choix $\mu_1 = (i + j + k)/\sqrt{3}$</i>	45
1.8	<i>Transformée de Fourier quaternionique à gauche appliquée sur une image jaune contenant un carré bleu (en haut) et sur la maison de Fourier couleur (en bas).</i>	48
1.9	<i>Illustration de la base $\{\mathbf{c}, \mathbf{e}_4, \mathbf{v}, \mathbf{w}\}$ de I_4 en utilisant la librairie GABLE [Dorst 02]</i>	57
1.10	<i>Système d’équations permettant la reconstruction de la transformée de Fourier Clifford</i>	59
1.11	<i>Visualisation de la transformée de Fourier Clifford couleur avec le bivecteur $B_r = \text{rouge} \wedge \mathbf{e}_4$</i>	60
1.12	<i>Reconstruction de l’image à partir du module (à gauche) et de la phase (à droite) de la transformée de Fourier Clifford couleur</i>	60
1.13	<i>Visualisation de la transformée de Fourier Clifford couleur à plusieurs étapes de la construction de la “maison de Fourier couleur”</i>	62
2.1	<i>Illustration de différentes transformations de l’image de Lenna</i>	69
2.2	<i>Illustration du module de la transformée de Fourier polaire \widehat{f}^{P^2} d’une image f de la base COIL-100 [Nene 96]</i>	76
2.3	<i>Calcul de GFD1 (à gauche) et calcul de GFD2 (à droite)</i>	82
2.4	<i>Extraction de la signature $a(k)$ d’une image de la base COIL-100</i>	84
2.5	<i>Extraction de descripteurs de Fourier selon la méthode marginale</i>	85
2.6	<i>Extraction des descripteurs GCFD1 d’une image couleur</i>	88
3.1	<i>Les 100 objets de COIL-100 (à gauche) - une image de la base (à droite)</i>	99
3.2	<i>Quelques images de la base color FERET</i>	99
3.3	<i>Quelques images de la base color FERET bruitée</i>	100
3.4	<i>Quelques images de la base ALOI</i>	100
3.5	<i>COIL-100 : Taux de reconnaissance pour GFD1 (à gauche) et GCFD1 (à droite) avec différentes valeurs de B.</i>	103
3.6	<i>color FERET : Taux de reconnaissance avec GCFD1 pour 100 bivecteurs aléatoires</i>	106
3.7	<i>COIL-100 : Taux de reconnaissance pour GCFD1 avec différentes valeurs de B_T, puis avec 3 bivecteurs (cas marginal et par SFFS)</i>	107

3.8	<i>color FERET : Taux de reconnaissance pour GCFD1 avec différentes valeurs de B_T, puis avec 3 bivecteurs (cas marginal et par SFFS)</i>	108
3.9	<i>ALOI : Taux de reconnaissance pour GCFD1 avec différentes valeurs de B_T, puis avec 3 bivecteurs (cas marginal et par SFFS)</i>	109
3.10	<i>La corrélation de phase pour les images en niveaux de gris</i>	111
3.11	<i>La corrélation de phase couleur avec la transformée de Fourier quaternionique</i>	113
3.12	<i>L'image de similarité calculée à partir de l'agrégation de deux scores avec un bivecteur B</i>	115
3.13	<i>L'image de similarité calculée à partir du cosinus entre multivecteurs</i>	116
3.14	<i>Scores de corrélation entre les images 1 et 2 pour des choix différents de \mathbf{c}_2 et B. Le symbole '-' décrit une valeur qui ne peut pas être calculée parce que la partie parallèle ou orthogonale est nulle.</i>	117
B.1	<i>Illustration des éléments de $\mathbb{R}_{3,0}$</i>	134
C.1	<i>La "maison de Fourier" couleur</i>	141

Bibliographie

- [Arbter 90] K. Arbter, W.E. Snyder, H. Burkhardt & G. Hirzinger. *Application of affine-invariant Fourier descriptors to recognition of 3-D objects*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pages 640–647, jul 1990. 73
- [Batard 09] T. Batard. *Géométrie différentielle des fibrés vectoriels et algèbres de Clifford appliquées au traitement d'images multicanaux*. PhD thesis, 2009. 9, 39, 40, 136, 137
- [Batard 10] T. Batard, M. Berthier & C. Saint-Jean. *Clifford Fourier Transform for Color Image Processing*. In E. Bayro-Corrochano & G. Scheuermann, editeurs, Geometric Algebra Computing in Engineering and Computer Science, chapitre 8, pages 135–161. Springer Verlag, 2010. 5, 11, 12, 17, 18, 49, 51, 61, 63, 87, 108, 123, 124
- [Bayer 76] Bryce E. Bayer. *Color imaging array*, July 1976. 108
- [Bayro-Corrochano 10] E. Bayro-Corrochano & G. Scheuermann. Geometric algebra computing : In engineering and computer science. Springer, 2010. 39
- [Bülow 99] T. Bülow. *Hypercomplex spectral signal representations for the processing and analysis of images*. PhD thesis, University of Kiel, 1999. 5, 9, 12, 17, 19, 34, 37, 38, 129
- [Bülow 01] T. Bülow, M. Felsberg & G. Sommer. Non-commutative hypercomplex fourier tranforms of multidimensional signals, pages 187–207. Springer-Verlag, London, UK, 2001. 37
- [Bornard 86] G. Bornard, J.-P. Gauthier & M. Silbermann. *Mouvements et traitement d'images*. Traitement du Signal, vol. 6, pages 281–290, 1986. 12, 18, 31, 33, 34, 74
- [Bracewell 86] R. Bracewell. The fourier transform and its applications. McGraw-Hill Companies, February 1986. 18, 19, 35, 89
- [Brackx 06] F. Brackx, N. De Schepper & F. Sommen. *The Two-Dimensional Clifford-Fourier Transform*. Journal of Mathematical Imaging and Vision, vol. 26, pages 5–18, 2006. 40
- [Casasent 76] D. Casasent & D. Psaltis. *Position, rotation, and scale invariant optical correlation*. Applied Optics, vol. 15, no. 7, pages 1795–1799, Jul 1976. 25
- [Chang 01] C.-C. Chang & C.-J. Lin. *LIBSVM : a library for support vector machines*, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 101
- [Chen 09] Y. Chen, E. K. Garcia, M. R. Gupta, A. Rahimi & L. Cazzanti. *Similarity-based Classification : Concepts and Algorithms*. Journal of Machine Learning Research, vol. 10, pages 747–776, 2009. 118, 119

- [Chen 10] B. Chen, H. Shu, H. Zhang, G. Chen & L. Luo. *Color Image Analysis by Quaternion Zernike Moments*. In Pattern Recognition (ICPR), 2010 20th International Conference on, pages 625–628, aug. 2010. 44
- [Choksuriwong 08] A. Choksuriwong, B. Emile, H. Laurent & C. Rosenberger. *Comparative study of global invariant descriptors for object recognition*. Journal of Lightwave Technology Electronic Imaging, vol. 17, no. 2, page 023015, 2008. 13
- [Cosgriff 60] R. L. Cosgriff. *Identification of shape*. Rapport technique ASTIA AD 254-792, Ohio State Univ. Res. Foundation, Columbus, Dec. 1960. 70
- [Denis 07] P. Denis. *Quaternions et Algèbres Géométriques, de nouveaux outils pour les images numériques couleur*. PhD thesis, Université de Poitiers, 12 2007. 49
- [Derrode 99] S. Derrode. *Représentation de formes planes à niveaux de gris par différentes approximations de Fourier-Mellin analytique en vue d'indexation de bases d'images*. PhD thesis, Université de Rennes 1, 1999. 26
- [Derrode 01] S. Derrode & F. Ghorbel. *Robust and Efficient Fourier-Mellin Transform Approximations for Gray-Level Image Reconstruction and Complete Invariant Description*. Computer Vision and Image Understanding, vol. 83, no. 1, pages 57–78, 2001. 25, 27, 70, 79, 85, 86, 92, 123
- [Dickson 19] L. E. Dickson. Annals of mathematics, volume 20, chapitre On quaternions and their generalization and the history of the eight square theorem, pages 155–171. 1919. 129
- [Dorst 02] L. Dorst, S. Mann & T. Bouma. *GABLE : A Matlab Tutorial for Geometric Algebra*, 2002. 57, 149
- [Dorst 07] L. Dorst, D. Fontijne & S. Mann. Geometric algebra for computer science : An object-oriented approach to geometry (the morgan kaufmann series in computer graphics). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007. 39
- [Duda 01] R. O. Duda, P. E. Hart & D. G. Stork. Pattern Classification (2nd Edition). Wiley-Interscience, 2 edition, November 2001. 11, 70, 101, 118
- [Ebling 05] J. Ebling & G. Scheuermann. *Clifford Fourier Transform on Vector Fields*. IEEE Transactions on Visualization and Computer Graphics, vol. 11, pages 469–479, 2005. 39, 110
- [Ell 92] T. A. Ell. *Hypercomplex spectral transformations*. PhD thesis, Minneapolis, MN, USA, 1992. 34, 37, 39, 40, 49
- [Ell 07] T. A. Ell & S. J. Sangwine. *Hypercomplex Fourier Transforms of Color Images*. IEEE Transactions on Image Processing, vol. 16, no. 1, pages 22–35, 2007. 11, 41, 42, 44, 48, 49, 104, 128, 129
- [Erturk 03] S. Erturk. *Digital image stabilization with sub-image phase correlation based global motion estimation*. IEEE Transactions on Consumer Electronics, vol. 49, no. 4, pages 1320 – 1325, nov. 2003. 110

- [Felsberg 02] M. Felsberg. *Low-Level Image Processing with the Structure Multivector*. PhD thesis, Inst. f. Informatik u. Prakt. Math. der Christian-Albrechts-Universität zu Kiel, 2002. 39
- [Flusser 09] J. Flusser, T. Suk & B. Zitova. *Moments and moment invariants in pattern recognition*. Wiley, 2009. 13
- [Fontijne 08] D. Fontijne. *Efficient algorithms for factorization and join of blades*. In Proceedings of the 3rd International Conference on Applications of Geometric Algebras in Computer Science and Engineering (AGACSE 2008), 2008. 55
- [Fourier 22] J.B.J. Fourier. *Théorie analytique de la chaleur*. Chez Firmin Didot, père et fils, 1822. 19
- [Gauthier 91] J.-P. Gauthier, G. Bornard & M. Silbermann. *Motions and pattern analysis : harmonic analysis on motion groups and their homogeneous spaces*. Systems, Man and Cybernetics, IEEE Transactions on, vol. 21, no. 1, pages 159 –172, jan/feb 1991. 12, 74, 80
- [Geusebroek 05] J. M. Geusebroek, G. J. Burghouts & A. W. M. Smeulders. *The Amsterdam library of object images*. Int. J. Comput. Vision, vol. 61, pages 103–112, 2005. 98
- [Ghorbel 93] F. Ghorbel. *Application de la transformée de Fourier généralisée au problème de l'invariance en reconnaissance de formes a niveaux de gris*. Groupe d'Etudes du Traitement du Signal et des Images (GRETSI) - Actes de Colloques, 1993. 18, 27, 70
- [Ghorbel 94] F. Ghorbel. *A complete invariant description for gray-level images by the harmonic analysis approach*. Pattern Recognition Letters, vol. 15, no. 10, pages 1043 – 1051, 1994. 74
- [Gonzalez 01] R. C. Gonzalez & R. E. Woods. *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2001. 11, 23, 26, 71, 72
- [Gourd 89] F. Gourd, J.-P.Gauthier & H. Younes. *Une méthode d'invariants de l'analyse harmonique en reconnaissance de formes*. Traitement du signal, vol. 6, pages 161–178, 1989. 31, 73
- [Grace 91] A. E. Grace & M. Spann. *A comparison between Fourier-Mellin descriptors and moment based features for invariant object recognition using neural networks*. Pattern Recognition Letters, vol. 12, no. 10, pages 635 – 643, 1991. 25
- [Guo 11] L.-Q. Guo & M. Zhu. *Quaternion Fourier-Mellin moments for color images*. Pattern Recognition, vol. 44, no. 2, pages 187 – 195, 2011. 45, 86, 101, 105
- [Hamilton 44] W. R. Hamilton. *On Quaternions, or on a New System of Imaginaries in Algebra*. Philosophical Magazine, vol. 25, no. 3, pages 489–495, 1844. 11

- [Helgason 01] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. American Mathematical Society, August 2001. 136
- [Hestenes 66] D. Hestenes. *Space-time algebra. Documents on modern physics*. Gordon and Breach, 1966. 39
- [Hestenes 84] D. Hestenes & G. Sobczyk. *Clifford algebra to geometric calculus*. Reidel, 1984. 11, 51, 52, 116, 133, 136
- [Hu 62] M.-K. Hu. *Visual pattern recognition by moment invariants*. IEEE Transactions on Information Theory, vol. 8, no. 2, pages 179–187, February 1962. 13
- [Isard 98] Michael Isard & Andrew Blake. *CONDENSATION - conditional density propagation for visual tracking*. International Journal of Computer Vision, vol. 29, pages 5–28, 1998. 124
- [Jain 97] A.K. Jain & D. Zongker. *Feature-Selection : Evaluation, Application, and Small Sample Performance*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19(2), pages 153–158, 1997. 102
- [Kakuta 97] M. Kakuta, T. Obi, M. Yamaguchi & N. Ohyama. *Optical implementation of color image correlation*. Optics Communications, vol. 142, no. 1-3, pages 19 – 25, 1997. 85, 111
- [Kovesi 03] P. Kovesi. *Phase Congruency Detects Corners and Edges*. In *Digital Image Computing : Techniques and Applications*, pages 309–318, 2003. 24, 125
- [Kunttu 05] I. Kunttu, L. Lepisto, J. Rauhamaa & A. Visa. *Color Fourier Descriptor for Defect Image Retrieval*. pages 415–422, 2005. 83, 90
- [Lounesto 01] P. Lounesto. *Clifford algebras and spinors*. Cambridge University Press, 2nd edition, May 2001. 42, 53, 115, 130, 132, 136
- [Lowe 99] D. G. Lowe. *Object Recognition from Local Scale-Invariant Features*. In *International Conference on Computer Vision 99*, pages 1150–1157, 1999. 13
- [Lu 93] Y. Lu, S. Schlosser & M. Janeczko. *Fourier descriptors and handwritten digit recognition*. Machine Vision and Applications, vol. 6, pages 25–34, 1993. 70
- [Mei 08] Y. Mei & D. Androustos. *Affine invariant shape descriptors : The ICA-Fourier descriptor and the PCA-Fourier descriptor*. pages 1–4, 2008. 71, 73
- [Mikolajczyk 05] K. Mikolajczyk & C. Schmid. *A performance evaluation of local descriptors*. IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 27, no. 10, pages 1615–1630, 2005. 13
- [Moxey 02] C. E. Moxey, T. A. Ell & S. J Sangwine. *Hypercomplex operators and vector correlation*. In *EUSIPCO 2002, Eleventh European Signal Processing Conference*, volume III, pages 247–250, 2002. 113

- [Moxey 03] C.E. Moxey, S.J. Sangwine & T.A. Ell. *Hypercomplex correlation techniques for vector images*. IEEE Transactions on Signal Processing, vol. 51, no. 7, pages 1941 – 1953, july 2003. 44
- [Moya-Sánchez 10] E. Moya-Sánchez & E. Bayro-Corrochano. *Quaternion Atomic Function Wavelet for Applications in Image Processing*. In Isabelle Bloch & Roberto Cesar, editeurs, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, volume 6419 of *Lecture Notes in Computer Science*, pages 346–353. Springer Berlin / Heidelberg, 2010. 38
- [Nene 96] S. A. Nene, S. K. Nayar & H. Murase. *Columbia Object Image Library (COIL-100)*, 1996. Technical Report CUCS-006-96. 27, 76, 98, 149
- [Phillips 98] P. J. Phillips, H. Wechsler, J. Huang & P. Rauss. *The FERET database and evaluation procedure for face recognition algorithms*. Image and Vision Computing, vol. 16, no. 5, pages 295–306, 1998. 98
- [Reddy 96] B.S. Reddy & B.N. Chatterji. *An FFT-based technique for translation, rotation, and scale-invariant image registration*. IEEE Transactions on Image Processing, vol. 5, no. 8, pages 1266 –1271, aug. 1996. 11, 110, 111
- [Rosten 10] E. Rosten, R. Porter & T. Drummond. *FASTER and better : A machine learning approach to corner detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pages 105–119, 2010. 13
- [Sangwine 96] S.J. Sangwine. *Fourier transforms of colour images using quaternion or hypercomplex numbers*. Electronics Letters, vol. 32, no. 21, pages 1979–1980, 1996. 39, 40
- [Sangwine 97] S. J. Sangwine. *The discrete quaternion Fourier transform*. Image Processing and Its Applications, Sixth International Conference on, vol. 2, pages 790–793, 1997. 40, 41
- [Sangwine 00] S. J. Sangwine & T. A. Ell. *The discrete Fourier transform of a colour image*. In J. M. Blackledge & M. J. Turner, editeurs, Image Processing II : Mathematical Methods, Algorithms and Applications, (Proceedings of Second IMA Conference on Image Processing, De Montfort University, Leicester, UK, September 1998), pages 430–441. Horwood Publishing for Institute of Mathematics and its Applications, 2000. 5, 9, 11, 12, 17, 18, 39, 40, 41, 49, 56, 63, 123
- [Sangwine 01] S. J. Sangwine, T. A. Ell & C. E. Moxey. *Vector phase correlation*. Electronics Letters, vol. 37, no. 25, pages 1513–1515, 2001. 112
- [Sangwine 05] S. J. Sangwine & N. Le Bihan. *QTFM : Quaternion Toolbox for Matlab*, 2005. 113, 114
- [Serra 09] J. Serra. *The "False Colour" Problem*. In Proceedings of the 9th International Symposium on Mathematical Morphology and Its Application to Signal and Image Processing, ISMM '09, pages 13–23, Berlin, Heidelberg, 2009. Springer-Verlag. 38

- [Sheng 86a] Y. Sheng & H. H. Arsenault. *Experiments on pattern recognition using invariant Fourier-Mellin descriptors*. The Journal of the Optical Society of America A, vol. 3, no. 6, pages 771–776, Jun 1986. 26, 78, 79
- [Sheng 86b] Y. Sheng & J. Duvernoy. *Circular-Fourier-radial-Mellin transform descriptors for pattern recognition*. The Journal of the Optical Society of America A, vol. 3, no. 6, pages 885–888, Jun 1986. 18, 25, 70, 74
- [Smach 07] F. Smach, J. Mitéran, M. Atri, J. Dubois, M. Abid & J.-P. Gauthier. *An FPGA-based accelerator for Fourier Descriptors computing for color object recognition using SVM*. Journal of Real-Time Image Processing, vol. 2, no. 4, pages 249–258, 2007. 80
- [Smach 08] F. Smach, C. Lemaître, J.-P. Gauthier, J. Mitéran & M. Atri. *Generalized Fourier Descriptors with Applications to Objects Recognition in SVM Context*. Journal of Mathematical Imaging and Vision, vol. 30, no. 1, pages 43–71, 2008. 11, 13, 22, 32, 69, 70, 74, 80, 81, 82, 84, 86, 87, 88, 99, 100, 105, 108, 123
- [Sommer 01] G. Sommer. *Geometric computing with clifford algebra : theoretical foundations and applications in computer vision and robotics*. Springer, 2001. 39
- [Teague 80] M. R. Teague. *Image analysis via the general theory of moments*. Journal of the Optical Society of America (1917-1983), vol. 70, pages 920–930, August 1980. 13
- [Torralba 03] A. Torralba & A. Oliva. *Statistics of natural image categories*. In Network : Computation in Neural Systems, pages 391–412, 2003. 23
- [Torres 99] L. Torres, J.Y. Reutter & L. Lorente. *The importance of the color information in face recognition*. In Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, volume 3, pages 627–631 vol.3, 1999. 84
- [Vapnik 98] V. N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, September 1998. 101
- [Vilenkin 68] N.J. Vilenkin. *Special functions and the theory of group representations*, volume 22 of *Translations of Mathematical Monographs*. American Mathematical Society, 1968. 12, 29, 30, 32, 123
- [Wu 05] G. Wu, E. Y. Chang & Z. Zhang. *An analysis of transformation on non-positive semidefinite similarity matrix for kernel machines*. In Proceedings of the 22nd International Conference on Machine Learning, 2005. 119
- [Zahn 72] C. T. Zahn & R. Z. Roskies. *Fourier Descriptors for Plane Closed Curves*. no. 3, pages 269–281, 1972. 70
- [Zernike 34] von F. Zernike. *Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode*. Physica, vol. 1, no. 7-12, pages 689 – 704, 1934. 13

- [Zhang 01] D. Zhang & G. Lu. *A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures*. Victoria, vol. 14, no. 1, pages 1–9, 2001. 72
- [Zhang 02a] D. Zhang & G. Lu. *Shape-based image retrieval using generic Fourier descriptor*. Signal Processing : Image Communication, vol. 17, no. 10, pages 825 – 848, 2002. 12, 74, 77, 85, 86
- [Zhang 02b] D. S. Zhang. *Image Retrieval Based on Shape*. PhD thesis, Monash University, 2002. 75
- [Zhang 05] D. Zhang & G. Lu. *Study and evaluation of different Fourier methods for image retrieval*. Image and Vision Computing, vol. 23, no. 1, pages 33–49, 2005. 73
- [Zickler 08] T. Zickler, S. P. Mallick, D. J. Kriegman & P. N. Belhumeur. *Color subspaces as photometric invariants*. International Journal of Computer Vision, vol. 79, pages 13–30, aug 2008. 69
- [Zwicke 83] P. E. Zwicke & I. Kiss. *A New Implementation of the Mellin Transform and its Application to Radar Classification of Ships*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-5, no. 2, pages 191 –199, 1983. 26

Méthodes fréquentielles pour la reconnaissance d'images couleur - Une approche par les algèbres de Clifford

Résumé : Dans cette thèse, nous nous intéressons à la reconnaissance d'images couleur à l'aide d'une nouvelle approche géométrique du domaine fréquentiel. La plupart des méthodes existantes ne traitent que les images en niveaux de gris au travers de descripteurs issus de la transformée de Fourier usuelle. L'extension de telles méthodes aux images multicanaux, comme par exemple les images couleur, consiste généralement à reproduire un traitement identique sur chacun des canaux. Afin d'éviter ce traitement marginal, nous étudions et mettons en perspective les différentes généralisations de la transformée de Fourier pour les images couleur. Ce travail nous oriente vers la transformée de Fourier Clifford pour les images couleur définie dans le cadre des algèbres géométriques. Une étude approfondie de celle-ci nous conduit à définir un algorithme de calcul rapide et à proposer une méthode de corrélation de phase pour les images couleur. Dans un deuxième temps, nous cherchons à généraliser à travers cette transformée de Fourier les définitions des descripteurs de Fourier de la littérature. Nous étudions ainsi les propriétés, notamment l'invariance à la translation, rotation et échelle, des descripteurs existants. Ce travail nous mène à proposer trois nouveaux descripteurs appelés "descripteurs de Fourier couleur généralisés" (*GCFD*) invariants en translation et en rotation.

Les méthodes proposées sont évaluées sur des bases d'images usuelles afin d'estimer l'apport du contenu fréquentiel couleur par rapport aux méthodes niveaux de gris et marginales. Les résultats obtenus à l'aide d'un classifieur *SVM* montrent le potentiel des méthodes proposées ; les descripteurs *GCFD* se révèlent être plus compacts, de complexité algorithmique moindre pour des performances de classification au minimum équivalentes. Nous proposons également des heuristiques pour le choix du paramètre de la transformée de Fourier Clifford.

Cette thèse constitue un premier pas vers une généralisation des méthodes fréquentielles aux images multicanaux.

Mots clés : Transformée de Fourier, Reconnaissance d'images, Descripteurs de Fourier couleur, Corrélation de phase couleur, Algèbre de Clifford, , Images couleur, Méthodes géométriques.

Frequency methods for color image recognition - An approach based on Clifford algebras

Summary :

In this thesis, we focus on color image recognition using a new geometric approach in the frequency domain. Most existing methods only process grayscale images through descriptors defined from the usual Fourier transform. The extension of these methods to multichannel images such as color images usually consists in reproducing the same processing for each channel. To avoid this marginal processing, we study and compare the different generalizations of color Fourier transforms. This work leads us to use the Clifford Fourier transform for color images defined in the framework of geometric algebra. A detailed study of it leads us to define a fast algorithm and to propose a phase correlation for color images. In a second step, with the aim of generalizing Fourier descriptors of the literature with this Fourier transform, we study their properties, including invariance to translation, rotation and scale. This work leads us to propose three new descriptors called "generalized color Fourier descriptors" (*GCFD*) invariant in translation and in rotation.

The proposed methods are evaluated on usual image databases to estimate the contribution of color frequency content compared with grayscale and marginal methods. The results obtained using an *SVM* classifier show the potential of the proposed methods ; the *GCFD* are more compact, have less computational complexity and give better recognition rates. We also propose heuristics for choosing the parameter of the color Clifford Fourier transform.

This thesis is a first step towards a generalization of frequency methods to multichannel images.

Keywords : Fourier transform, Image recognition, Color Fourier descriptors, Color phase correlation, Clifford algebra, Color images, Geometric methods.

Laboratoire Mathématiques,
Image et Applications
Avenue Michel Crépeau
17042 La Rochelle Cedex 01