

N° d'ordre : 2266

THÈSE

présentée

DEVANT L'UNIVERSITÉ DE RENNES I

pour obtenir

le grade de : **DOCTEUR DE L'UNIVERSITÉ DE RENNES I**

Mention : *Traitement du Signal et Télécommunications*

par

Stéphane DERRODE

Équipes d'accueil : *Groupe de Recherche Images et Formes (ENIC/INT)*
& *Dpt Image et Traitement de l'Information (ENST-Bretagne)*

École doctorale : *Informatique, Traitement du Signal et Télécommunications*

Composante universitaire : *Structure et Propriétés de la Matière*

Représentation de formes planes à niveaux de gris par différentes approximations de Fourier-Mellin analytique en vue d'indexation de bases d'images

soutenue le 15 décembre 1999 devant la commission d'examen :

Président : M. René Collorec, Professeur (Univ. de Rennes I)
Rapporteurs : M. Bruno Choquet, Docteur d'État (CNET/DIH)
M. Alain Le Méhauté, Professeur (Univ. de Nantes)
Examineurs : M. Christian Roux, Professeur (Enst Bretagne)
M. Pascal Haigron, Maître de conf. (Univ. de Rennes I)
Invité : M. Alain Hillion, Directeur scientifique (Enst Bretagne)
Directeur de thèse : M. Faouzi Ghorbel, Professeur (Univ. de Tunis II)

Résumé

CETTE thèse est consacrée à l'étude de la représentation globale des images planes à niveaux de gris dans le contexte de l'indexation de bases d'objets par le contenu. L'approche proposée dans ce mémoire est basée sur la description des objets par la transformée de Fourier-Mellin (TFM).

Pour des raisons de divergence de l'intégrale de Mellin, nous abordons, dans un premier temps, l'approximation de la TFM. Nous utilisons le prolongement analytique de la transformée de Fourier-Mellin (TFMA) pour proposer trois algorithmes originaux qui diffèrent par la manière dont les images sont ré-échantillonnées sur de nouvelles grilles et par les méthodes d'intégration adéquates pour ces grilles. Pour valider les algorithmes, nous examinons le problème inverse de la reconstruction d'une image depuis son spectre numérique.

Ensuite, un procédé original de normalisation de la TFMA est présenté pour l'extraction des familles complètes et intégrables de descripteurs invariants par similitudes. Une vraie distance sur l'espace des formes est utilisée pour mesurer la similarité des objets indépendamment de leur orientation et de leur taille. La performance de ces familles est évaluée par des expériences portant sur l'invariance numérique effective, la discrimination et la robustesse des descripteurs de formes.

Nous proposons également une méthode d'appariement permettant d'estimer le mouvement global entre deux objets de même forme. Celle-ci est fondée sur la distance euclidienne définie sur l'espace transformé de Fourier-Mellin et consiste à minimiser une fonction des paramètres de mouvement apparents entre les objets. Le minimum de cette fonction fournit également une mesure de similarité entre formes. Cela permet de déterminer si les paramètres estimés entre deux objets quelconques sont significatifs.

Ce mémoire présente enfin l'application des descripteurs de forme en vue de l'indexation de bases d'objets par le contenu. Les résultats de recherche sur deux bases d'objets à niveaux de gris montrent que ces descripteurs sont une alternative intéressante aux descripteurs de formes globaux utilisés dans les moteurs de recherche actuels.

Mots clés : Transformée de Fourier-Mellin, Approximation, Images à niveaux de gris, Reconnaissance de formes, Descripteurs invariants, Similitudes, Complétude, Distance entre formes, Estimation de mouvement, Indexation.

Abstract

THIS thesis is devoted to gray-level shape representation in the context of content-based image retrieval. The approach we used in this dissertation is based on the description of objects by the Fourier-Mellin Transform (FMT).

Due to the divergence of the Mellin transform, we are first interested in the approximation of the FMT. From the Analytical extension of the Fourier-Mellin Transform (AFMT), we propose three original algorithms. They differ in re-sampling methods of the discrete image on new grids and in numerical integration performed on these grids. To validate the algorithms, we examine the inverse problem of reconstructing an image from its numerical spectrum.

Then, a new process for the normalisation of the AFMT is used to derive complete and integrable sets of features which are invariant under planar similarities. A true distance between shapes is then used to measure the similarity between objects independently of their orientation and scale. The performances of the new sets are analyzed with experiments on the numerical invariancy, the discrimination and the robustness of the shape features.

Besides, we propose a matching process to estimate the global motion parameters between objects with similar shape. The process is based on the Euclidean distance expressed in the Fourier-Mellin domain. It consists in minimizing a function that depends on apparent motion parameters between objects. The minimum of the function also gives a similarity measure between the shape of the objects. This can be used to determine whether or not the estimated motion parameters are significant.

Finally, we apply the gray-level shape descriptors for content-based retrieval and image database indexation. Experimental results on two object databases show that these features are more efficient than those used in existing retrieval engines.

Key words : Fourier-Mellin transform, Approximation, Gray-level images, Pattern recognition, Invariant, Similarity, Completeness, Distance between shapes, Motion estimation, Content-based retrieval.

Remerciements

Ce travail de thèse a été réalisé au sein du *Groupe de Recherche Images et Formes* (GRIF) de l'École Nouvelle d'Ingénieurs en Communication et du *département Image et Traitement de l'Information* de l'École Nationale Supérieure des Télécommunications de Bretagne, grâce au soutien du centre de recherche et développement de France Télécom (CNET).

Je tiens tout d'abord à adresser mes plus vifs remerciements à mon directeur de thèse et encadrant, M. GHORBEL, directeur des Études et Recherches du *Centre des Études et Recherche des Télécommunications* (Tunisie Télécom) et responsable du *Groupe de Recherche Images et Formes de Tunisie* (ENSI). Il a su m'épauler, aux moments opportuns, par ses conseils, ses réflexions et son enthousiasme, dans l'avancement de ce travail. Son influence sur ce document n'est ni mesurable ni quantifiable (contrairement aux descripteurs de formes !) et je désire par ces quelques mots lui témoigner toute mon amitié et ma reconnaissance.

Mes remerciements vont également à M. COLLOREC, Professeur à l'université de Rennes 1, pour l'honneur qu'il me fait en acceptant de présider le jury de cette thèse. Je remercie tout particulièrement M. LE MÉHAUTÉ, Professeur à l'université de Nantes, et M. CHOQUET, responsable Innovation, Prospective et Veille (CNET/DIH), pour l'intérêt qu'ils ont porté à ce travail en acceptant d'en être les rapporteurs. Je remercie également M. ROUX, Professeur à l'ENST Bretagne et responsable du département ITI, M. HILLION, Directeur scientifique de l'ENST Bretagne, ainsi que M. HAIGRON, Maître de conférence à l'université de Rennes 1, pour les questions et remarques judicieuses et pour l'intérêt qu'ils ont manifesté pour ce travail.

Je suis reconnaissant envers MM. P. GIORGINI et J.C. DAMIEN, Directeurs successifs de l'ENIC, de m'avoir fait confiance en me permettant de réaliser cette thèse au sein de l'école qui m'a formé aux domaines des télécommunications. Parmi les membres de l'équipe GRIF, je tiens à remercier plus particulièrement Lamia, Stanislaw, Mohamed, Ahmed et Mourad pour avoir partagé avec eux durant trois années *la longue pénitence de la thèse ...*

C'est également pour son accueil au sein du département ITI que j'adresse mes remerciements les plus sincères à C. Roux. J'ai apprécié de réaliser la rédaction de ce manuscrit dans de très bonnes conditions, et je salue tous les membres de l'équipe parmi lesquels Jean-Michel, Laurent, Grégoire, Mathieu, Renaud, Jean-Marc, Laurent et Sylvain qui ont contribué par leur compétence, leur patience et leur bonne humeur au bon déroulement mon séjour en Bretagne. Je n'oublie pas Martin, Patrick et Olivier pour les bons moments passés ensemble.

C'est également avec un très grand plaisir que je remercie les équipes qui m'ont invité au cours de plusieurs séjours et ont participé ainsi à l'enrichissement de mon travail.

Je pense en particulier à Olivier Avaro (actuellement à Deutsche Telekom - Berkom) et Gérard Eude (direction scientifique du CNET) pour leur accueil pendant cinq mois au sein du service *Communication de Groupe et Visiophonie* (DSE/SGV).

Mes pensées vont également aux membres du GRIFT de l'ENSI et du groupe *Projet en Imagerie Numérique et Applications* (PINA) du CERT pour leur accueil chaleureux à l'occasion de mes séjours en Tunisie. Je salue en particulier Rim, Fatma, Leila, Najla, Asma, Tarak, Moez et Moez, *vivement un TAIMA bis à Hammamet ...*

Je veux par ailleurs remercier R. LENZ, Professeur à l'université de Linköping (Suède), pour son accueil durant un mois au sein de son groupe de recherche. Qu'il trouve ici une contribution de l'application de la théorie des groupes en analyse d'images.

Table des matières

Résumé	i
Abstract	iii
Remerciements	v
Introduction	1
1 Représentations et Descriptions Globales des Objets 2D	5
1.1 Les transformées de Fourier des fonctions	6
1.1.1 Rappels préliminaires sur la transformée de Fourier	6
1.1.2 Extension de la transformée de Fourier	8
1.1.3 Les TF des groupes de transformations d'une variable	10
1.1.3.1 Les coefficients de Fourier	10
1.1.3.2 La transformée de Fourier discrète	13
1.1.3.3 La transformée de Mellin	13
1.2 La représentation des images par les coefficients de Fourier	15
1.2.1 Les fonctions harmoniques circulaires	15
1.2.2 Les moments des images à niveaux de gris	16
1.2.2.1 Les moments complexes	16
1.2.2.2 Les moments de Zernike	17
1.2.2.3 Les moments de Fourier-Mellin	19
1.3 La transformée de Fourier-Mellin	19
1.3.1 La transformée de Fourier-Mellin standard	20
1.3.2 Existence de la TFM pour des images	21
1.4 Conclusion du chapitre	22
2 Approximations de la TFM et Reconstruction des Images	23

2.1	La transformée de Fourier-Mellin analytique	24
2.1.1	La TFM analytique des objets à niveaux de gris	24
2.1.2	Quelques résultats sur la TFMA	25
2.2	Trois méthodes d'approximations de la TFMA	27
2.2.1	Fonctions et spectres discrets	28
2.2.2	Approximation directe par échantillonnage polaire	29
2.2.3	Approximation rapide par échantillonnage log-polaire	32
2.2.4	Approximation cartésienne sans ré-échantillonnage préalable	34
2.3	Reconstruction des objets : comparaison des approximations	35
2.3.1	Expériences sur l'image d'une lettre binaire	36
2.3.2	Expériences sur l'image d'un papillon	38
2.4	Conclusion du chapitre	43
3	Reconnaissance Invariante des Formes par Similitudes	45
3.1	Unicité des descripteurs invariants	46
3.2	Représentation des formes par la TFMA	47
3.2.1	L'égalité de Parseval sur le groupe des similitudes	48
3.2.2	Action du groupe des similitudes sur la TFMA	48
3.2.3	Caractérisation des formes et interprétation	49
3.3	Extraction de nouvelles familles complètes de descripteurs invariants	52
3.3.1	Description du procédé de normalisation	52
3.3.1.1	Principe de normalisation de la TFMA	52
3.3.1.2	Unicité et sommabilité des invariants - Distance entre formes	53
3.3.2	Premier exemple de paramètres de normalisation	55
3.3.2.1	Famille initiale	55
3.3.2.2	Les descripteurs invariants de l'image d'un demi-disque	56
3.3.3	Interprétation des paramètres : objet canonique d'une forme	57
3.4	Extension des paramètres et représentation numérique des formes	59
3.4.1	Robustesse des paramètres de normalisation	59
3.4.2	Extension de la famille initiale	61
3.4.3	Représentation numérique des formes et objets canoniques	62
3.5	Reconnaissance de formes et approximations de la TFMA	64
3.5.1	Images de test et présentation des résultats	64
3.5.2	Distance entre des objets similaires	65
3.5.3	Évaluation de la discrimination des deux familles	67
3.5.4	Robustesse des descripteurs pour des formes presque similaires	69

3.6	Conclusion du chapitre	73
4	Estimation du Mouvement Global entre Objets	75
4.1	Appariement d'objets et estimation des paramètres de mouvement	76
4.1.1	Principe de la méthode d'appariement	77
4.1.1.1	Fonction d'appariement	77
4.1.1.2	Mesure de correspondance et estimation des paramètres de mouvement	78
4.1.1.3	Exemple de deux demi-disques	79
4.1.2	Optimisation de la fonction $\mathbf{E}(\rho, \psi)$	80
4.1.2.1	Appariement d'objets similaires et d'objets différents	80
4.1.2.2	Estimation de l'écart entre formes	80
4.2	Sous-groupe des rotations vectorielles et distance de Hausdorff	84
4.2.1	Rappels sur la distance de Hausdorff	84
4.2.2	Expression de la distance de Hausdorff sur l'espace des formes	85
4.2.3	Application à la détection des symétries d'une image par rotation	86
4.2.3.1	La TFMA des objets symétriques	86
4.2.3.2	Algorithme de détection des axes de symétrie d'une image	88
4.3	Approximations de la TFMA et estimation des paramètres de mouvement	89
4.3.1	Estimation de mouvement entre des objets similaires	90
4.3.2	Évaluation de la discrimination de l'écart	92
4.3.3	Robustesse de l'écart aux faibles déformations	93
4.4	Conclusion du chapitre	93
5	Application à la Recherche d'Images en vue d'Indexation de Bases par le Contenu	95
5.1	Principes de la recherche d'images par le contenu	96
5.1.1	Architecture d'un moteur de recherche	96
5.1.2	Description du contenu des images	97
5.1.3	Indexation des bases d'images	98
5.1.4	Coopération des descripteurs	98
5.2	Résultats expérimentaux de recherche par le contenu	99
5.2.1	Algorithme de recherche et bases de test	100
5.2.2	Principe du système de recherche et interface utilisateur	101
5.2.3	Quelques résultats pour les trois méthodes d'approximation de la TFMA	103
5.2.4	Comparaison avec des descripteurs des moments	106
5.2.4.1	Base de Papillons	106

5.2.4.2	Base de Columbia	108
5.3	Conclusion du chapitre	112
Conclusion générale		113
A Petite biographie sur J. Fourier et H. Mellin		117
A.1	Joseph Fourier	117
A.2	Hjalmar Mellin	118
B Les moments géométriques et les invariants de Hu		119
C Bases d'objets utilisées pour les tests		121
C.1	Base de papillons	121
C.2	Bases d'objets de l'université de Columbia	122
Liste des figures		123
Liste des tableaux		127
Bibliographie		129

Introduction

AVEC le développement des applications de l'analyse d'images encouragé par la mise en service de nouveaux réseaux (Internet, ATM), par l'évolution des capacités des ordinateurs (mémoire de masse et puissance) et par la baisse du prix des équipements matériels, nous observons un regain d'intérêt important pour la description de scènes. Les applications touchent différents domaines tels que les télécommunications (codage, indexation multimédia, tatouage), l'imagerie médicale (morphométrie des organes, indexation), ou bien encore la télédétection (analyse des fonds marins, cartographie pour l'aménagement du territoire). Différentes manières de représenter les scènes ont été abordées, qu'il s'agisse de la description des contours, des images planes ou des images tridimensionnelles (surfaces et volumes). Nous nous intéressons au cas des images planes à niveaux de gris. La représentation globale et la description du mouvement d'objets 2D à niveaux de gris est un problème d'un grand intérêt. Il est cependant reconnu qu'il présente des difficultés majeures.

L'indexation de bases d'images est un nouveau champ d'application pour les méthodes de description des formes. Cela fait l'objet d'une attention particulière par le groupe MPEG dans le cadre de l'établissement de la norme MPEG7. Une autre application en pleine expansion dans le domaine des télécommunications concerne les nouvelles tendances en codage audiovisuel avec la norme MPEG4 (télévision numérique interactive, échange de flux multimédia à travers les réseaux). D'autres applications visent les domaines plus traditionnels de la vision artificielle, l'imagerie médicale, la télédétection et l'analyse de documents, par exemple.

L'étude présentée dans ce mémoire a été initiée par une convention de recherche entre le CNET et l'ENIC/INT portant sur l'application des descripteurs de formes pour la représentation des objets multimédia en vue d'applications en codage et en indexation. La première partie de cette action a porté sur la représentation des formes contours et son application au codage de séquences d'images [88, 30]. La seconde partie, qui est le prolongement de ce dernier travail, traite d'un cas plus général qui est celui de la représentation globale des objets à niveaux de gris.

La perception et la compréhension des formes des objets d'une scène est une tâche automatique chez l'homme qui la réalise de manière intelligente et avec un haut niveau hiérarchique de représentation. Dans la communauté scientifique de l'imagerie, il a été montré que de telles tâches nécessitent une caractérisation précise des objets permettant de dégager deux notions, à savoir [56] :

- La forme des objets,
- Les transformations géométriques qui opèrent globalement sur les objets.

C'est pourquoi, nous avons généralement recours à une transcription de la représentation originale (matrice de pixels) sous la forme de descripteurs invariants par des transformations

géométriques, ces dernières décrivant le mouvement global des objets dans l'image. Il existe de nombreux outils pour décrire les formes, ceux-ci dépendent de la nature des objets étudiés :

- Les contours plans : les descripteurs de Fourier [135, 54], les points d'intérêts [51], les courbures [128], la géométrie discrète [25].
- Les images planes à niveaux de gris : les moments de l'image [86, 101], les points d'intérêts [135], la transformée de Fourier-Mellin [115, 55], la M-transformée [52, 50].
- Les volumes 3D à niveaux de gris : les moments [109], l'analyse harmonique [136].
- Les surfaces fermées 3D : géométrie différentielle [44, 80, 15], moments surfaciques [132].

Comme nous traitons le cas des objets 2D à niveaux de gris, un outil adapté est la transformée de Fourier-Mellin (TFM). Celle-ci permet de décrire les formes indépendamment des translations, rotations et facteurs d'échelle [114]. Malheureusement, elle présente des limitations liées à son approximation [137]. Celles-ci ont largement déprécié son utilisation en analyse de scènes par rapport à des méthodes classiques telles que les moments de l'image. Le travail traité dans ce mémoire s'articule autour de quatre grands thèmes qui sont :

1. L'approximation numérique de la transformée de Fourier-Mellin pour des images à niveaux de gris.
2. La représentation des formes et leur description invariante par similitude issue de la TFM.
3. L'estimation des paramètres du mouvement apparents entre deux objets de même forme.
4. L'application à l'indexation de bases d'objets à niveaux de gris pour la recherche d'images par le contenu.

Ce mémoire s'organise donc de la manière suivante

L'objectif du **chapitre 1** est d'aboutir à la définition rigoureuse de la transformée de Fourier-Mellin pour la description globale des objets à niveaux de gris. Du point de vue de la théorie des groupes et de l'analyse harmonique, celle-ci se présente comme la transformée de Fourier sur le groupe des similitudes planes. Plutôt que de développer les aspects théoriques, nous adopterons une démarche progressive permettant d'intégrer les principales méthodes de description globale que l'on rencontre dans la littérature. Ce chapitre met finalement en évidence les problèmes numériques cruciaux rencontrés lors de l'approximation de la transformée de Fourier-Mellin.

Avant d'envisager les applications de la transformée de Fourier-Mellin pour la représentation des objets à niveaux de gris, il est important de proposer des approximations robustes de la TFM. Dans le **chapitre 2**, nous utiliserons le prolongement analytique de la TFM (TFMA) comme solution rigoureuse au problème de divergence des intégrales. Nous développons ainsi trois méthodes d'approximation de la TFM et de la transformée inverse. Ces algorithmes se distinguent par la manière dont les images discrètes sont ré-échantillonnées sur de nouvelles grilles et par la méthode d'intégration numérique adéquate sur ces grilles. Pour valider et comparer ces trois méthodes, nous examinerons le problème inverse de la reconstruction d'une image à niveaux de gris depuis son spectre discret.

La présentation d'un procédé original de normalisation de la TFMA permettant d'extraire des familles complètes et intégrables de descripteurs invariants par similitudes planes directes

fera l'objet du **chapitre 3**. Une vraie distance entre formes est utilisée pour mesurer le degré de similarité des objets indépendamment de leur position, leur orientation et leur taille dans les images. La performance de ces familles au regard de trois méthodes d'approximations de la TFMA est évaluée et illustrée par des expériences de reconnaissance de formes portant sur l'invariance numérique effective aux similitudes planes, la discrimination des formes et la robustesse des descripteurs vis-à-vis des faibles déformations de formes.

Le **chapitre 4** présentera une seconde application de la TFMA pour l'estimation des paramètres de mouvement globaux entre deux objets de même forme. L'algorithme de correspondance que nous proposons est fondé sur la distance euclidienne exprimée sur l'espace de représentation de Fourier-Mellin et consiste à minimiser une fonction des paramètres de mouvement apparents entre les objets. Le minimum de cette fonction fournit également une mesure de similarité entre formes, ce qui permet de déterminer si les paramètres estimés entre deux objets quelconques ont un sens. Enfin, les expériences proposées visent à tester la robustesse des trois méthodes d'approximation de la TFMA quant aux problèmes d'isotropie de la grille cartésienne vis-à-vis des similitudes.

Finalement, nous présenterons dans le **chapitre 5** l'application des familles complètes de descripteurs invariants pour la recherche d'images par le contenu en vue d'indexer des bases d'objets à niveaux de gris. Nous présenterons les résultats de recherche effectués sur une base originale d'images de papillons et sur la base d'objets de l'université de Columbia. Nous confronterons les résultats, d'une part, avec les invariants de Hu, que l'on retrouve dans la grande majorité des systèmes d'indexation actuels, et d'autre part, avec une nouvelle famille complète de descripteurs invariants par similitudes planes issue de la représentation des images par les moments complexes.

Représentations et Descriptions Globales des Objets 2D

Le principal objectif de ce chapitre est d'aboutir à la définition rigoureuse de la transformée de Fourier-Mellin pour la description et l'analyse des objets à niveaux de gris. Du point de vue de la théorie des groupes et de l'analyse harmonique, celle-ci se présente comme la transformée de Fourier sur le groupe des similitudes planes vectorielles. Plutôt que de présenter les aspects théoriques, nous avons préféré adopter une démarche progressive scindée en trois étapes permettant d'intégrer plusieurs méthodes classiques de description globale et de reconnaissance de formes 2D.

Nous commençons par rappeler quelques propriétés intéressantes de la transformée de Fourier pour la description des fonctions définies sur \mathbb{R} . Dans le cadre de l'analyse harmonique, ces résultats s'étendent notamment aux signaux numériques avec la transformée de Fourier discrète, aux signaux périodiques avec les coefficients de Fourier et aux fonctions causales avec la transformée de Mellin (section 1.1).

Ces trois outils sont à l'origine de plusieurs méthodes de description globale des objets à niveaux de gris que nous présentons dans la section 1.2. Les fonctions harmoniques circulaires, comme application des coefficients de Fourier aux images, permettent d'étudier les objets 2D en rotation dans le plan. En combinant cette représentation avec des moments radiaux, nous aboutissons aux méthodes classiques de représentation des images par les moments complexes et les moments de Zernike. Ces techniques ont été très largement étudiées en reconnaissance de formes.

Finalement, la section 1.3 aboutit à la présentation de la transformée de Fourier-Mellin qui est à l'origine des développements en reconnaissance de formes présentés dans ce mémoire. Celle-ci peut être vue comme l'association des fonctions circulaires harmoniques de l'image et de la transformée de Mellin radiale.

1.1 Les transformées de Fourier des fonctions

La transformée de Fourier, et ses voisines telles que les séries de Fourier, les coefficients de Fourier et la transformée de Fourier discrète, sont parmi les outils de base les plus utilisés pour l'analyse des signaux mono-dimensionnels. La transformée de Fourier fournit une représentation fréquentielle unique et inversible qui permet d'accéder à certaines caractéristiques importantes d'un signal (bande-passante, composante continue, ...) qu'il est plus difficile d'extraire depuis la représentation initiale. Dans cette section, nous rappelons certaines propriétés que nous serons amenés à étendre, dans les sections suivantes, pour la représentation et la description des objets à niveaux de gris.

1.1.1 Rappels préliminaires sur la transformée de Fourier

Parmi les propriétés de la transformée de Fourier, l'opérateur à décalage de temps permet d'extraire certaines primitives du signal, indépendamment du système de coordonnées dans lequel il est décrit. Nous rappelons également, dans ce paragraphe, l'expression de l'identité de Parseval et le lien entre la transformée de Fourier et l'opérateur de corrélation utilisé pour comparer deux signaux décalés dans le temps.

La transformée de Fourier sur \mathbb{R}

Considérons une fonction mono-dimensionnelle $f(x)$ définie et à valeurs dans \mathbb{R} , intégrable par rapport à la mesure de Lebesgue. Sa *Transformée de Fourier* (TF) est donnée par :

$$\forall u \in \mathbb{R}, \quad \mathcal{F}_f(u) = \int_{-\infty}^{+\infty} f(x) e^{-iux} dx. \quad (1.1)$$

Cette représentation est unique et inversible. Il est, en effet, possible de reconstruire complètement un signal depuis sa représentation de Fourier, en utilisant la transformée inverse \mathcal{F}^{-1} :

$$\forall x \in \mathbb{R}, \quad f(x) = \mathcal{F}^{-1}(\mathcal{F}_f(u))(x) = \int_{-\infty}^{+\infty} \mathcal{F}_f(u) e^{iux} du. \quad (1.2)$$

En supposant que f est de carré sommable sur \mathbb{R} , une relation importante est l'*identité de Parseval* :

$$\int_{-\infty}^{+\infty} \|\mathcal{F}_f(u)\|^2 du = \int_{-\infty}^{+\infty} |f(x)|^2 dx. \quad (1.3)$$

La transformée de Fourier est un isomorphisme sur l'espace vectoriel $\mathbf{L}^2(\mathbb{R})$ ¹. Cette identité est connue sous le nom de *théorème de conservation de l'énergie* en traitement du signal.

Le théorème du retard

Soit maintenant la fonction f_a obtenue par translation de f , i.e. $f_a(x) = f(x + a)$. Sa transformée de Fourier s'écrit :

$$\mathcal{F}_{f_a}(u) = \int_{-\infty}^{+\infty} f(x + a) e^{-iux} dx.$$

¹— $\mathbf{L}^2(\mathbb{R})$ désigne l'espace vectoriel normé des fonctions de carré intégrables par rapport à la mesure de Lebesgue dx .

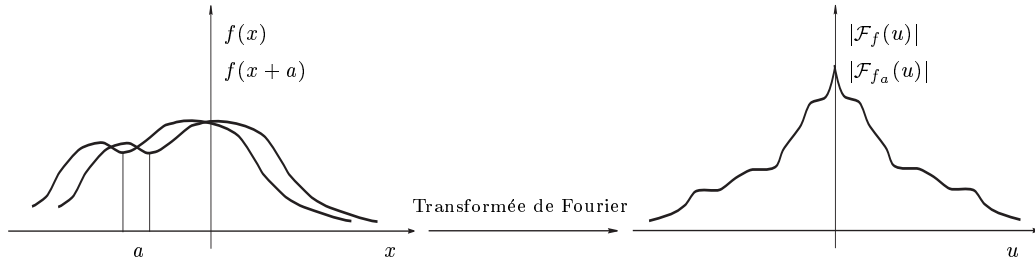


Figure 1.1 — La transformée de Fourier des fonctions à support sur \mathbb{R} et illustration théorème du retard. À gauche : Le signal original et son translaté. À droite : le module de la transformée de Fourier est identique pour les deux signaux.

En effectuant le changement de variable $y = x + a$, nous obtenons :

$$\begin{aligned}
 \mathcal{F}_{f_a}(u) &= \int_{-\infty}^{+\infty} f(y) e^{-iu(y-a)} dy \\
 &= e^{iua} \int_{-\infty}^{+\infty} f(y) e^{-iuy} dy \\
 &= e^{iua} \mathcal{F}_f(u).
 \end{aligned} \tag{1.4}$$

Ainsi une translation dans le domaine original produit un déphasage dans le domaine de Fourier. Cette relation est appelée *théorème à décalage de temps* ou bien encore *théorème du retard* de la transformée de Fourier. Son importance découle du résultat suivant : si on décompose en module et en phase l'expression précédente,

$$\forall u \in \mathbb{R}, \quad \begin{cases} |\mathcal{F}_f(u)| = |\mathcal{F}_{f_a}(u)|, \\ \arg \mathcal{F}_f(u) = \arg \mathcal{F}_{f_a}(u) - ua, \end{cases}$$

on constate que :

- D'une part, l'amplitude de la TF est invariante sous l'action des translations. On peut donc extraire un ensemble de descripteurs caractérisant la forme d'un signal indépendamment des translations qu'il a subit (cf. schéma 1.1). Cependant, l'information contenue dans la phase de la TF est perdue, et il est clair que deux signaux de formes différentes peuvent avoir les mêmes descripteurs. Ainsi, cette caractérisation n'est pas unique et peut engendrer des confusions entre les formes.
- D'autre part, la translation est codée dans la phase de chaque composante harmonique. Ceci donne un moyen rapide pour estimer le paramètre de translation entre deux signaux dont on sait, a priori, qu'ils ont la même forme.

La corrélation par translation

Une autre méthode bien connue pour comparer deux signaux, indépendamment d'un certain décalage dans le temps, consiste à utiliser la technique de corrélation. Celle-ci permet de s'affranchir de l'étape d'extraction de descripteurs invariants et donne, en plus, une mesure sur la position relative des deux signaux. La corrélation \mathcal{C}_{fg} en a entre deux fonctions f et g

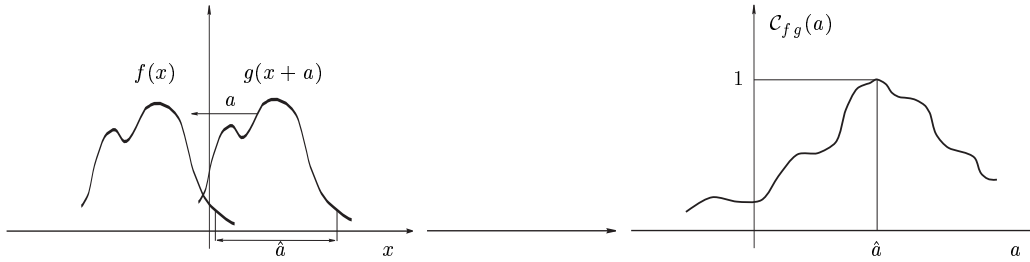


Figure 1.2 — La corrélation des fonctions à support sur \mathbb{R} . À gauche : le signal original f et le signal à appairer g_a ; à droite : la fonction de corrélation entre les deux signaux.

est donnée par :

$$\mathcal{C}_{fg}(a) = (f \star g)(a) = \int_{-\infty}^{+\infty} f(x) g_a(x) dx, \quad (1.5)$$

en notant $g_a(x) = g(x+a)$, avec a le paramètre de translation.

En parcourant l'ensemble des translations possibles, on dresse la fonction de corrélation. Celle-ci atteint la valeur maximale pour le paramètre \hat{a} de la translation qui apparie « au mieux » les deux signaux. La corrélation fournit une mesure de la position relative entre f et g . Si, de plus, les fonctions sont normalisées par leur énergie moyenne, le maximum de corrélation atteint la valeur 1 pour deux signaux de même forme. Ainsi, la corrélation peut être considérée comme une méthode de reconnaissance de formes (cf. figure 1.2). Il est bien connu que la corrélation est une technique robuste en présence d'un bruit additif. Néanmoins, cette technique s'avère coûteuse puisqu'il est nécessaire de réitérer l'opération de manière systématique pour toutes les translations possibles a entre l'image f et le modèle g .

La relation entre la corrélation et la transformée de Fourier sur \mathbb{R} est connu sous le nom de *Théorème de Convolution*. En calculant la transformée de Fourier de l'équation (1.5) et en appliquant le théorème du retard (1.4) sur le modèle, on montre que la fonction de corrélation est la transformée de Fourier inverse du produit suivant :

$$\mathcal{C}_{fg}(a) = \mathcal{F}^{-1}(\overline{\mathcal{F}_f} \mathcal{F}_{g_a}) = \int_{-\infty}^{+\infty} e^{iau} \overline{\mathcal{F}_f(u)} \mathcal{F}_g(u) du, \quad (1.6)$$

où la notation \bar{c} désigne le conjugué du nombre complexe c . L'unicité de la transformée de Fourier assure qu'aucun faux appariement ne peut être enregistré. Moyennant certaines hypothèses sur les signaux, des techniques de calcul rapide de la TF peuvent être appliquées pour réduire les temps de calcul. On peut également s'affranchir de la confrontation systématique entre le signal et le modèle par des techniques d'optimisation qui permettent de converger rapidement vers le maximum global de corrélation [100, chapitre 10].

1.1.2 Extension de la transformée de Fourier

L'importance de la *Théorie des Groupes* en traitement d'images est largement reconnue depuis une vingtaine d'années. Son intérêt pour la reconnaissance de formes est évident : il est important de reconnaître des images indépendamment de la pose des objets, des paramètres

de calibration de la caméra, des conditions d'éclairage, ... La plupart de ces problèmes ont une structure commune qui peut être décrite par des groupes de transformations. Plusieurs travaux ont été récemment publiés, dans lesquels des méthodes de la théorie de la représentation des groupes sont utilisées pour résoudre des problèmes en vision et en traitement d'images. Pour une introduction plus complète et acquérir une vue d'ensemble, le lecteur pourra consulter les ouvrages suivants [96, 70, 75, 89]. Un état de l'art des applications de la théorie des groupes de Lie en reconnaissance de formes peut être consulté dans [45].

Dans le cadre plus restreint des *Groupes Localement Compacts et Abéliens* (GLCA, noté \mathcal{G}), la *théorie de l'analyse harmonique commutative* fournit un outil particulièrement intéressant, la *transformée de Fourier sur \mathcal{G}* , comme extension de la transformée de Fourier sur le groupe des translations $(\mathbb{R}, +)$. Sans entrer dans le formalisme mathématique de l'analyse harmonique², la transformée de Fourier sur un GLCA se construit à l'aide des représentations unitaires et irréductibles de \mathcal{G} (le noyau de la transformée) et de la mesure normalisée, invariante et positive associée (mesure de Haar). Si nous reprenons le cas simple de la transformée de Fourier classique, la mesure de Haar devient la mesure de Lebesgue dx , et la représentation du groupe est donnée par le noyau exponentiel complexe $\{e^{iux}\}$ avec $u \in \mathbb{R}$. Cet exemple met en évidence deux droites réelles, une droite pour la fonction (paramètre x) et une autre droite pour la transformée de Fourier (paramètre u), chacune d'entre elles étant le groupe dual de l'autre.

Les concepts de base et les théorèmes évoqués pour la transformée de Fourier peuvent être transférés à ce contexte général, notamment :

- L'ensemble des fonctions admissibles pour lequel la transformée de Fourier sur un groupe \mathcal{G} existe constitue un espace vectoriel normé des fonctions intégrables sur \mathcal{G} , i.e. $\mathbf{L}^1(\mathcal{G}, d_{\mathcal{G}})$ où $d_{\mathcal{G}}$ est la mesure de Haar du groupe.
- L'opérateur à décalage pour le groupe de transformations \mathcal{G} est une extension du théorème du retard pour le groupe \mathbb{R} .
- Une transformée inverse qui est définie comme la transformée de Fourier sur le groupe dual $\hat{\mathcal{G}}$.
- Le théorème de Plancherel-Parseval qui assure que la transformée de Fourier est une isométrie pour la classe des fonctions de carré intégrable $\mathbf{L}^2(\mathcal{G}, d_{\mathcal{G}})$.

Depuis une vingtaine d'années, des développements théoriques fondés sur l'analyse harmonique ont été réalisés dans le domaine de la reconnaissance de formes [12, 52, 112, 55, 56]. Le principal intérêt de la transformée de Fourier provient de sa faculté à dissocier l'information sur la forme d'un objet (contours ou objets plan dans le cas 2D, surfaces ou volumes dans le cas 3D) de sa position relative dans le plan ou l'espace. Par des traitements appropriés, elle permet alors d'extraire des descripteurs de formes invariants par rapport à certaines transformations géométriques, parmi lesquelles les translations, rotations, dilatations, étirements, et certaines combinaisons de ces groupes.

Dans ce contexte théorique, la transformée de Fourier-Mellin est la transformée de Fourier sur le groupe des similitudes planes vectorielles, produit direct du groupe des rotations par le groupe des dilatations du plan. Elle est particulièrement bien adaptée pour l'étude des objets

²— Les lecteurs intéressés par ces aspects théoriques peuvent consulter les ouvrages mathématiques de W. Rudin [106] et J. Dieudonné [41] ainsi que les travaux de F. Ghorbel [55, 56] pour l'application de l'analyse harmonique à la reconnaissance de formes.

à niveaux de gris et, en particulier, pour l'extraction de descripteurs de formes invariants par rapport à ce groupe. La transformée de Fourier-Mellin est la composée des coefficients de Fourier et de la transformée de Mellin. Ces transformées de Fourier sont définies pour des groupes de transformations d'une variable, nous commençons par les présenter individuellement.

1.1.3 Les TF des groupes de transformations d'une variable

Dans ce paragraphe, nous présentons trois transformées de Fourier de fonctions définies sur un espace \mathbf{X} homogène au groupe de transformations \mathcal{G} , *i.e.* $\mathbf{X} = \mathcal{G}$. Ces exemples nous seront utiles dans les sections suivantes. Il s'agit des coefficients de Fourier pour le groupe des rotations, de la transformée de Fourier discrète pour le groupe des translations discrètes et de la transformée de Mellin pour le groupe des dilatations. Tous ces groupes de transformations sont des groupes topologiques, compacts ou localement compacts, et commutatifs. Nous mettons l'accent sur l'existence de ces transformées, leur transformée inverse et le théorème à décalage approprié pour chacun des cas présentés. Ces transformées sont illustrées dans la figure 1.4 page 12.

1.1.3.1 Les coefficients de Fourier

Considérons une fonction mono-dimensionnelle $f(\theta)$ 2π -périodique. Le paramètre θ représente généralement un angle de rotation et le groupe de transformations qui agit sur ces fonctions peut être assimilé à $\mathbb{S}^1 = [0; 2\pi[$. La transformée de Fourier adaptée à ce type de fonction est donnée par les *Coefficients de Fourier* (CF) de f :

$$\forall k \in \mathbb{Z}, \quad \mathcal{F}_f(k) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-ik\theta} d\theta. \quad (1.7)$$

La paramètre k décrit l'ensemble des entiers relatifs, et $\hat{\mathcal{G}} = \mathbb{Z}$. Les CF de f existent si :

$$\int_0^{2\pi} |f(\theta)| d\theta < \infty. \quad (1.8)$$

L'ensemble des fonctions vérifiant cette condition constitue un espace vectoriel normé noté $\mathbf{L}^1(\mathbb{S}^1, d\theta)$. La norme d'une fonction appartenant à cet espace est, par définition, la valeur de l'intégrale (1.8). Les coefficients de Fourier sont à l'origine des fonctions harmoniques circulaires qui permettent d'étudier le mouvement d'objets bi-dimensionnels en rotation dans le plan. Celles-ci sont présentées dans le paragraphe 1.2.1 page 15.

La transformée inverse d'une fonction représentée par ses CF est définie par les *séries de Fourier*. Celles-ci correspondent à la transformée de Fourier sur \mathbb{Z} du point de vue de l'analyse harmonique commutative :

$$\forall \theta \in \mathbb{S}^1, \quad f(\theta) = \sum_{k=-\infty}^{+\infty} \mathcal{F}_f(k) e^{ik\theta}. \quad (1.9)$$

Les séries de Fourier, si elles existent³, permettent de reconstruire une fonction périodique depuis l'ensemble infini de ses coefficients de Fourier.

³— La sommabilité de la fonction f n'implique pas la sommabilité de ses séries de Fourier.

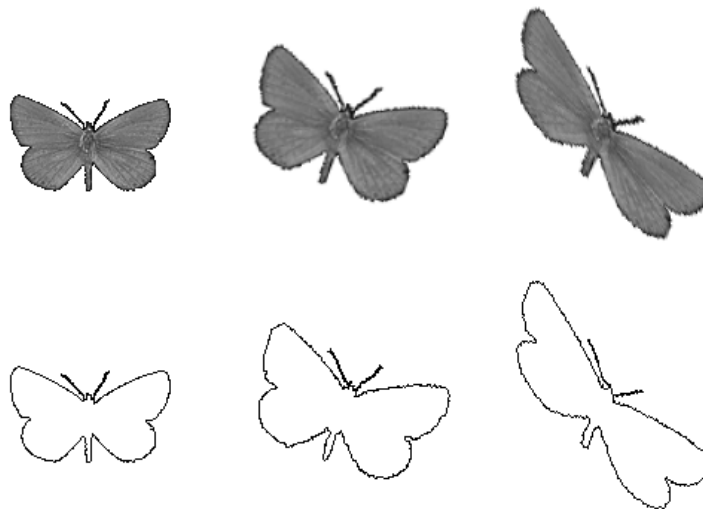


Figure 1.3 — Action d'une similitude et d'une transformation affine sur la représentation d'un papillon et de son profil extérieur.

Soit maintenant la fonction f_β obtenue par rotation d'angle β de la fonction f , *i.e.* $f_\beta(\theta) = f(\theta + \beta)$. Avec une démarche similaire à celle proposée pour la transformée de Fourier, on montre que :

$$\forall k \in \mathbb{Z}, \quad \mathcal{F}_{f_\beta}(k) = e^{ik\theta} \mathcal{F}_f(k).$$

Les modules des CF de ces deux fonctions sont donc identiques et fournissent des descripteurs invariants par rapport au groupe \mathbb{S}^1 pour les fonctions périodiques (cf. figure 1.4).

Les coefficients de Fourier sont particulièrement bien adaptés à l'étude d'objets représentés de manière satisfaisante par leur profil extérieur (les caractères manuscrits, les micro-organismes, ...). L'idée de base est qu'une courbe simple et fermée peut être représentée par une fonction périodique d'un paramètre continu (paramétrisation), ou, de manière équivalente, par les coefficients de Fourier de cette fonction (généralement appelés les *descripteurs de Fourier*). L'adéquation de la paramétrisation à la classe de transformations étudiée permet d'extraire des descripteurs de Fourier invariants à cette classe et au point de description initial du contour.

Dans le cas des *transformations euclidiennes* du plan, *i.e.* une translation suivie d'une similitude, une paramétrisation adéquate est la longueur d'arc normalisée. Celle-ci a permis d'extraire de nombreux ensembles de descripteurs invariants aux similitudes [135] qui ont été appliqués à la reconnaissance de caractères manuscrits [60, 99] ou à la classification d'organismes biologiques [57] par exemple. Ces descripteurs ont été étendus par :

- T.R. Crimmins [29] à des familles de descripteurs *complètes*, garantissant que toute l'information sur la forme des objets est conservée.
- F. Ghorbel [54] à des familles *complètes et stables*, garantissant, de plus, la robustesse des descripteurs vis-à-vis des faibles variations de formes.

Ces familles ont permis de définir une vraie distance entre formes qui a été utilisée pour le suivi et le codage des objets [88, 37].

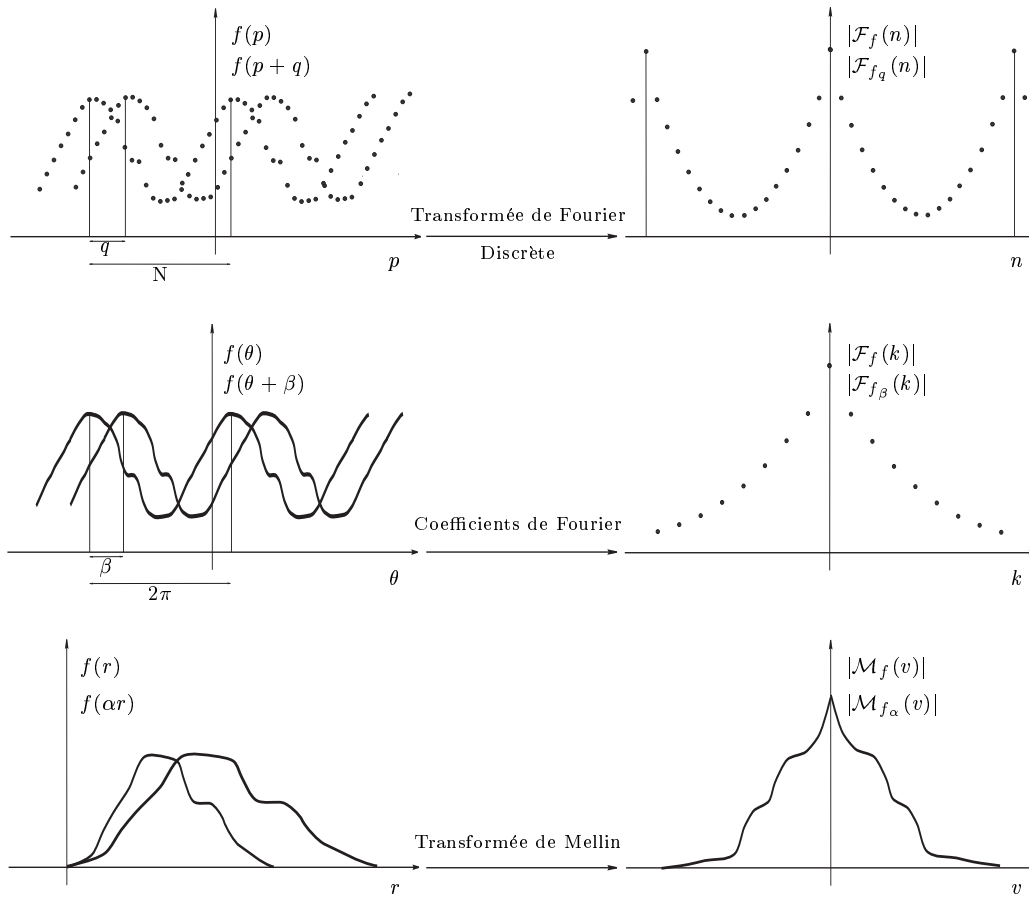


Figure 1.4 — La représentation originale des signaux discrets, périodiques et causaux, et action des groupes de transformations sur leur transformée de Fourier respective. À gauche : les signaux originaux et leur translaté au sens du groupe de transformations ; à droite : l'amplitude de leur transformée de Fourier respective.

Dans le cas des *transformations affines* du plan, *i.e.* une transformation euclidienne et un étirement, une paramétrisation adéquate est donnée par la longueur équi-affine normalisée ou équivalente [9]. À partir des coefficients de Fourier de cette paramétrisation, il a été possible d'extraire des descripteurs invariants sous l'action de cette transformation. Ceux-ci ont été appliqués à la reconnaissance de silhouettes d'avions dont le mouvement apparent dans le plan image est assimilé à une transformation affine [10]. Plus récemment, une famille complète et stable de descripteurs invariants a été proposée [56].

Les limites des descripteurs de Fourier sont atteintes lorsque le contour extérieur n'est pas suffisant pour caractériser un objet particulier parmi un ensemble d'objets. C'est par exemple le cas des images de certains papillons qui présentent des contours sensiblement identiques sans pour autant appartenir à la même famille. L'information discriminante est, dans ce cas, située dans la texture des ailes et des descripteurs fondés sur les niveaux de gris de l'image doivent permettre de distinguer de telles images.

1.1.3.2 La transformée de Fourier discrète

Considérons maintenant le cas d'un signal discret et périodique composé de N échantillons. Le support de la fonction est contenu dans $\mathbb{Z}/N\mathbb{Z}$. Le groupe de transformations est le groupe cyclique des entiers modulo N , muni de l'addition. La transformée de Fourier correspondante est donnée par la *Transformée de Fourier Discrète* (TFD) dont l'expression est la suivante :

$$\forall n \in \mathbb{Z}/N\mathbb{Z}, \quad \mathcal{F}_f(n) = \frac{1}{N} \sum_{p=0}^{N-1} f(p) e^{-i \frac{np}{N}}. \quad (1.10)$$

Cette transformée est bien adaptée à l'étude des images numériques échantillonnées sur des grilles cartésiennes. De plus, la TFD peut être estimée grâce à l'algorithme de la FFT (*Fast Fourier Transform*). Notons enfin que les modules des transformées de Fourier discrètes de deux signaux translatés l'un par rapport à l'autre d'un nombre entier q restent identiques (cf. figure 1.4).

1.1.3.3 La transformée de Mellin

Parmi les transformées de Fourier, la transformée de Mellin⁴ permet d'étudier certaines fonctions indépendamment des dilatations qu'elles ont subies. Dans le domaine du traitement du signal, elle est par exemple utilisée pour la détection, la reconnaissance et la classification de bateaux à partir de leur profil radar [68] et pour le traitement des signaux acoustiques [8, 69].

Soit $f(r)$ une fonction causale ($r > 0$). Le groupe de transformations adapté à ce type de fonction est le groupe des dilatations $\mathcal{G} = \mathbb{R}_+^*$ pour la loi multiplicative et la transformée de Fourier est appelée *Transformée de Mellin* (TM) de f :

$$\forall v \in \mathbb{R}, \quad \mathcal{M}_f(v) = \int_0^{\infty} f(r) r^{-iv} \frac{dr}{r}, \quad (1.11)$$

où $\{r^{-iv}\}$ est la représentation de \mathcal{G} et $\frac{dr}{r}$ sa mesure de Haar. Le paramètre v décrit la droite réelle et le groupe dual est dans ce cas : $\hat{\mathcal{G}} = \mathbb{R}$. La TM d'une fonction causale existe si l'intégrale suivante est finie :

$$\int_0^{\infty} |f(r)| \frac{dr}{r} < \infty. \quad (1.12)$$

L'ensemble des fonctions vérifiant cette condition constitue un espace vectoriel normé noté $\mathbf{L}^1(\mathbb{R}_+^*, \frac{dr}{r})$. La norme sur cet espace est donnée par la valeur de l'intégrale (1.12).

La transformée de Mellin fournit une représentation unique des fonctions. Le signal original peut être reconstruit grâce à la *Transformée de Mellin Inverse* :

$$\forall r \in \mathbb{R}_+^*, \quad f(r) = \int_{-\infty}^{+\infty} \mathcal{M}_f(v) r^{iv} dv. \quad (1.13)$$

Comme dans le cas des coefficients de Fourier, l'existence de la transformée de Mellin inverse n'est pas garantie par l'existence de la transformée directe. Il est donc nécessaire de vérifier

⁴— Hjalmar R. Mellin, 1854-1933. Une brève bibliographie des mathématiciens J. Fourier et H. Mellin est présentée en annexe A page 117.

au préalable l'existence de l'intégrale, c'est-à-dire si $\mathcal{M}_f \in \mathbf{L}^1(\mathbb{R}, dv)$.

La TM se distingue des transformées de Fourier évoquées auparavant, d'une part, par la forme du noyau, et d'autre part, par la nature de la mesure d'intégration à densité $\frac{1}{r}$ par rapport à la mesure de Lebesgue dr . Ces deux caractéristiques font que la TM est adaptée pour la représentation et l'étude des dilatations d'un signal. En effet, si f_α est le signal f dilaté d'un facteur d'échelle α , *i.e.* $\forall r \in \mathbb{R}_+^*$, $f_\alpha(r) = f(\alpha r)$, alors sa transformée de Mellin s'écrit⁵ :

$$\mathcal{M}_{f_\alpha}(v) = \int_0^\infty f(\alpha r) r^{-iv} \frac{dr}{r}.$$

En effectuant le changement de variable $\rho = \alpha r$, nous obtenons :

$$\begin{aligned} \mathcal{M}_{f_\alpha}(v) &= \int_0^\infty f(\rho) \left(\frac{\rho}{\alpha}\right)^{-iv} \frac{d\rho \alpha}{\alpha \rho}, \\ &= \alpha^{iv} \int_0^\infty f(\rho) \rho^{-iv} \frac{d\rho}{\rho}, \\ &= \alpha^{iv} \mathcal{M}_f(v). \end{aligned}$$

C'est le théorème du retard de la TM relativement au groupe des dilatations. Comme dans les cas précédents, l'influence de la transformation géométrique se répercute sur la phase de la transformée de Mellin. Le module de la TM est donc invariant sous l'action des dilatations. Ces descripteurs ont, par exemple, été utilisés pour caractériser le type d'un navire indépendamment de l'angle sous lequel celui-ci est observé [68].

Dans la pratique, les signaux sont acquis de manière discrète, avec un pas d'échantillonnage constant. Cela pose le problème de l'approximation numérique de la transformée de Mellin. Une première solution consiste à estimer directement l'intégrale de Mellin à partir des échantillons. En général, les résultats numériques ne sont pas satisfaisants et on préfère utiliser la solution suivante. En effectuant le changement de variable $t = \ln(r)$, pour $r \in \mathbb{R}_+^*$, dans l'expression de la transformée de Mellin (1.11), nous obtenons :

$$\forall v \in \mathbb{R}, \quad \mathcal{M}_f(v) = \int_{-\infty}^{+\infty} f(e^t) e^{-itv} dt = 2\pi \mathcal{F}_g(v),$$

en notant $g(t) = f(e^t)$. Ainsi la TM de f s'exprime comme la transformée de Fourier du signal original échantillonné à pas exponentiel.

Ce changement de variable transforme la demi-droite positive en une droite et la dilatation est convertie en une translation. Ce lien étroit entre les deux transformées est à l'origine, d'une part, d'un théorème d'échantillonnage adapté à la transformée de Mellin [58], et d'autre part, d'un algorithme d'approximation de la transformée de Mellin fondé sur la TFD [134], après ré-échantillonnage exponentiel des données numériques. Les aspects numériques de l'approximation de la transformée de Mellin seront détaillés dans le chapitre 2.

⁵— Par convention, nous utilisons la notation $f_\alpha(r) = f(\alpha r)$ plutôt que $f(\frac{r}{\alpha})$. Si $\alpha > 1$ alors la fonction f_α est « plus petite » que f , et inversement, si $\alpha < 1$ alors f_α est « plus grande » que f .

1.2 La représentation des images par les coefficients de Fourier

Dans la section précédente nous avons présenté trois transformées de Fourier définies pour des fonctions mono-dimensionnelles, dont le support et le groupe de transformations sont identiques. Dans le cas des images 2D, les objets sont généralement décrits selon leurs coordonnées cartésiennes (x, y) ou leurs coordonnées polaires $(r = \sqrt{x^2 + y^2}, \theta = \arctan \frac{y}{x})$ ⁶. La représentation cartésienne convient à l'étude des objets en translation dans le plan (transformée de Fourier 2D). Un cas plus intéressant, qui concerne plus de situations concrètes de la vision, est l'étude des objets en rotation dans le plan. Les fonctions circulaires harmoniques de l'image procurent alors une représentation pertinente. En les associant avec des moments radiaux, il est également possible de définir des descripteurs invariants par rotation et dilatation.

1.2.1 Les fonctions harmoniques circulaires

Les *Fonctions Harmoniques Circulaires* (FHC) d'une image $f(r, \theta)$ se présentent comme les coefficients de Fourier (1.7) de chaque cercle de rayon r de l'image (cf. figure 1.5) :

$$\forall k \in \mathbb{Z}, \quad f_k(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) e^{-ik\theta} d\theta. \quad (1.14)$$

La décomposition d'un objet en FHC dépend du choix de l'origine des coordonnées. Le centre du repère est généralement fixé au centre de masse de l'objet, ce qui permet d'obtenir une représentation invariante par translation.

L'utilisation des fonctions harmoniques circulaires en traitement d'images a été initiée dans le domaine de la reconnaissance de formes optique dans les années 80, notamment par des méthodes de corrélation optique entre les FHC d'un objet et d'un modèle par un filtre holographique invariant par rotation [121]. Du fait de certaines difficultés de mise en œuvre optique, notamment l'estimation du centre de rotation des images, ces techniques ont été utilisées en vision numérique pour la reconnaissance invariante de formes par rotation [130, 131].

Une image est entièrement déterminée par l'ensemble infini de ses FHC et la représentation originale peut être reconstruite à l'aide de ses séries de Fourier (1.9) :

$$\forall (r, \theta) \in \mathcal{G}, \quad f(r, \theta) = \sum_{k \in \mathbb{Z}} f_k(r) e^{ik\theta}.$$

En pratique, seul un nombre fini de FHC sont utilisées. Dans [11], les auteurs proposent des résultats de reconstruction numérique d'une image à niveaux de gris à partir d'un nombre fini $k \in [-m; m]$ de fonctions circulaires harmoniques. En prenant suffisamment d'harmoniques, l'erreur de reconstruction provient essentiellement de la double conversion des coordonnées cartésiennes-polaires.

Les FHC intègrent l'objet en fonction de la coordonnée angulaire θ , et l'objet reste inchangé pour la coordonnée radiale r . Un ensemble de descripteurs invariants par rotation

⁶— Le changement de variable entre coordonnées cartésiennes et coordonnées polaires n'est valable que sur le plan privé de son centre $(0, 0)$. Cette remarque prendra toute son importance dans la prochaine section.

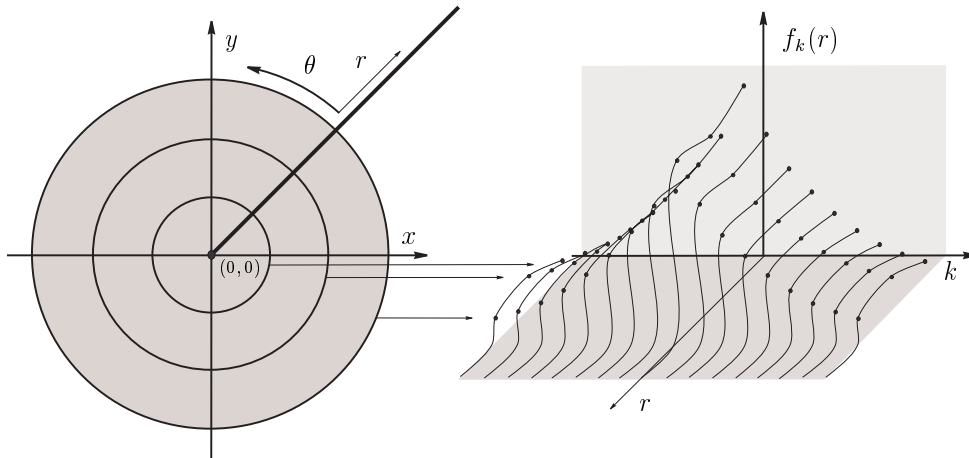


Figure 1.5 — Représentation des images par les fonctions circulaires harmoniques.

est obtenu en prenant les modules des FHC. Lorsqu'un objet est agrandi, la représentation angulaire de l'objet ne varie pas et les FHC répercutent simplement le facteur d'échelle. Pour obtenir des descripteurs également invariants par facteur d'échelle, les fonctions circulaires harmoniques peuvent être combinées avec des moments radiaux comme nous le présentons dans le paragraphe suivant.

1.2.2 Les moments des images à niveaux de gris

Les méthodes de représentation des images par les moments sont parmi les premières à avoir été appliquées en reconnaissance de formes. La principale motivation étant d'extraire des descripteurs de formes invariants par rapport à certaines transformations linéaires générales pour la classification d'images. Depuis le travail initial de H. Ming-Kuel [86] en 1962 sur les invariants issus des moments géométriques de l'image (cf. annexe B page 119), plusieurs approches ont été proposées. La plupart des moments ainsi définis s'expriment comme des moments radiaux des fonctions circulaires harmoniques de l'image. Ces méthodes se distinguent principalement par la forme du noyau radial utilisé, qui s'avère plus ou moins bien adapté à l'extraction de descripteurs invariants par similitudes planes.

Les nombreuses approches proposées ont conduit à plusieurs états de l'art, notamment [101, 129, 103], qui mettent en évidence le lien entre les différents moments, leurs propriétés et leurs applications pratiques. Dans cette synthèse, nous mettons l'accent sur une propriété qui nous semble souhaitable pour obtenir une caractérisation pertinente des objets à niveaux de gris : l'unicité de la représentation et des descripteurs invariants extraits à partir de celle-ci. Ceci garanti d'une part la possibilité de reconstruire un objet depuis sa représentation par les moments, et d'autre part, limite la confusion d'objets de formes différentes.

1.2.2.1 Les moments complexes

Les *Moments Complexes* (MC) de l'image ont été développés par Davis [31] et par Y. Abu-Mostafa et D. Psaltis [5, 6] comme une extension de la méthode des moments géométriques

pour la représentation des images :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad c_f(p, q) = \iint_{\mathbb{R}^2} f(x, y) (x - iy)^p (x + iy)^q dx dy.$$

En effectuant un changement de variables cartésien-polaire, les MC s'expriment comme les moments radiaux des fonctions harmoniques circulaires de f :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad c_f(p, q) = \int_0^\infty r^{p+q+1} f_{p-q}(r) dr. \quad (1.15)$$

Les moments complexes sont particulièrement bien adaptés pour traiter simultanément les changements d'orientation et d'échelle des objets. Soient deux images f et g identiques à une similitude près de paramètres (α, β) , *i.e.* $g(r, \theta) = f(\alpha r, \theta + \beta)$. On montre que les MC d'ordre $(p + q)$ de f et g sont liés par la relation suivante :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad c_g(p, q) = \alpha^{-\gamma} e^{i(p-q)\beta} c_f(p, q) \quad (1.16)$$

avec $\gamma = p + q + 2$.

Ainsi, une similitude dans le domaine de représentation original se traduit par une multiplication complexe dans le domaine de représentation des moments complexes. Par analogie avec la transformée de Fourier, cette relation est appelée théorème du retard des MC. L'orientation relative des objets est codée dans la phase et leur différence de taille est codée dans l'amplitude des MC. En prenant les modules des moments complexes, nous obtenons une famille non complète de descripteurs invariants uniquement par rotation. L'invariance par facteur d'échelle est obtenue par un procédé de normalisation identique à celui qui permet d'obtenir les moments réduits à partir des moments centrés (cf. annexe B page 119).

L'ensemble des descripteurs

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad J_f(p, q) = c_f(0, 0)^{\frac{\gamma}{2}} |c_f(p, q)|, \quad (1.17)$$

forme une famille non complète de descripteurs invariants par similitudes planes. Une solution pour atteindre la complétude de la famille a été récemment proposée [82]. Celle-ci est obtenue par des combinaisons linéaires des moments géométriques et par un procédé de normalisation similaire à celui décrit dans le chapitre 3.

On montre que les MC s'expriment comme des combinaisons linéaires des moments géométriques de l'image [31], ce qui permet, en principe, de reconstruire une image depuis ses moments complexes. Cependant, les moments complexes n'ont pas permis de restituer une image car la transformée inverse pose des problèmes d'instabilités numériques identiques aux moments géométriques (cf. annexe B).

1.2.2.2 Les moments de Zernike

Les *Moments de Zernike* (MZ) de l'image ont été introduits par M.R. Teague [123]. Il a proposé d'utiliser les polynômes complexes de Zernike, orthogonaux à l'intérieur du cercle unité, que l'on peut écrire sous la forme :

$$V_{nm}(r, \theta) = R_{nm}(r) e^{-im\theta}, \quad \text{avec} \quad \begin{cases} n \in \mathbb{N} \\ m \in \mathbb{Z} \text{ tel que } n - |m| \text{ impair et } |m| \leq n \end{cases} \quad (1.18)$$

avec

$$R_{nm}(r) = \sum_{s=0}^{n-\frac{|m|}{2}} \frac{(-1)^s (n-s)! r^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!}.$$

Les moments de Zernike $\{A_f(n, m)\}$ d'une image f sont les projections de l'image sur la base $\{V_{nm}(r, \theta)\}$ et donc :

$$A_f(n, m) = \frac{n+1}{\pi} \int_0^1 f_m(r) R_{nm}(r) dr. \quad (1.19)$$

Les MZ se présentent donc comme les moments radiaux, pour la base des polynômes orthogonaux $\{R_{nm}(r)\}$, des fonctions circulaires harmoniques $h_m(r)$ de l'image f .

On montre que les moments de Zernike des deux images f et g ne se différenciant que d'une rotation sont liés par la relation suivante :

$$A_g(n, m) = e^{+im\beta} A_f(n, m), \quad (1.20)$$

en prenant en compte le fait que les parties radiales et angulaires des polynômes de Zernike sont séparables. Une fois encore l'angle de rotation est codé dans la phase des moments et la forme la plus simple des descripteurs invariants par rotation est obtenue en prenant les modules des moments de Zernike (ensemble non-complet) [123]. Récemment, Å. Wallin et O. Kübler [126] ont proposé une famille de descripteurs composée d'un ensemble d'invariants par rotation et d'un ensemble de pseudo-invariants par réflexion⁷. Cette famille capture l'ensemble de l'information sur la forme des objets. Cependant, l'unicité des descripteurs purement invariants par rotation n'a pas été atteinte. L'invariance par facteur d'échelle est, quant à elle, difficile à réaliser du fait de la complexité de l'expression des polynômes radiaux.

Plusieurs études [124, 17] rapportent que certains invariants issus des moments de Zernike donnent des résultats de reconnaissance et classification de formes supérieurs à ceux obtenus par les moments géométriques et les moments complexes. Il est également connu que les moments de Zernike sont plus robustes au bruit [72]. La principale raison évoquée est que la base de projection radiale utilisée est orthogonale. Ainsi, les moments orthogonaux caractérisent les images de manière optimale dans la mesure où chaque moment apporte une information supplémentaire sur l'image, contrairement aux moments non-orthogonaux qui sont en parti redondants. D'autres polynômes orthogonaux à l'intérieur du cercle unité provenant de la famille des polynômes de Jacobi [103] ont également été utilisés. Les moments obtenus restent essentiellement identiques aux moments de Zernike.

Une image peut être reconstruite directement à partir de l'ensemble infini de ses moments de Zernike :

$$f(x, y) = \sum_n \sum_m A_f(n, m) V_{nm}(x, y).$$

⁷— Le terme *pseudo-invariant* désigne un descripteur qui change de signe sous l'action d'une transformation géométrique.

avec n et m définis dans l'équation (1.18). Un nombre fini de moments de Zernike a permis de reconstruire des images avec une qualité satisfaisante. Plusieurs équipes ont étudié le comportement numérique des moments de Zernike pour la reconstruction des images en fonction du nombre de moments utilisés et des diverses approximations numériques nécessaires à leur implémentation (discrétisation, ré-échantillonnage des images et approximation des intégrales) [71, 77, 78]. Bien que ces moments soient adaptés pour la description des images à niveaux de gris, les expériences rapportées dans ces travaux se limitent généralement à des images binaires telles que des lettres typographiques, ce qui ne permet pas de juger de la reconstruction des objets à niveaux de gris de forme et texture complexes.

1.2.2.3 Les moments de Fourier-Mellin

Les *Moments de Fourier-Mellin* (MFM) de l'image ont été introduits en 1986 par H. ARSENAULT et Y. SHENG comme des moments radiaux de la représentation en fonctions circulaires harmoniques des images. Deux types de descripteurs ont été proposés. Ceux-ci dépendent de la base radiale utilisée :

- En utilisant une base radiale non-orthogonale [115, 117], les MFM sont essentiellement identiques aux moments complexes. À partir de cette représentation, un nouvel ensemble de descripteurs invariants par similitude plane a été constitué par combinaisons non-linéaires des MFM [76]. Cette famille n'a, toutefois, pas permis d'atteindre la complétude ;
- En utilisant une base radiale orthogonale [114, 118] de la famille des polynômes de Jacobi, les MFM orthogonaux sont essentiellement identiques aux moments de Zernike. Le problème de l'invariance par facteur d'échelle n'a toutefois pas été résolu du fait de la complexité des polynômes radiaux utilisés.

Le terme de *moment de Fourier-Mellin* est, semble-t-il, incorrect car la base radiale utilisée ne correspond pas à celle de la transformée de Mellin. En effet, dans le premier cas, le noyau est du type r^k avec k un entier et, dans le second cas, le noyau de Mellin est du type r^{iv} avec v un nombre réel. La nature de la représentation s'en trouve sensiblement modifiée. Les moments de Fourier-Mellin proviennent de la théorie des moments algébriques et non de la théorie de l'analyse harmonique comme leur nom semble l'indiquer. Dans la section suivante, nous présentons la transformée de Fourier-Mellin dans le cadre théorique de l'analyse harmonique comme une autre méthode de représentation des images à niveaux de gris.

1.3 La transformée de Fourier-Mellin

La *Transformée de Fourier-Mellin* (TFM) peut être vue comme l'association des fonctions harmoniques circulaires et de la transformée de Mellin que nous avons présentées dans les sections précédentes. Dans le cadre rigoureux de l'analyse harmonique commutative sur les groupes que nous avons évoqué dans le paragraphe 1.1.2, la transformée de Fourier-Mellin est la transformée de Fourier sur le produit direct du groupe des dilatations vectorielles \mathbb{R}_+^* , par le groupe des rotations vectorielles \mathbb{S}^1 de même centre. Ce groupe, noté $\mathcal{G} = \mathbb{R}_+^* \times \mathbb{S}^1$ ⁽⁸⁾, est

⁸— La lettre \mathcal{G} désignera désormais le groupe des similitudes.

appelé *groupe des similitudes vectorielles du plan* pour la loi de composition suivante :

$$(\alpha, \theta) \circ (\rho, \psi) = (\alpha\rho, \theta + \psi).$$

\mathcal{G} est un groupe topologique, localement compact et abélien. Il est équivalent à l'espace des coordonnées polaires, ce qui en fait un groupe de transformations transitif sur cet espace.

1.3.1 La transformée de Fourier-Mellin standard

La transformée de Fourier-Mellin standard se présente comme l'association des fonctions circulaires harmoniques et de la transformée de Mellin radiale. La représentation $\{r^{-iv} e^{-ik\theta}\}$ du groupe \mathcal{G} est le produit des représentations individuelles des groupes des rotations et des homothéties vectorielles. La mesure invariante $d\theta \frac{dr}{r}$ est également le produit des mesures invariantes de chacun des groupes.

La TFM standard d'une fonction f définie en coordonnées polaires, lorsqu'elle existe, est donnée par⁹ :

$$\forall (k, v) \in \mathbb{Z} \times \mathbb{R}, \quad \mathcal{M}_f(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{-iv} e^{-ik\theta} d\theta \frac{dr}{r}. \quad (1.21)$$

L'espace des paramètres (k, v) de la TFM définit le groupe dual de \mathcal{G} : $\hat{\mathcal{G}} = \mathbb{Z} \times \mathbb{R}$.

La transformée de Fourier-Mellin, comme transformée de Fourier sur le groupe des similitudes planes, possède un certain nombre de propriétés qui la rend adaptée pour l'analyse des objets à niveaux de gris soumis à l'action des rotations et des dilatations. De plus, la transformée de Fourier-Mellin fournit une représentation unique des images. Comme \mathcal{G} est commutatif, la transformée de Fourier sur $\hat{\mathcal{G}}$ existe et est appelée *transformée de Fourier-Mellin inverse* de f :

$$\forall (r, \theta) \in \mathcal{G}, \quad f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) r^{iv} e^{ik\theta} dv, \quad (1.22)$$

$\delta k \otimes dv$ étant la mesure invariante du groupe $\hat{\mathcal{G}}$. La TFM standard permet, en principe, de reconstruire une image à partir de sa TFM. De manière symétrique à la TFM, la TFM inverse peut être vue comme l'association des séries de Fourier (1.9) avec la transformée de Mellin inverse (1.13).

Bien qu'inversible, la TFM standard n'a pas permis, à notre connaissance, de reconstruire un objet à niveaux de gris. Les problèmes numériques rencontrés sont essentiellement issus de l'approximation de la TFM directe, comme nous le justifions dans le paragraphe suivant.

⁹— Nous utilisons la même lettre \mathcal{M} pour désigner la TM et la TFM. Cependant, il ne peut pas y avoir de confusion entre ces deux transformées car la première dépend d'un paramètre unique (v), et la seconde dépend de deux paramètres (k et v).

1.3.2 Existence de la TFM pour des images

La transformée de Fourier-Mellin d'un objet existe si sa représentation f est intégrable sur $\mathbb{R}_+^* \times \mathbb{S}^1$, c'est-à-dire si :

$$\int_0^\infty \int_0^{2\pi} \left| f(r, \theta) r^{-iv} e^{-ik\theta} \right| d\theta \frac{dr}{r} = \int_0^\infty \int_0^{2\pi} \frac{1}{r} f(r, \theta) d\theta dr < \infty, \quad (1.23)$$

puisque f est positive. L'ensemble des fonctions réelles et positives pour lesquelles la TFM existe constitue un espace vectoriel normé, noté $\mathbf{L}^1(\mathbb{R}_+^* \times \mathbb{S}^1, d\theta \frac{dr}{r})$, ou plus simplement, $\mathbf{L}^1(\mathcal{G})$, la mesure de Haar étant implicite. La valeur de l'intégrale (1.23) définit la norme d'une fonction sur cet espace : $\|f\|_{\mathbf{L}^1(\mathcal{G})}$.

En réalité, la majeure partie des images à niveaux de gris ne peuvent pas être considérées comme des fonctions de $\mathbf{L}^1(\mathcal{G})$. L'existence de l'intégrale à l'infini ne pose pas de problèmes puisque f représente un objet d'étendue spatiale finie. Par contre, l'existence de l'intégrale n'est pas garantie au centre des coordonnées pour des objets à niveaux de gris. En effet, la TFM standard est définie pour des fonctions équivalentes à Kr^λ au voisinage de l'origine des coordonnées polaires, pour $\lambda > 0$ et pour une constante positive K . Cependant, un objet ne réalise généralement pas cette hypothèse. En pratique, l'origine des coordonnées polaires est située au centre de masse de l'objet pour obtenir une représentation de l'image invariante par translation. Ainsi, au voisinage du point $(0, 0)$, la fonction f est équivalente au niveau de gris du centre de masse, valeur généralement non nulle. En conséquence, f n'est pas sommable sur \mathcal{G} et l'intégrale (1.23) diverge.

Pour résoudre le problème de singularité à l'origine des coordonnées, plusieurs auteurs ont proposé une approche numérique. La méthode consiste à supprimer un petit disque de rayon ε au centre de masse de l'image. Ainsi, le problème de divergence disparaît et la TFM standard peut être estimée. Sur ce principe, P.E. Zwicke et Z. Kiss [137] ont proposé plusieurs algorithmes pour estimer la TFM. Cependant, cette technique a de graves conséquences sur l'approximation numérique de la TFM, principalement à cause des deux raisons suivantes :

- Elle engendre des problèmes de déformations des images lorsque celles-ci changent de taille. En particulier, lorsqu'un objet est agrandi d'un facteur inconnu, quel rayon ε' doit-on prendre pour le disque ? Si le diamètre du disque est identique pour toutes les images, on ne retranche pas la même quantité d'information de chaque image.
- Les pixels de l'image proches du centre de masse ont une influence plus importante dans le calcul de la TFM que les pixels éloignés du centre. Ceci est dû à la densité $\frac{1}{r}$ dans la mesure invariante du groupe. Ainsi, lorsqu'on supprime ne serait-ce qu'un petit disque au centre du repère polaire, une quantité importante d'information pour le calcul de la TFM est perdue.

L'intérêt de la TFM pour la reconnaissance de forme provient essentiellement du théorème du retard par similitude qu'elle hérite, par extension, de la transformée de Fourier classique. On montre en effet aisément que les modules des harmoniques de Fourier-Mellin de deux objets ne se distinguant que par leur orientation et leur taille dans l'image sont identiques. Les descripteurs issus de cette représentation sont appelés *Descripteurs de Fourier-Mellin* (DFM),

qu'il ne faut pas confondre avec les moments de Fourier-Mellin que nous avons évoqués dans le paragraphe 1.2.1. Cependant, étant donné les problèmes évoqués ci-dessus, ils ont été peu utilisés étant donné leur très forte sensibilité numérique. Citons néanmoins l'étude de A.E. Grace et M. Spann [59], dont les expériences montrent que les taux de reconnaissance d'images binaires à partir des DFM s'avèrent supérieurs à ceux obtenus par les invariants des moments de Hu, surtout lorsque les images sont bruitées. En pratique, seules quelques harmoniques ont été utilisées. La reconstruction d'un objet depuis sa transformée de Fourier-Mellin standard devient impossible étant donné du faible nombre d'harmoniques calculées et de leur fiabilité numérique.

1.4 Conclusion du chapitre

Guidés par quelques résultats de l'analyse harmonique, nous avons présenté la transformée de Fourier-Mellin comme la transformée de Fourier sur le groupe des similitudes planes. Cette démarche nous a permis de décrire quelques unes des méthodes de descriptions globales parmi les plus utilisées en reconnaissance de formes, notamment les descripteurs de Fourier dans le cas des objets-contour et les moments dans le cas des images à niveaux de gris.

Nous avons également pu constater que la transformée de Fourier-Mellin, dans sa définition standard, n'est pas adaptée pour les fonctions représentant des objets à niveaux de gris. Elle pose des problèmes numériques cruciaux qui ont largement déprécié son application tant en reconnaissance de formes que pour son application en indexation de bases d'images, en tatouage d'images et en codage orienté objet. Dans le chapitre suivant, nous présentons le prolongement analytique de la transformée de Fourier-Mellin comme une solution rigoureuse au problème de l'approximation de la TFM.

Approximations de la TFM et Reconstruction des Images

Comme nous l'avons constaté, la transformée de Fourier-Mellin pose des problèmes numériques théoriques pour son application en traitement d'images. Dans ce chapitre, nous présentons le prolongement analytique de la TFM qui permet de résoudre le problème de divergence des intégrales lorsqu'elles sont calculées sur des objets à niveaux de gris (section 2.1).

Fort de cette solution, nous proposons dans la section 2.2 trois méthodes pour estimer les transformées de Fourier-Mellin directe et inverse : l'approximation directe, l'approximation rapide et l'approximation cartésienne. Ces algorithmes se distinguent par la manière dont les images discrètes sont ré-échantillonnées sur de nouvelles grilles, et par la méthode numérique d'intégration adéquate pour ces grilles.

Pour valider et comparer les trois algorithmes, le problème inverse de la reconstruction d'une image est examiné dans la section 2.3. À partir d'un spectre de Fourier-Mellin discret et d'étendue finie, des expériences sur la reconstruction d'images à niveaux de gris sont présentées. La comparaison des trois méthodes est réalisée par une inspection visuelle des images restituées et par le calcul d'une erreur de reconstruction. Dans les deux chapitres suivants, la comparaison des différentes approximations est complétée par d'autres critères portant sur la reconnaissance de formes et l'estimation de mouvement.

2.1 La transformée de Fourier-Mellin analytique

La section 1.3 page 19, nous avons mis en évidence que l'estimation de la TFM standard, pour des objets décrits par rapport à leur centre d'inertie, conduisait à des difficultés de réalisation numériques. Les solutions proposées ne sont pas satisfaisantes tant d'un point de vue théorique que numérique. Dans cette section, nous présentons la solution proposée par F. GHORBEL [55]. Cette méthode s'appuie sur une modification réversible de l'image préalable au calcul de la TFM standard, conduisant à définir la *Transformée de Fourier-Mellin Analytique* (TFMA). Nous présentons également quelques propriétés de la TFMA qui nous seront utiles dans les sections suivantes.

2.1.1 La TFM analytique des objets à niveaux de gris

Une solution rigoureuse au problème d'existence de l'intégrale de Fourier-Mellin a été récemment proposée [55]. Cette méthode consiste à modifier la fonction f (représentant l'objet) comme étape préalable au calcul de la transformée de Fourier-Mellin, selon la transformation suivante :

$$\forall(r, \theta) \in \mathbb{R}_+^* \times \mathbb{S}^1, \quad f_\sigma(r, \theta) = r^\sigma f(r, \theta). \quad (2.1)$$

où σ est un réel strictement positif fixé. La modification est réversible et nous permet de retrouver l'originale f en effectuant la transformation inverse :

$$\forall(r, \theta) \in \mathbb{R}_+^* \times \mathbb{S}^1, \quad f(r, \theta) = r^{-\sigma} f_\sigma(r, \theta).$$

L'image étant à support compact, la fonction f_σ est bornée. Cette modification permet d'éviter la particularité de f à l'origine des coordonnées polaires et la fonction f_σ est alors intégrable sur \mathcal{G} . La TFM standard de f_σ est appelée *transformée de Fourier-Mellin analytique* (TFMA) de f , que l'on peut directement écrire sous la forme suivante :

$$\boxed{\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r}} \quad (2.2)$$

Le paramètre σ assure l'existence de la TFM pour les images à niveaux de gris. Le lien entre la TFM standard et la TFM analytique est analogue à la relation entre la transformée de Fourier sur \mathbb{R} et la transformée de Laplace. En ce sens, la TFM analytique peut être vue comme la transformée de Laplace sur le groupe \mathcal{G} du point de vue de l'analyse harmonique sur les groupes.

En appliquant la transformée de Fourier-Mellin inverse à \mathcal{M}_{f_σ} (1.22), nous obtenons :

$$\forall(r, \theta) \in \mathbb{R}_+^* \times \mathbb{S}^1, \quad f_\sigma(r, \theta) = r^\sigma f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) r^{iv} e^{ik\theta} dv,$$

que l'on peut directement écrire, en intégrant r^σ dans l'intégrale, sous la forme suivante :

$$\boxed{\forall(r, \theta) \in \mathcal{G}, \quad f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) r^{-\sigma+iv} e^{ik\theta} dv} \quad (2.3)$$

Cette dernière équation permet de reconstruire un objet à partir de sa représentation de Fourier-Mellin. Il faut cependant noter que l'existence de la TFM n'implique pas l'existence de la transformée inverse. Ainsi, nous supposons que $\mathcal{M}_{f_\sigma} \in \mathbf{L}^1(\hat{\mathcal{G}})$, où $\mathbf{L}^1(\hat{\mathcal{G}})$ désigne l'espace vectoriel normé des fonctions complexes intégrables¹ sur $\hat{\mathcal{G}}$:

$$\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_{f_\sigma}(k, v)| dv < \infty \quad (2.4)$$

Dans le paragraphe suivant, nous examinons le cas particulier de la TFMA d'un demi-disque et présentons quelques propriétés qui nous serviront lors de la prochaine section.

2.1.2 Quelques résultats sur la TFMA

✓ Exemple analytique : La TFMA d'un demi-disque

À titre d'exemple, examinons le cas d'une fonction représentant un demi-disque de niveaux de gris constant L :

$$f(r, \theta) = \begin{cases} L & \text{pour } (r, \theta) \in [0; R] \times [\delta; \delta + \pi] \\ 0 & \text{ailleurs} \end{cases}$$

où R désigne le rayon du disque, et δ l'orientation initiale du demi-disque. La figure 2.1 représente le demi-disque f et la fonction modifiée f_σ en coordonnées cartésiennes et polaires. Sa transformée de Fourier-Mellin est donnée par :

$$\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma}(k, v) = L \frac{R^{\sigma - iv}}{-\sigma + iv} \begin{cases} -\frac{1}{2} & \text{pour } k = 0 \\ \frac{i}{k\pi} e^{-ik\delta} & \text{pour } k \text{ impair} \\ 0 & \text{ailleurs} \end{cases} \quad (2.5)$$

Dans la figure 2.2, nous avons représenté la phase (en radians) et l'amplitude de la TFMA du demi-disque. Les harmoniques suivant l'axe v sont connectées pour schématiser une variable réelle. Par contre, les valeurs suivant l'axe k ne le sont pas pour rendre compte de l'aspect discret de la variable. On constate que la transformée de Fourier-Mellin du demi-disque converge partout. Si nous avons utilisé la TFM standard, ce qui consiste à remplacer σ par 0 dans l'équation (2.2), alors le spectre aurait divergé pour $v = 0$ et k quelconque.

L'exemple du demi-disque permet d'illustrer les quelques propriétés de la TFMA que nous présentons ci-dessous. Il nous servira également à illustrer d'autres résultats dans la suite de ce mémoire.

¹— Une condition suffisante pour que la fonction \mathcal{M}_f soit intégrable sur $\hat{\mathcal{G}}$ consiste à considérer les objets comme des fonctions intégrables, *définies positives et continues* sur \mathcal{G} [106, chapitre 1].

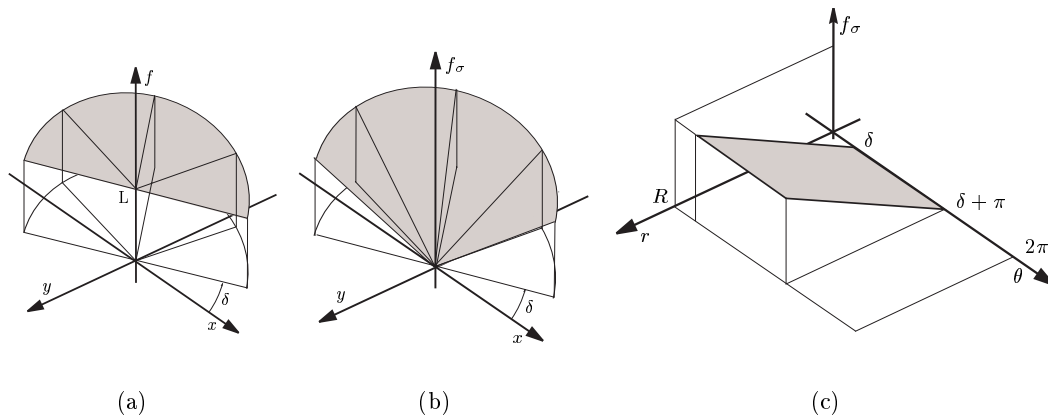


Figure 2.1 — Représentation 2,5D de la surface d'un demi-disque en coordonnées cartésiennes et polaires. (a) Surface cartésienne; (b) Surface après modification ($\sigma = 1$); (c) Représentation polaire de (b).

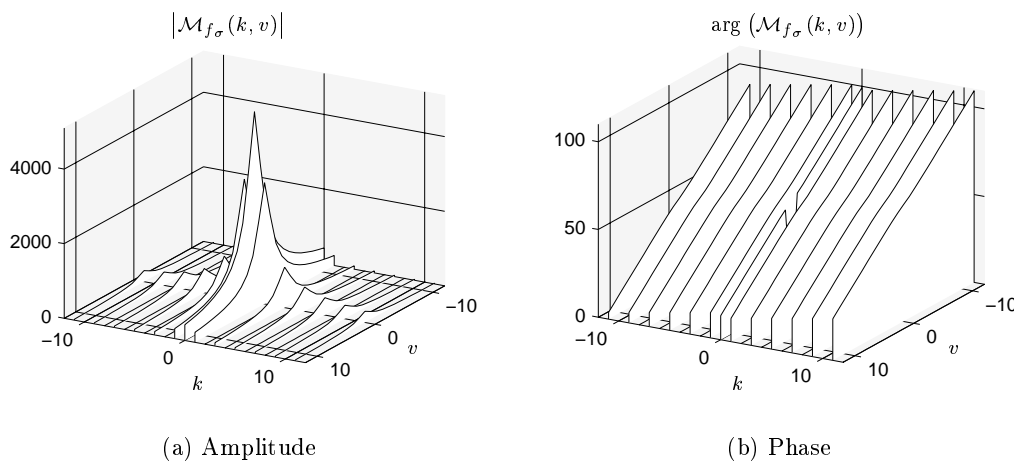


Figure 2.2 — Amplitude et phase (en radians) de la TFM analytique du demi-disque de la figure 2.1 avec $\sigma = 1$, $R = 30$ et $\delta = \frac{\pi}{5}$.

✓ Symétrie hermitienne de la TFM

Soit $f_\sigma \in \mathbf{L}^1(\mathcal{G})$ et, quelque soit $(k, v) \in \mathcal{G}$, nous avons :

$$\begin{aligned}
 \mathcal{M}_{f_\sigma}(-k, -v) &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma+iv} e^{+ik\theta} d\theta \frac{dr}{r} \\
 &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) \overline{r^{\sigma-iv} e^{-ik\theta}} d\theta \frac{dr}{r} \\
 &= \frac{1}{2\pi} \overline{\int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r}} \\
 &= \overline{\mathcal{M}_{f_\sigma}(k, v)}
 \end{aligned} \tag{2.6}$$

La transformée de Fourier-Mellin analytique possède une symétrie hermitienne. Seule la moitié du spectre pourra donc être considérée pour son approximation numérique.

✓ Majoration de la TFM

Les valeurs de la fonction f représente la luminance de l'image, c'est donc une fonction réelle et positive et $\mathcal{M}_{f_\sigma}(0, 0)$ est une nombre réel positif. Comme extension du résultat classique pour la transformée de Fourier, on montre également que, quelque soit $(k, v) \in \mathcal{G}$, nous avons :

$$|\mathcal{M}_{f_\sigma}(k, v)| \leq \mathcal{M}_{f_\sigma}(0, 0) \quad (2.7)$$

La transformée de Fourier-Mellin analytique est majorée par sa composante centrale.

✓ Théorème de Riemann-Lebesgue

Soit $f_\sigma \in \mathbf{L}^1(\mathcal{G})$ et \mathcal{M}_{f_σ} sa transformée de Fourier-Mellin. Le théorème de Riemann-Lebesgue dans le cas du groupe \mathcal{G} s'écrit :

$$\mathcal{M}_{f_\sigma}(k, v) \longrightarrow 0 \text{ lorsque } (k, v) \rightarrow \pm\infty \quad (2.8)$$

Comme l'étendue des objets à niveaux de gris est limitée dans l'espace, sa TFM est, au contraire, d'étendue infinie (principe de dualité). L'intérêt de ce résultat apparaît avec l'approximation numérique de la TFM que nous présentons dans la section suivante.

Comme extension de la transformée de Fourier sur le groupe des similitudes, la TFM possède un certain nombre de propriétés particulièrement intéressantes pour l'analyse de formes à niveaux de gris. Il est notamment possible de redéfinir le théorème du retard, l'identité de Parseval et le théorème de corrélation pour le groupe des similitudes planes vectorielles. Ces propriétés sont exploitées pour la reconnaissance de formes dans le chapitre 3, et pour l'estimation des paramètres de mouvement dans le chapitre 4. Auparavant, pour poursuivre ce chapitre, il convient de proposer des approximations numériques de la TFMA pour son application sur des images réelles.

2.2 Trois méthodes d'approximations de la TFMA

Dans la section précédente, nous avons présenté la méthode du prolongement analytique comme une solution au problème de divergence de l'intégrale de Fourier-Mellin. Les transformées de Fourier-Mellin directe et inverse sont définies pour des fonctions à paramètres réels et continus ($r \in \mathbb{R}_+^*$, $\theta \in \mathbb{S}^1$ et $v \in \mathbb{R}$). Comme il n'existe pas de transformée discrète, il convient maintenant de proposer des approximations robustes de ces transformées pour leur implémentation algorithmique avec des données numériques.

Notre travail a conduit à la réalisation de trois algorithmes d'estimation de la TFMA et de son inverse. Ces approximations diffèrent : (i) par la manière dont les images sont ré-échantillonnées sur de nouvelles grilles et, (ii) par la méthode numérique d'intégration adéquate pour ces grilles. La TFM, dans sa version « continue », fournit une représentation complète (et donc unique) des images. Dans sa version numérique, il convient de connaître quelle est l'influence des différentes approximations, notamment la troncature d'un spectre, sur la représentation des objets. Pour vérifier la validité des différentes approximations, nous

illustrons les algorithmes proposés en présentant les résultats de l'échantillonnage, l'amplitude des spectres de Fourier-Mellin et la reconstruction d'une image à niveaux de gris représentant un papillon.

2.2.1 Fonctions et spectres discrets

Un objet à niveaux de gris est considéré comme une fonction $f(p, q)$ définie sur une grille cartésienne et d'étendue finie (le support de l'objet ne dépasse pas le cadre de l'image). Le centre du repère est fixé au barycentre de l'objet, ce qui permet de s'affranchir de la position de l'objet dans l'image. Comme le support de f est borné, il est important de noter que sa transformée de Fourier-Mellin s'étend, en principe, à l'infini. Cependant, dans la pratique, nous ne pouvons travailler qu'avec un nombre fini de coefficients, et la troncature du spectre est inévitable pour son implémentation. En conséquence une partie de l'information originale sur l'objet est perdue. Néanmoins, pour les images, l'amplitude de la TFM est évanescence lorsqu'on s'éloigne du centre du repère². Le théorème de Riemann-Lebesgue (équation (2.8)) précise que la TFM s'annule à l'infini. En prenant en compte un spectre plus important, l'information perdue par troncature du spectre peut être aussi faible que souhaitée.

Le spectre numérique de Fourier-Mellin d'une image, noté $\widehat{\mathcal{M}}_{f_\sigma}(k, v)$, est calculé pour $k \in [-K, K]$, $v \in [-V, V]$, avec un pas d'échantillonnage selon l'axe v fixé à 1. La valeur du paramètre de convergence n'est pas fixée a priori et sera étudiée dans certaines expériences. Notons toutefois que la valeur $\sigma = \frac{1}{2}$ est intéressante car elle donne un poids équitable aux pixels situés au centre des coordonnées par rapport aux pixels situés à la périphérie de l'objet [58]. C'est la valeur par défaut à moins que l'expérience numérique ne précise explicitement une autre valeur. Comme la TFM est symétrique (cf. paragraphe 2.1.2), la taille effective d'un spectre numérique est donc de $S_{K,V} = \frac{(2K+1)(2V+1)+1}{2}$.

Finalement, la reconstruction d'une image numérique à partir d'un spectre fini et discret $\widehat{\mathcal{M}}_{f_\sigma}(k, v)$ est notée par $\widehat{f}_{K,V}(p, q)$. Le choix des valeurs de K et V peut être guidé par le contenu spectrale de l'image ainsi que le type d'application, mais aucune règle générale ne peut être définie. Par exemple, les dimensions d'un spectre peuvent être inférieures pour une application de reconnaissance de formes que pour une application où la reconstruction est nécessaire³.

Dans la suite, nous présentons trois algorithmes d'approximation des TFM directe et inverse fondés sur la réécriture des équations (2.2) et (2.3). Le premier est une approximation directe de l'intégrale de Fourier-Mellin par échantillonnage polaire de l'image. Le second est basé sur un ré-échantillonnage log-polaire et le calcul de transformées de Fourier. Ces deux méthodes ont déjà été utilisées pour l'approximation directe de la TFM standard. Nous les adaptons ici au cas de la TFMA. Enfin, le dernier algorithme est original et s'exprime directement à partir des coordonnées cartésiennes de l'image. Les résultats des différents algorithmes sont illustrés à l'aide du papillon de la figure 2.3. Il faut noter que cette image est particulièrement complexe car la texture des ailes présente de forts contrastes sur de petites zones non régulières. Les techniques de ré-échantillonnage et d'intégration numérique utilisées sont linéaires. Les résultats expérimentaux seraient sans doute améliorés en utilisant des méthodes

²— C'est une extension du résultat classique de la transformée de Fourier des fonctions non singulières.

³— Dans le chapitre 3, nous étudions l'influence de la troncature des spectres sur le taux de reconnaissance et de discrimination des formes.

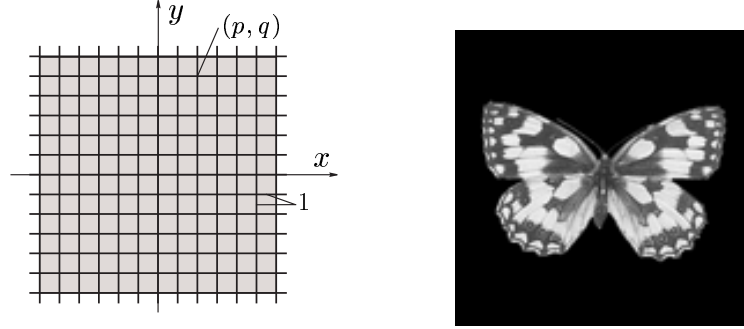


Figure 2.3 — Représentation de la grille cartésienne et image du papillon servant à illustrer les trois méthodes d'approximation.

d'ordre supérieur. Néanmoins, pour juger du potentiel de chaque approximation à estimer la TFMA, il est surtout essentiel d'utiliser des méthodes semblables.

2.2.2 Approximation directe par échantillonnage polaire

L'*approximation directe* de la TFMA (TFMA-D) consiste à ré-échantillonner l'image $f(p, q)$ selon une grille polaire et à estimer directement les intégrales de Fourier-Mellin, qui sont rappelées ici pour commodités :

$$\begin{cases} \mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r} \\ f(r, \theta) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) r^{-\sigma+iv} e^{ik\theta} dv. \end{cases} \quad (2.9)$$

Les différentes étapes de l'algorithme, que nous présentons maintenant, sont illustrées dans la figure 2.4 à la page suivante.

La grille d'échantillonnage polaire est construite à partir de l'intersection de M cercles concentriques régulièrement espacés, avec N rayons dont l'origine se situe au centre des cercles. Les pas d'échantillonnage radial et angulaire valent respectivement :

$$\Delta\theta = \frac{2\pi}{M}, \quad \Delta\rho = \frac{R}{N},$$

où R est le rayon du plus petit disque nécessaire pour contenir entièrement l'objet.

En général, la position des points polaires ne correspond pas à des coordonnées cartésiennes discrètes. Dans ce cas, la valeur de niveaux de gris du point polaire est estimée à l'aide d'une interpolation bi-linéaire utilisant les 4 points voisins de la grille cartésienne. Ainsi, la représentation polaire d'une image est une matrice de dimensions $[N, M]$ dont les valeurs correspondent à :

$$\hat{f}(\hat{\rho}_n, \hat{\theta}_m), \quad \text{avec } n \in [0, N - 1] \text{ et } m \in [0, M - 1],$$

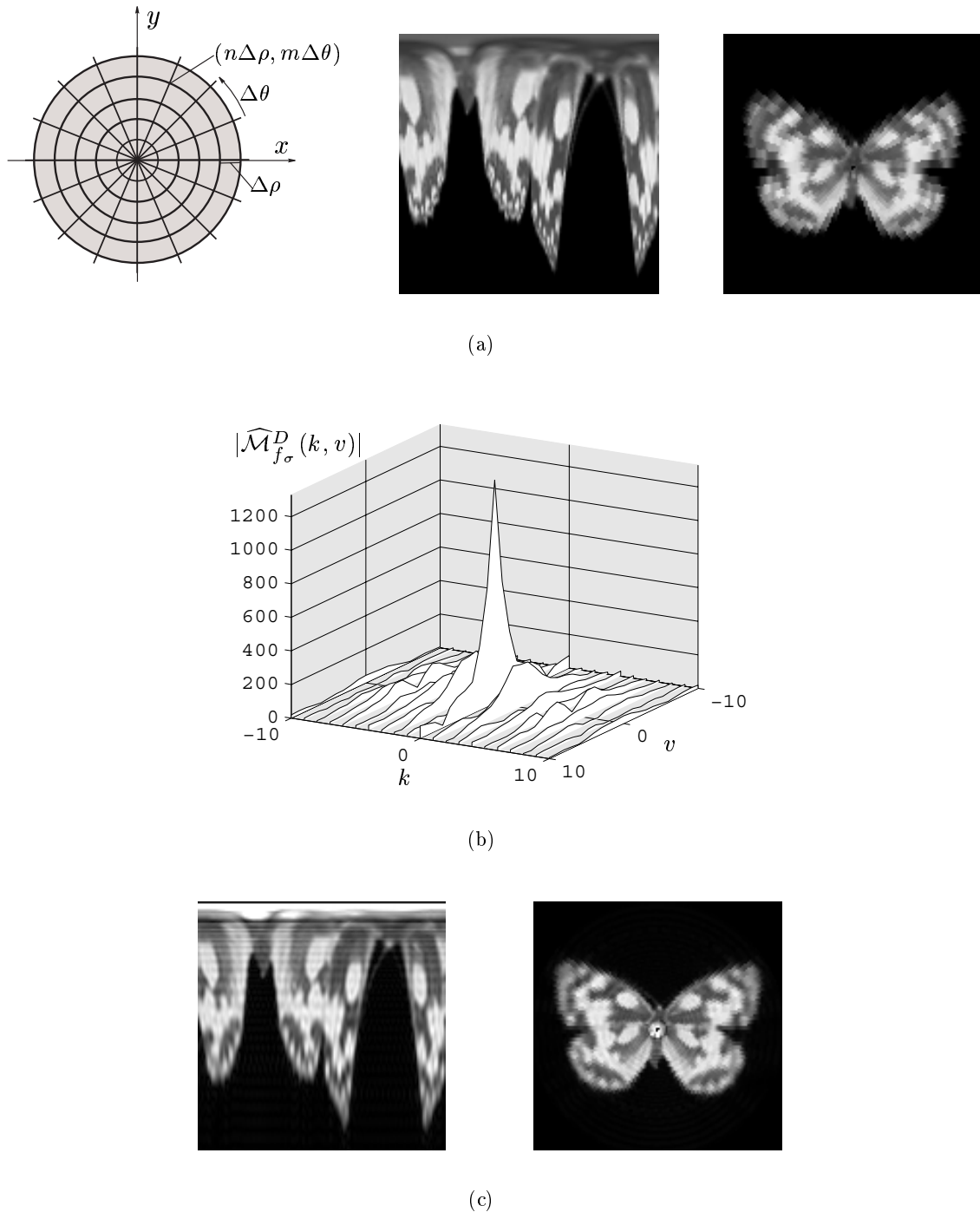


Figure 2.4 — Illustration de l'algorithme direct d'approximation des TFMA directe et inverse. (a) Grille polaire d'échantillonnage, représentation polaire du papillon ($N = M = 128$) et effet de l'échantillonnage polaire ($N = M = 64$) par interpolation bi-linéaire; (b) Amplitude du spectre de la TFMA avec $K = V = 10$ et $\sigma = 0,5$; (c) Reconstructions de l'image polaire et cartésienne avec $N = M = 128$ et $K = V = 60$, soit $S_{K,V} = 7321$ harmoniques.

où $\hat{\rho}_n = n\Delta\rho$ et $\hat{\theta}_m = m\Delta\theta$ désignent respectivement les rayons d'indice n et les angles d'indice m de la grille polaire (cf. figure 2.4(a)).

Les fonctions harmoniques circulaires de l'image sont estimées par la transformée de Fourier discrète de chaque cercle :

$$\hat{\mathcal{F}}_k(\hat{\rho}_n) = \Delta\theta \sum_{m=0}^{M-1} \hat{f}(\hat{\rho}_n, \hat{\theta}_m) e^{-i\frac{km}{M}}, \quad (2.10)$$

L'intégration suivant les rayons est réalisée en remplaçant l'intégrale de Mellin par des sommes sur les TFD de l'image⁴. Nous obtenons finalement l'approximation de la TFMA directe $\widehat{\mathcal{M}}_{f_\sigma}^D$ de f :

$$\widehat{\mathcal{M}}_{f_\sigma}^D(k, v) = \Delta\rho\Delta\theta \sum_{n=0}^{N-1} \hat{\mathcal{F}}_k(\hat{\rho}_n) (\hat{\rho}_n)^{\sigma-iv-1}, \quad (2.11)$$

pour tout $k \in [-K, K]$ et $v \in [-V, V]$. La figure 2.4(b) représente l'amplitude des composantes harmoniques centrales du spectre de Fourier-Mellin calculées grâce à l'algorithme TFMA-D.

L'algorithme réciproque qui consiste à reconstruire une image $f(p, q)$ à partir de sa représentation discrète de Fourier-Mellin $\widehat{\mathcal{M}}_{f_\sigma}^D(k, v)$ par l'estimation de la TFMA inverse, est similaire à l'approximation directe. Les étapes de l'algorithme sont les suivantes. D'abord, les fonctions harmoniques circulaires sont reconstruites en estimant la transformée de Mellin inverse. Puis, en calculant la TFD inverse, on restitue la représentation polaire de l'image $\hat{f}_{K,V}^D(\hat{\rho}_n, \hat{\theta}_m)$. Enfin, par un ré-échantillonnage polaire-cartésien, on reconstruit l'image dans sa représentation originale. Notons que l'algorithme de la FFT peut être utilisé pour accélérer les temps de calcul si le nombre de points de description circulaire M est une puissance de 2.

La reconstruction de l'image du papillon est illustrée dans la figure 2.4(c). L'aspect global de l'image restituée est de qualité correcte. Cependant, on observe également que les niveaux de gris divergent sur un disque de petite taille situé au centre des coordonnées. L'erreur provient essentiellement de l'approximation de la transformée de Mellin directe. Pour $k = 0$ et v quelconque, le spectre montre des valeurs sensiblement différentes de celles obtenues avec les deux algorithmes présentés plus loin. Lorsque v grandit, l'amplitude des harmoniques a tendance à diverger. Ces valeurs sont à l'origine du trou central car nous avons pu constater que le diamètre du disque diminuait légèrement lorsqu'on réduisait le pas d'échantillonnage radial (augmentant ainsi la précision de l'estimation de la TM). Ainsi, plus l'échantillonnage radial est fin, plus le trou dans l'image diminue jusqu'à sa disparition totale. Néanmoins, en augmentant le nombre de points de description radial, on augmente sensiblement le temps de calcul.

La représentation polaire met également en exergue la présence de certaines ondulations circulaires dont la fréquence diminue au fur et à mesure que l'on s'écarte du centre de l'objet. Ce phénomène fait penser aux ondulations que l'on obtient en reconstruisant un signal dont on

⁴— Dans les algorithmes, nous avons utilisé une méthode d'intégration par trapèzes. Ainsi, nous supposons que les données discrètes suivant l'axe v sont connectées par des segments de droite (interpolation linéaire).

a supprimé abruptement les hautes fréquences (présence de lobes en forme de sinus cardinaux). Ce phénomène justifie également le flou sur le contour externe de l'objet. Il faut également ajouter les erreurs dues à la double conversion des coordonnées par interpolation bilinéaire.

2.2.3 Approximation rapide par échantillonnage log-polaire

L'*approximation rapide* de la TFMA (TFMA-F) consiste à ré-échantillonner l'image $f(p, q)$ selon une grille log-polaire et à estimer les transformées de Fourier résultantes. C'est sans doute l'algorithme le plus utilisé pour l'approximation de la TFM standard. Nous l'adaptions ici au cas du prolongement analytique. En utilisant le fait que la transformée de Mellin peut se réécrire sous la forme d'une transformée de Fourier par changement de la variable d'intégration de r à $t = \ln(r)$, les transformées de Fourier-Mellin peuvent se réécrire sous la forme de deux transformées de Fourier :

$$\begin{cases} \mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_0^{2\pi} e^{t\sigma} f(e^t, \theta) e^{-i(k\theta + tv)} d\theta dt \\ f(e^t, \theta) = e^{-t\sigma} \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) e^{i(k\theta + tv)} dv. \end{cases} \quad (2.12)$$

de l'image déformée $e^{t\sigma} f(e^t, \theta)$. Les différentes étapes de l'algorithme, que nous présentons maintenant, sont illustrées dans la figure 2.5 à la page suivante.

La conversion numérique des coordonnées cartésiennes vers les coordonnées log-polaires est réalisée en utilisant un algorithme similaire à celui-ci présentée précédemment, excepté pour les cercles qui sont espacés de manière exponentielle (cf. Figure 2.5(a)). La transformée de Fourier-Mellin directe est alors estimée par une double transformée de Fourier discrète sur l'image déformée :

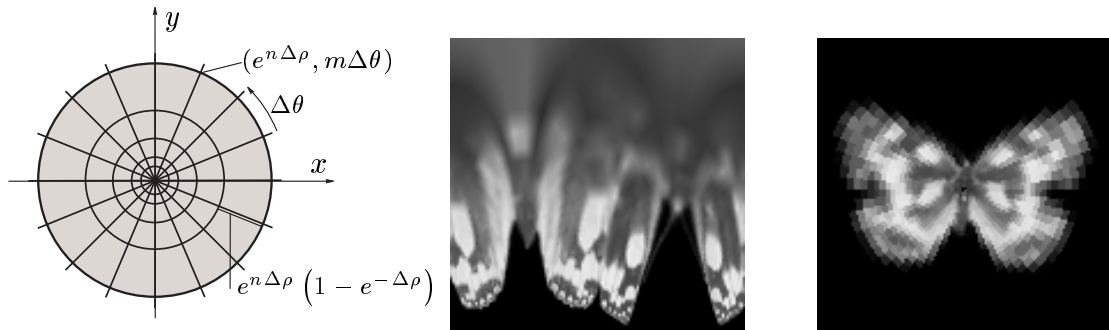
$$\widehat{\mathcal{M}}_{f_\sigma}^F(k, v) = \Delta\theta \sum_{n=0}^{N-1} e^{\sigma\hat{\rho}_n} \sum_{m=0}^{M-1} \widehat{f}(e^{\hat{\rho}_n}, \hat{\theta}_m) e^{-i(k\hat{\theta}_m + v\hat{\rho}_n)}. \quad (2.13)$$

Nous en verrons les conséquences sur l'image reconstruite. La figure 2.5(b) représente l'amplitude des composantes harmoniques centrales du spectre de Fourier-Mellin estimées grâce à l'algorithme de la TFMA-F. On constate que l'amplitude du spectre converge rapidement vers 0, même pour $k = 0$.

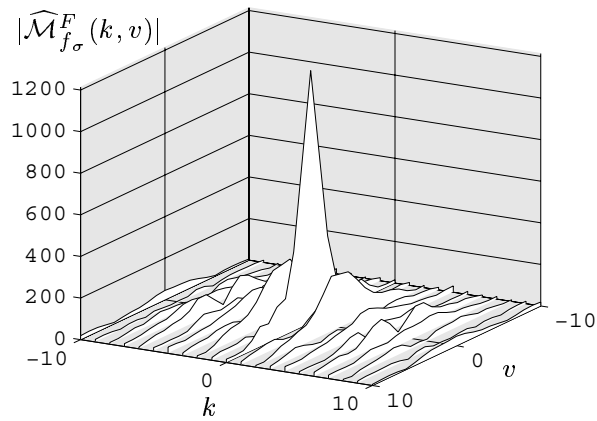
Depuis l'équation (2.12), la TFM inverse est estimée en calculant une TFD bi-dimensionnelle inverse sur le spectre de Fourier-Mellin. Ainsi, nous reconstruisons l'image déformée en coordonnées log-polaires :

$$\widehat{f}_{K,V}^F(e^{\hat{\rho}_n}, \hat{\theta}_m) = e^{-\sigma\hat{\rho}_n} \sum_{v=-V}^V \sum_{k=-K}^K \widehat{\mathcal{M}}_{f_\sigma}^F(k, v) e^{i(k\hat{\theta}_m + v\hat{\rho}_n)}. \quad (2.14)$$

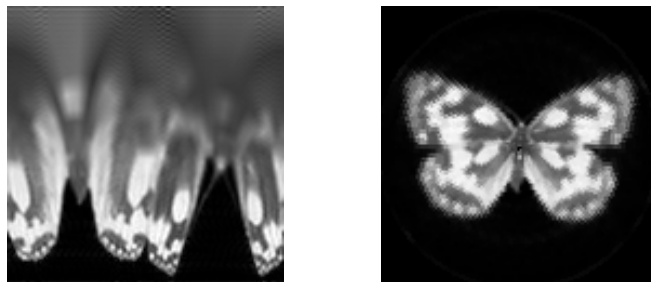
La dernière étape est la conversion inverse des coordonnées pour retrouver la représentation cartésienne. Si le nombre de points radiaux et angulaires sont des puissances de 2, alors



(a)



(b)



(c)

Figure 2.5 — Illustration de l'algorithme rapide d'approximation des TFMA directe et inverse. (a) Grille log-polaire d'échantillonnage, représentation log-polaire du papillon ($N = M = 128$) et effet de l'échantillonnage log-polaire ($N = M = 64$) par interpolation bilinéaire; (b) Amplitude du spectre de la TFMA avec $K = V = 10$ et $\sigma = 0,5$; (c) Reconstitutions de l'image log-polaire et cartésienne avec $N = M = 128$ et $K = V = 60$, soit $S_{K,V} = 7321$ harmoniques.

on peut utiliser une double FFT pour approcher la TFMA. De manière symétrique dans l'algorithme inverse, $(2K + 1)$ et $(2V + 1)$ sont des puissances de 2 et la FFT peut également être utilisée pour réduire les temps de calcul. Ceci justifie le qualificatif d'approximation rapide de la méthode.

La figure 2.4(c) présente l'image reconstruite du papillon dans ses coordonnées log-polaires et cartésiennes. La principale différence avec l'algorithme précédent est la disparition du disque centrale. Le calcul de la TFD sur les coordonnées logarithmiques semble être une meilleure approximation de la TM que l'approximation par intégration directe. Par contre, l'aspect flou de l'image est légèrement plus important, surtout à la périphérie de l'objet. Outre les effets de la troncature abrupte du spectre, la principale raison provient de l'échantillonnage log-polaire. Les zones situées à la périphérie de l'objet sont sous-échantillonnées et la double conversion des coordonnées par interpolation bi-linéaire reconstruit une image d'aspect lissée où les zones de forts gradients sont moins saillantes.

Notons finalement que l'utilisation d'une transformée de Fourier discrète pour estimer la transformée de Fourier sur \mathbb{R} implique que nous considérons l'image comme périodique selon toutes les directions radiales. Les effets d'une telle hypothèse sur la reconstruction sont mieux observés dans les expériences du paragraphe 2.3.

2.2.4 Approximation cartésienne sans ré-échantillonnage préalable

La dernière méthode d'estimation est appelée *approximation cartésienne* de la TFMA (TFMA-C). Il est en effet possible d'effectuer un changement des variables d'intégration dans l'équation (2.9) qui aboutit directement à des expressions des transformées de Fourier-Mellin en coordonnées cartésiennes :

$$\begin{cases} \mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) (x + iy)^{-k} (x^2 + y^2)^{\frac{k-2+\sigma-iv}{2}} dx dy \\ f(x, y) = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \mathcal{M}_{f_\sigma}(k, v) (x + iy)^k (x^2 + y^2)^{\frac{-k-\sigma+iv}{2}} dv. \end{cases} \quad (2.15)$$

Dans ce cas, aucun ré-échantillonnage n'est nécessaire et les TFMA sont estimées directement à partir de la grille rectangulaire.

Les approximations directe et inverse sont calculées en remplaçant les intégrales par des sommes :

$$\begin{cases} \widehat{\mathcal{M}}_{f_\sigma}^C(k, v) = \frac{1}{2\pi} \sum_{q=Q_{min}}^{Q_{max}} \sum_{p=P_{min}}^{P_{max}} f(p, q) (p + iq)^{-k} (p^2 + q^2)^{\frac{k-2+\sigma-iv}{2}} \\ \widehat{f}_{K,V}^C(p, q) = (p^2 + q^2)^{\frac{-\sigma}{2}} \sum_{v=-V}^V \sum_{k=-K}^K \widehat{\mathcal{M}}_{f_\sigma}^C(k, v) (p + iq)^k (p^2 + q^2)^{\frac{-k+iv}{2}}. \end{cases} \quad (2.16)$$

Les coordonnées p et q sont centrées par rapport au barycentre de l'image et P_{min} , P_{max} , Q_{min} , et Q_{max} désignent le plus petit rectangle englobant de l'image. Pour des raisons de compati-

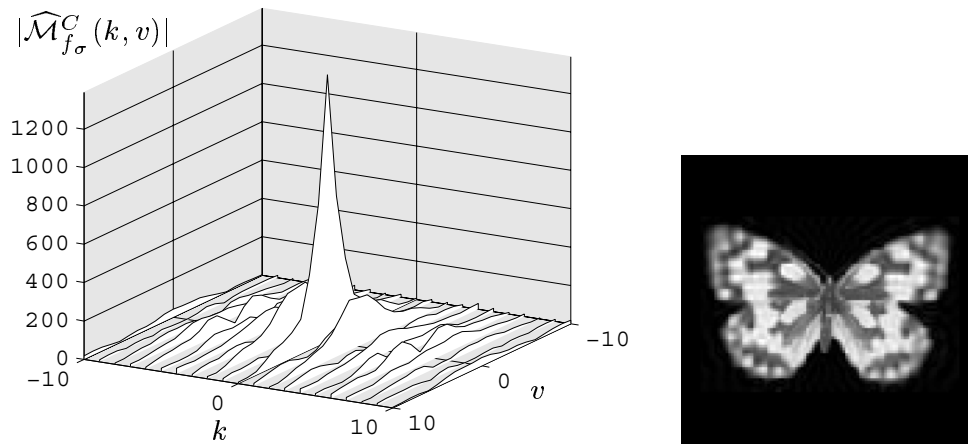


Figure 2.6 — Illustration de l'algorithme cartésien d'approximation des TFMA directe et inverse. À gauche : Amplitude des composantes harmoniques centrales du spectre de la TFMA avec $K = V = 10$ et $\sigma = 0,5$. À droite : Image reconstruite du papillon avec $K = V = 60$, soit $S_{K,V} = 7321$ harmoniques.

bilité avec les méthodes précédentes, les intégrales sont estimées par la technique des trapèzes.

L'image discrète est reconstruite directement en coordonnées cartésiennes. La figure 2.6 montre le spectre de Fourier-Mellin et la reconstruction d'un papillon obtenus grâce à l'algorithme TFMA-C. L'intégration des fonctions ne pose pas de réelles difficultés numériques, et, pour un spectre de taille comparable, la reconstruction semble être la meilleur des trois. Notamment, les contours externe du papillon sont plus nets. Par contre, cette approximation est la plus lente des trois car aucun algorithme rapide n'est disponible. Un léger aspect flou persiste à la périphérie de l'objet comme conséquence de la troncature abrupte des harmoniques hautes fréquences. L'ajout de composantes supplémentaires devraient réduire ces effets. C'est ce que nous nous proposons d'étudier dans le paragraphe suivant.

2.3 Reconstruction des objets : comparaison des approximations

Une fonction peut être, en principe, parfaitement reconstruite à partir de sa TFM. Nous avons vu dans le paragraphe précédent que les reconstructions obtenues à partir d'un spectre discret de dimensions finies sont altérées par différents phénomènes numériques. Certaines de ces dégradations liées à la troncature du spectre peuvent être progressivement gommées en utilisant la contribution d'harmoniques supplémentaires. Il est donc intéressant de connaître l'influence individuelle des dimensions de Fourier et Mellin pour la reconstruction des images.



Figure 2.7 — Images de test utilisées pour estimer l'erreur de reconstruction.

Pour cela, nous calculons l'Erreur Quadratique Moyenne Normalisée (EQMN), donnée par :

$$\varepsilon^2(\hat{f}_{K,V}) = \frac{\sum_p \sum_q |f(p, q) - \hat{f}_{K,V}(p, q)|^2}{\sum_p \sum_q f(p, q)^2}, \quad (2.17)$$

où $\hat{f}_{K,V}$ désigne une version reconstruite de l'image $f(p, q)$, obtenue par l'une des trois méthodes d'approximation, et pour un spectre de dimensions $K \times V$. Lorsque l'erreur de reconstruction est nulle, la reconstruction de l'image originale est parfaite. Ce n'est bien sûr jamais exactement le cas et cette erreur fournit une mesure de la fidélité de reconstruction des algorithmes.

Les expériences reportées dans la suite utilisent, d'une part, l'image de la lettre typographique « F » (niveau de gris constant égale à 128 sur un fond noir), et d'autre part, l'image d'un second papillon (cf. figure 2.7). En fonction des paramètres K et V , nous présentons un premier tableau comprenant les images reconstruites $\hat{f}_{K,V}$, et un second tableau représentant les images obtenues par différence entre l'image originale et ses versions reconstruites, augmentées de 128 : $f - \hat{f}_{K,V} + 128$. Ces tableaux nous permettent de constater visuellement du gain de reconstruction obtenu par la contribution d'harmoniques supplémentaires. Enfin, un graphique 3D, présente les résultats du calcul de l'erreur moyenne quadratique $\varepsilon(\hat{f}_{K,V})$ sur les images du 1^{er} tableau.

2.3.1 Expériences sur l'image d'une lettre binaire

La figure 2.8 présente les différents résultats que nous avons obtenus sur la lettre F à partir de l'approximation cartésienne de la TFMA. Comme attendu, l'erreur de reconstruction diminue au fur et à mesure que les dimensions du spectre augmentent. Lorsque la taille du spectre est peu importante, le tableau de l'image reconstruite montre l'influence radiale de la transformée de Mellin et l'influence circulaire de la transformée de Fourier sur la reconstruction. En augmentant les deux dimensions, ces influences se combinent pour donner une image reconstruite qui finalement apparaît de très bonne qualité pour un spectre large.

Le graphique sur l'erreur de reconstruction corrobore ces résultats. Le gain est très important jusqu'à ce que K et V atteignent la valeur 0,15 environ (pente de la surface importante). Puis, le gain de reconstruction est de plus en plus faible jusqu'à atteindre la valeur 0,006 pour

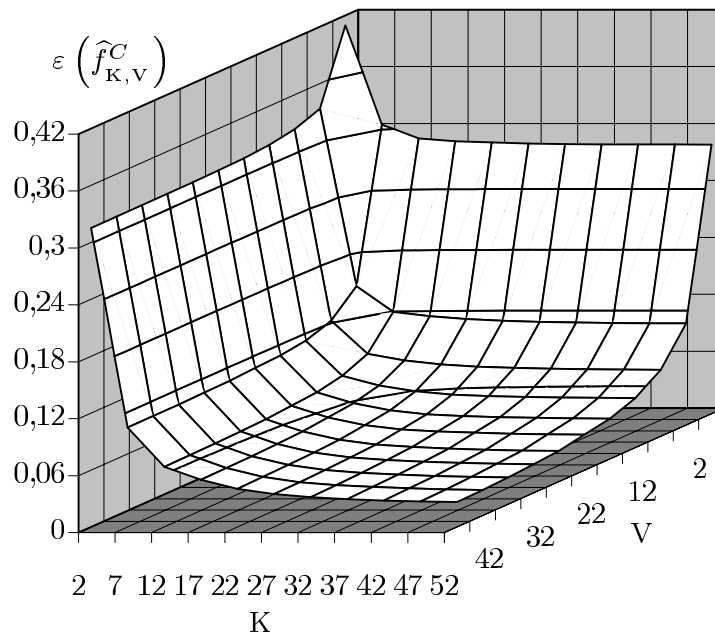
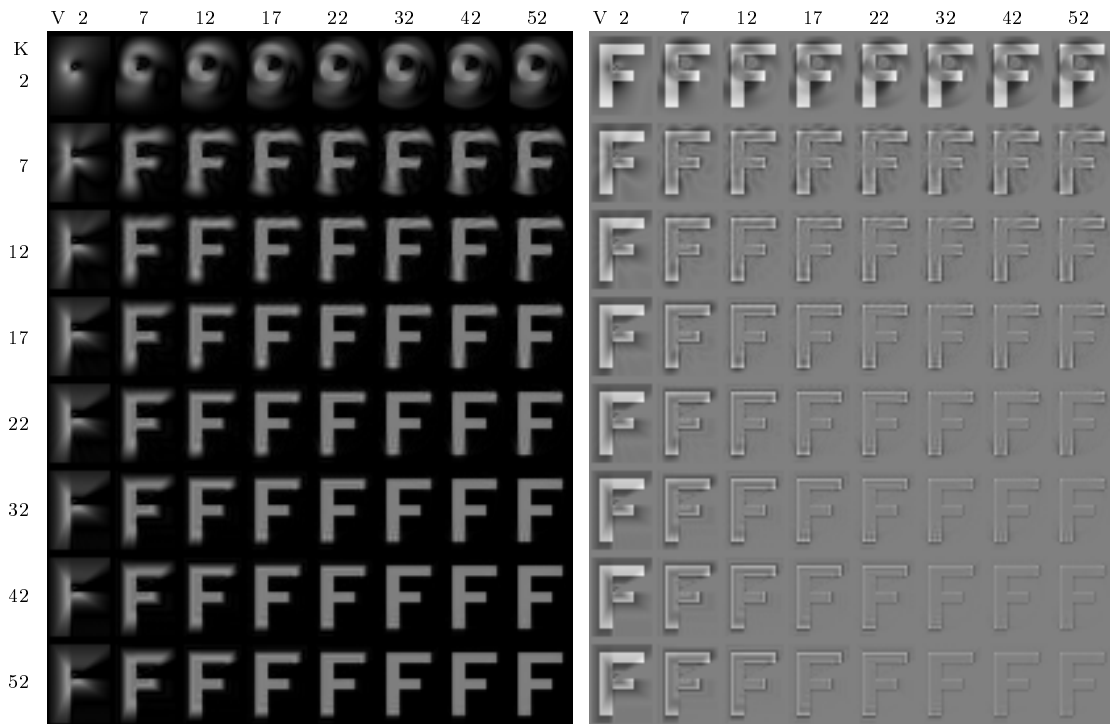


Figure 2.8 — Reconstruction de l'image de la lettre F en fonction de K et V (algorithme TFMA-C) En haut : tableaux des images reconstruites; en bas : graphique 3D représentant l'erreur de reconstruction.

$K = V = 52$. Il est important de noter que la contribution de la transformée de Mellin a une importance égale à celle de la transformée de Fourier, contrairement à certains commentaires émis lors de l'utilisation de la TFM en reconnaissance de formes.

2.3.2 Expériences sur l'image d'un papillon

Les figures 2.9, 2.10 et 2.11 représentent respectivement les tableaux de reconstruction du papillon la figure 2.7 pour les algorithmes TFMA-D, TFMA-F et TFMA-C. La figure 2.12 présente quant à elle l'erreur de reconstruction, calculée à partir de ces tableaux d'images, pour chacune de trois méthodes.

Les résultats de reconstruction du papillon confortent ceux obtenus à partir de la lettre F. Au fur et à mesure que les dimensions des spectres augmentent, les contours et la texture des ailes du papillon se précisent pour aboutir finalement à une image reconstruite de bonne qualité. Les images de différences mettent en exergue la présence de halos lorsque le spectre est de dimensions trop faibles. Ces halos diminuent lorsque K et V deviennent importants et les images de différences sont de plus en plus estompées, montrant le gain en terme de reconstruction.

En ce qui concerne l'approximation directe de la TFMA, on retrouve une zone non-reconstruite au centre des coordonnées polaires lorsque l'ordre V est important. On constate d'ailleurs sur les graphes (a) et (d) de la figure 2.12 que l'erreur de reconstruction augmente lorsque V est supérieur à 44, ce qui tente à montrer que le problème provient essentiellement de l'approximation de la transformée de Mellin.

Pour les algorithmes TFMA-F et TFMA-C, les résultats de reconstruction ne présentent pas de différences visuelles notables. Seule l'erreur de reconstruction présentée dans la figure 2.12(d) semble montrer que l'algorithme TFMA-C fournit une meilleur approximation des harmoniques d'ordres élevées.

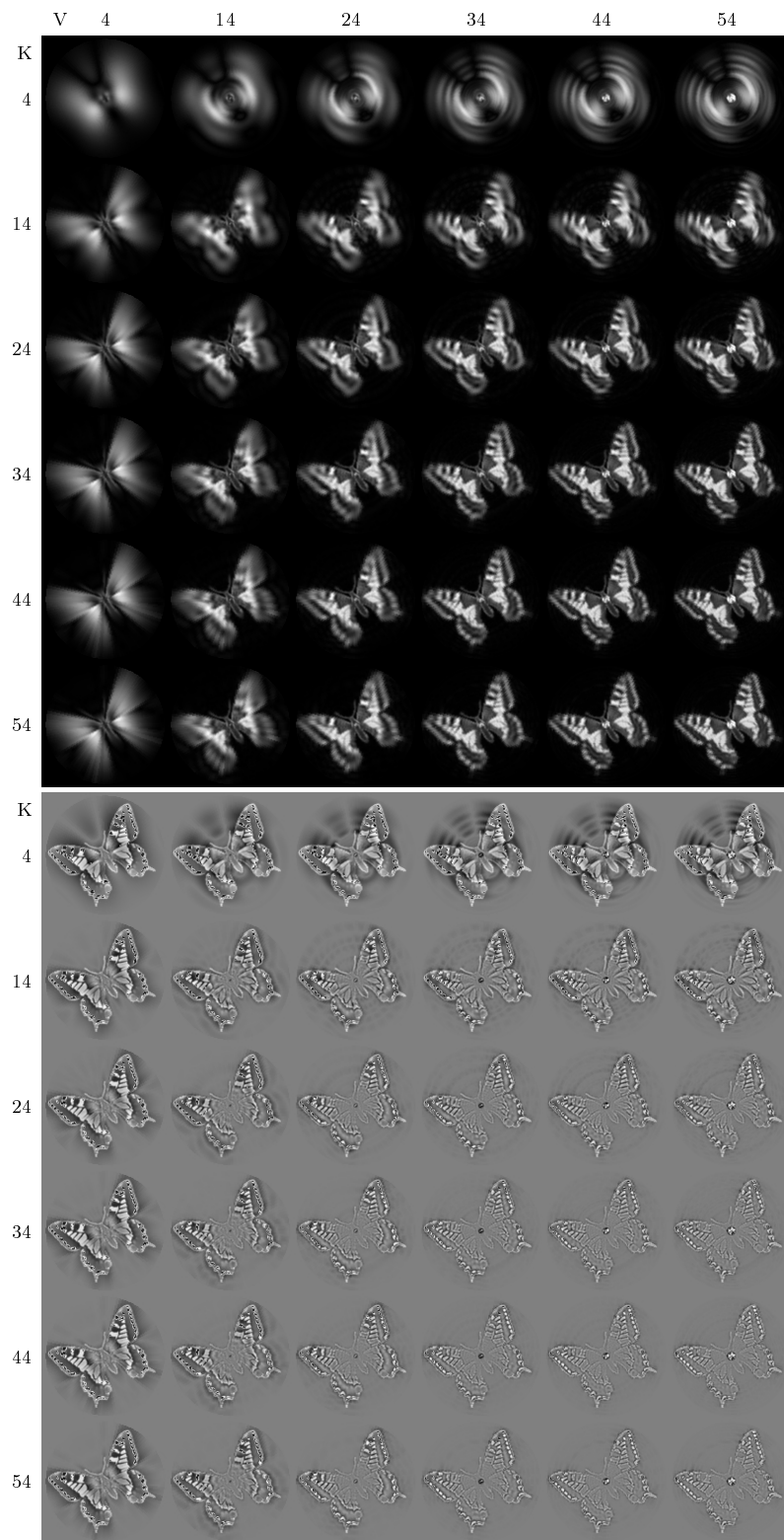


Figure 2.9 — Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-D). En haut : images reconstruites ; en bas : différence avec l'image originale rehaussée de 128.

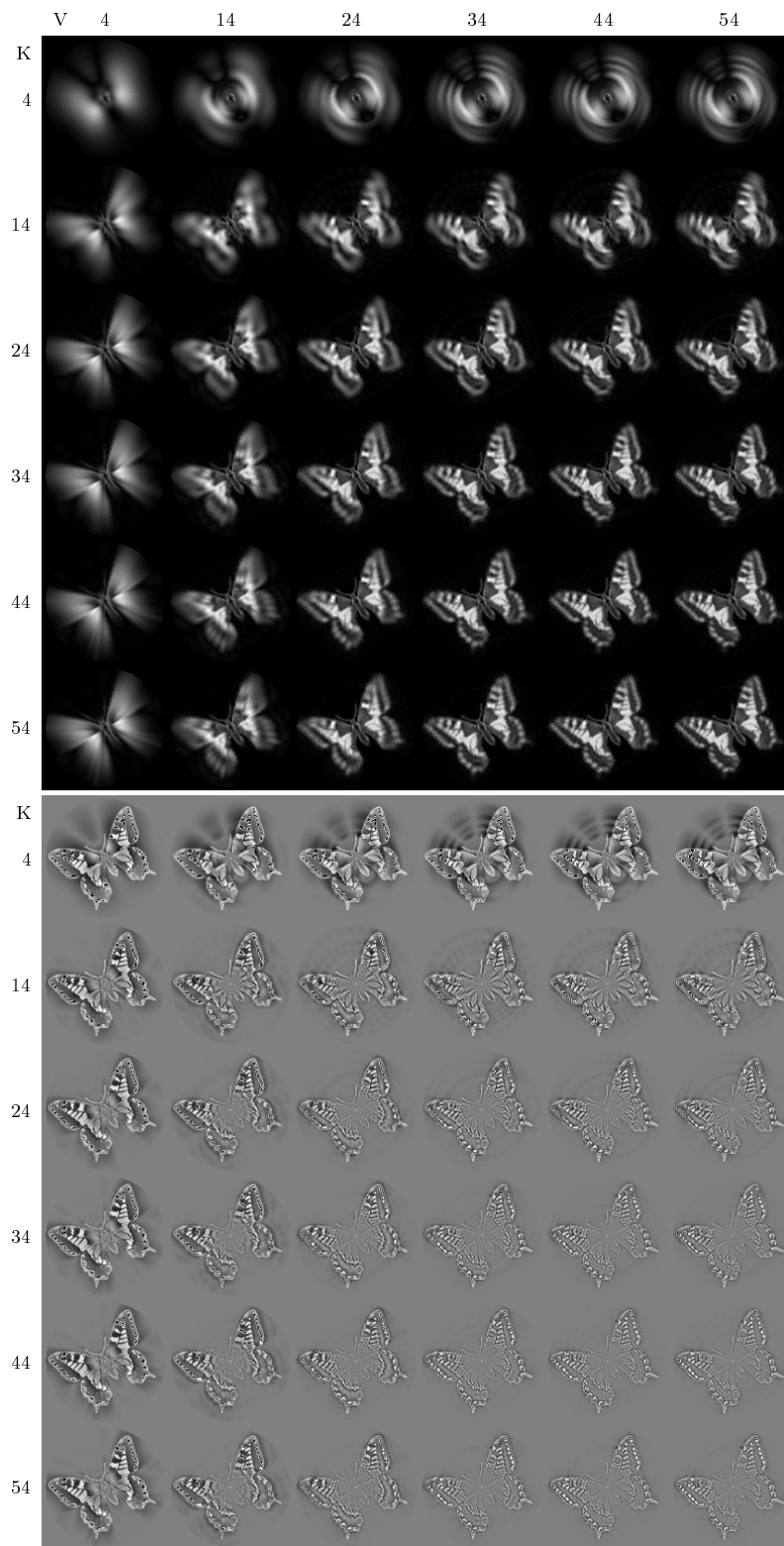


Figure 2.10 — Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-F). En haut : images reconstruites ; en bas : différence avec l'image originale rehaussée de 128.

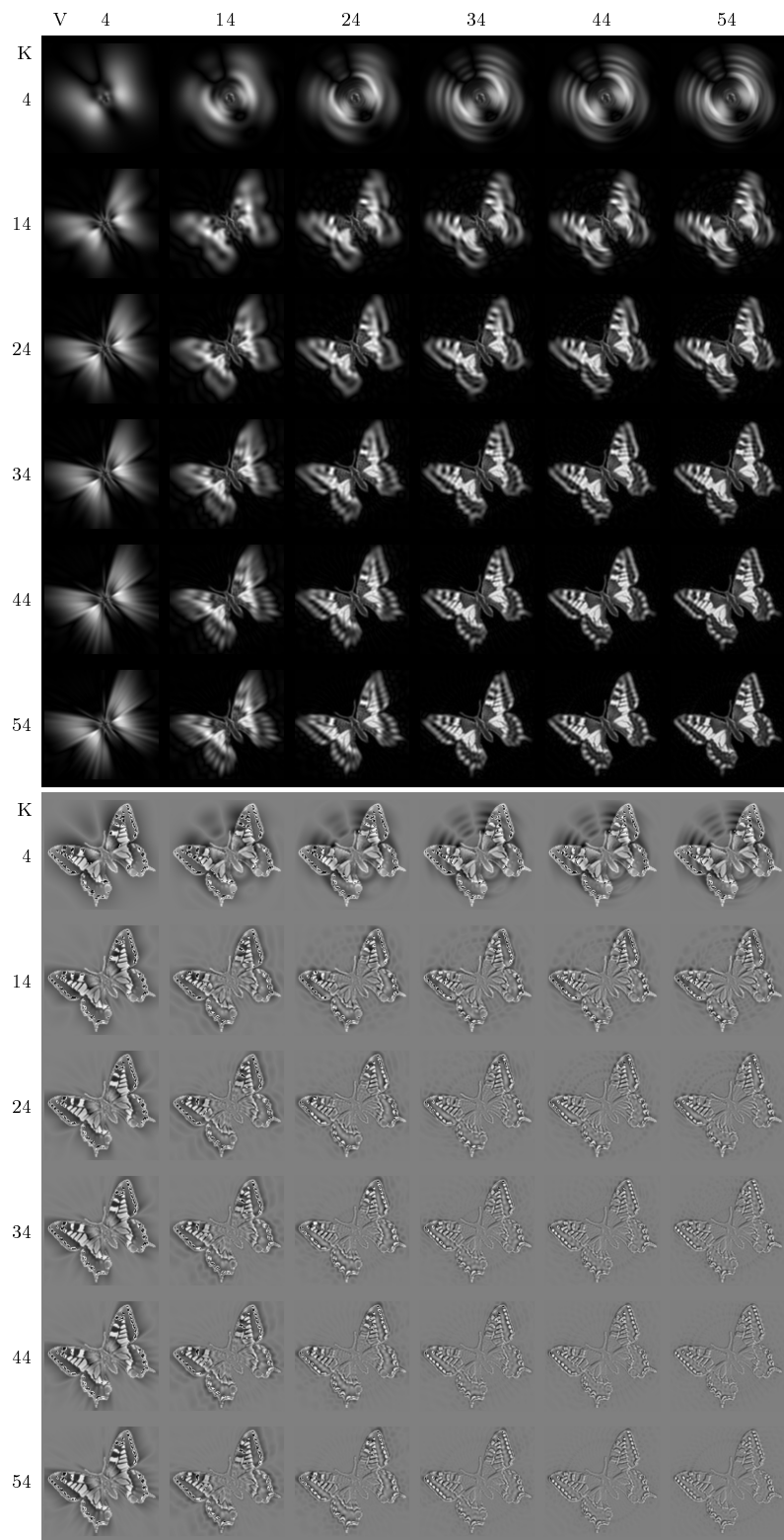


Figure 2.11 — Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-C). En haut : images reconstruites ; en bas : différence avec l'image originale rehaussée de 128.

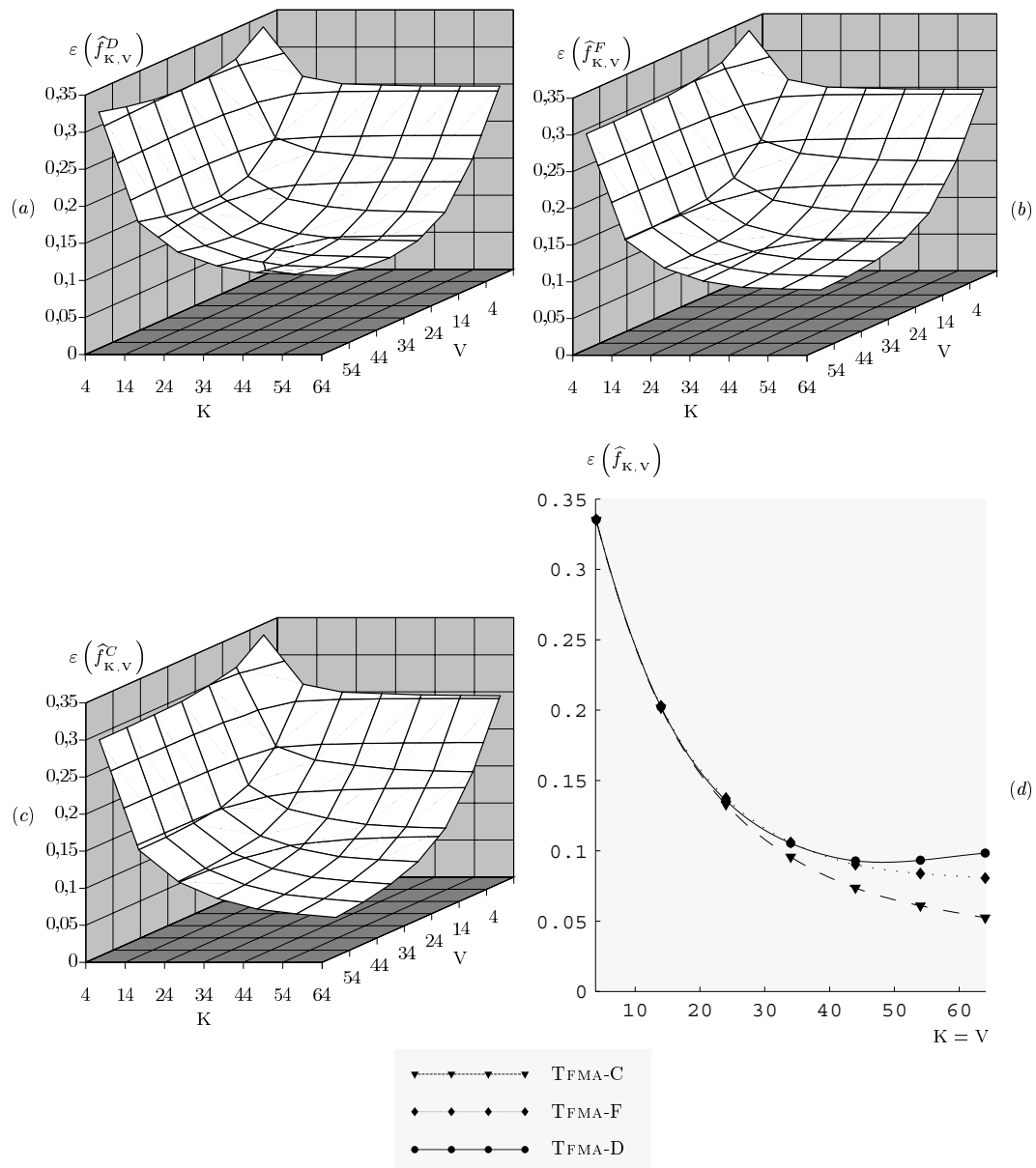


Figure 2.12 — Représentation de l'erreur de reconstruction pour l'image du papillon de la figure 2.7 en fonction des trois méthodes d'approximation. (a) algorithme TFMA-D (fig. 2.9) ; (b) algorithme TFMA-F (fig. 2.10) ; (c) algorithme TFMA-C (fig. 2.11) ; (d) erreur de reconstruction des trois méthodes pour $K = V$.

2.4 Conclusion du chapitre

La transformée de Fourier-Mellin analytique a été utilisée pour remédier aux problèmes théoriques et numériques de divergence de l'intégrale lorsque celle-ci est calculée sur des objets à niveaux de gris. Nous avons ainsi pu proposer trois algorithmes originaux pour estimer les TFMA directe et inverse. Ces algorithmes ont ensuite été testés et comparés par reconstruction des images à partir de leur spectre numérique. L'amélioration de l'estimation de la TFM nous a permis de mettre en évidence que la transformée de Mellin a un rôle d'une importance équivalente aux coefficients de Fourier dans la transformée globale.

Les trois algorithmes fournissent des résultats très satisfaisants. Cependant, l'approximation cartésienne TFMA-C présente des résultats de qualité légèrement supérieure, essentiellement du point de vue de l'erreur de reconstruction. Visuellement, la qualité des images reconstruites est comparable à celle des images obtenues avec l'approximation rapide de la TFMA. L'algorithme TFMA-F a l'avantage important d'être beaucoup plus rapide à calculer. Une certain nombre d'études peuvent compléter ce travail sur la reconstruction. Notamment, nous avons vu que la troncature abrupte des fréquences était à l'origine de la présence de halos dans les images restituées. Ainsi, une perspective consistera à étendre les méthodes de filtrage développées dans le cadre de la transformée de Fourier 2D en utilisant le lien qu'elle entretient avec la transformée de Fourier-Mellin analytique par le biais du changement de variables log-polaire.

La reconstruction d'un objet est un premier critère pour vérifier la robustesse et l'inversibilité numérique des trois approximations proposées. Cependant, ces expériences ne permettent pas de juger si les spectres numériques obtenus représentent effectivement la transformée de Fourier-Mellin d'un objet. Dans le chapitre suivant, la comparaison est poursuivie à l'aide d'expériences en reconnaissance de formes à partir de nouvelles familles de descripteurs invariants issues de la TFMA.

Reconnaissance Invariante des Formes par Similitudes

Comme nous l'avons souligné à plusieurs reprises dans le premier chapitre, les descripteurs invariants extraits de la transformée de Fourier-Mellin d'une image ne tiennent généralement compte que de l'amplitude de la TFM. L'information contenue dans la phase n'est, en général, pas utilisée. De tels invariants ne représentent que partiellement les objets, ce qui apparaît insuffisant et parfois critique pour les applications récentes des descripteurs de formes, notamment le codage orienté objet, le tatouage d'images et l'indexation de bases d'images par le contenu (cf. section 3.1). Pour résoudre ce problème, nous présentons un procédé original de normalisation permettant d'extraire des familles complètes de descripteurs invariants par similitude à partir du prolongement analytique de la transformée de Fourier-Mellin.

La section 3.2 présente quelques propriétés de la TFMA qui sont obtenues par extension des résultats classiques concernant la transformée de Fourier (notamment l'identité de Parseval et le théorème du retard). Nous aboutissons à une caractérisation précise des formes dans l'espace de représentation de Fourier-Mellin.

Un procédé original de normalisation de la TFMA est ensuite appliqué aux formes. Celui-ci permet de construire des familles complètes et intégrables de descripteurs globaux, invariants par similitude (section 3.3). Ces familles dépendent de deux paramètres de normalisation que l'on peut choisir parmi un large éventail de possibilités. Des critères numériques sont alors introduits dans la section 3.4 permettant de construire des descripteurs invariants robustes vis-à-vis des approximations et des faibles variations de formes.

Afin de mesurer le degré de similarité des objets indépendamment de leur position, de leur orientation et de leur taille dans les images, une distance entre formes est ensuite utilisée. Les performances de ces familles de descripteurs, au regard des trois approximations de la TFMA et de la robustesse des paramètres de normalisation, sont analysées à l'aide d'expériences de reconnaissance et de discrimination d'objets dans la section 3.5.

3.1 Unicité des descripteurs invariants

Les transformées de Fourier-Mellin standard de deux objets de même forme ne se distinguent que d'un facteur de phase qui contient à la fois l'orientation et la taille relatives aux objets (cf. paragraphe 1.3.2 page 21). Dans ce cas, nous avons vu qu'une technique classique d'extraction de descripteurs invariants par similitude consiste à ne considérer que le module de la TFM standard des objets [23, 59]. Nous avons déjà fait remarquer qu'un tel ensemble d'invariants n'est pas complet puisque l'information contenue dans la phase de la transformée n'est pas utilisée. De plus, étant donnés les problèmes numériques rencontrés pour estimer la transformée de Fourier-Mellin, la plupart des expériences de reconnaissance de formes n'utilise qu'un faible nombre d'harmoniques. Ainsi, les descripteurs invariants ne fournissent qu'une représentation superficielle de la forme des objets. La comparaison entre les descripteurs invariants est réalisée par des méthodes telles que le calcul d'une erreur quadratique entre les descripteurs [115], des réseaux de neurones [116, 127], ou bien encore des méthodes statistiques [114, 32, 59].

Ces méthodes peuvent s'avérer satisfaisantes dans le cas de bases d'images où les objets à reconnaître sont relativement simples et de formes bien distinctes, comme un alphabet de lettres d'imprimerie par exemple. Néanmoins, l'essor des systèmes d'acquisition (caméra numériques, scanner, IRM, ...), des capacités de stockage (mémoire de masse), des systèmes de transmission et de diffusion (réseau internet, réseaux spécialisés à hauts débits) a favorisé le développement de bases d'images volumineuses et de natures très diverses. Ceci pose le problème de l'exploitation et de la consultation de ces bases par le contenu (cf. chapitre 5). Dans le cas de bases spécialisées (bases d'images médicales ou biologiques, bases de visages), une légère distorsion peut s'avérer caractéristique de la forme d'un objet. Or, à cause de la non-unicité des descripteurs, des objets de formes différentes peuvent être décrits par des invariants identiques, ce qui peut engendrer des confusions dans le processus de classification de certaines images.

L'importance de l'information contenue dans la phase des coefficients de Fourier a été mis en évidence par T.R. Crimmins dans le cas d'objets-contours et du groupe des similitudes [29]. Il a montré, à l'aide d'un exemple, que les coefficients de Fourier de deux courbes polygonales fermées et de formes différentes peuvent avoir même amplitude. Ainsi, les invariants construits à partir des modules des coefficients de Fourier ne permettent pas de discriminer les deux contours. Pour remédier à ce problème, il a proposé une *famille complète* de descripteurs invariants tenant compte à la fois de la phase et du module des coefficients de Fourier. La représentation des formes est alors unique et assure théoriquement une discrimination totale des objets de formes différentes.

L'information contenue dans la phase de la TFMA peut être illustrée à l'aide d'une expérience simple¹ qui consiste à reconstruire une image à partir d'un spectre combinant l'amplitude de la TFMA d'une première image avec la phase d'une seconde. Les résultats sont illustrés dans la figure 3.1. Les deux images reconstruites montrent essentiellement l'objet dont on a utilisé la phase. Ceci met en évidence l'importance relative de la phase de la TFMA par rapport à son amplitude. On rencontre également plusieurs techniques d'appariement d'images utilisant exclusivement l'information contenue dans la phase de la transformée de

¹— Cette expérience s'inspire du travail réalisé avec la transformée de Fourier [95].

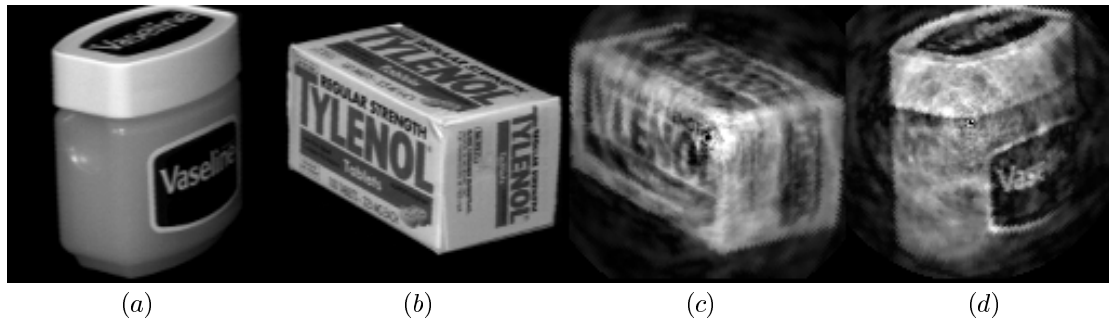


Figure 3.1 — Illustration de l'importance de la phase dans la représentation des images par la TFMA (algorithme TFMA-F). (a) et (b) : images originales utilisées pour l'expérience; (c) : image reconstruite à l'aide TFMA inverse de l'amplitude de (a) et la phase de (b); (d) : image reconstruite à l'aide TFMA inverse de l'amplitude de (b) et la phase de (a).

Fourier-Mellin (méthodes de flot optique [47] et de corrélation de phase [48, 26]).

Dans le cas de la description des objets par la transformée de Fourier-Mellin, le critère de complétude est difficile à satisfaire étant donné la complexité des outils mathématiques mis en œuvre. Une première famille complète de descripteurs invariants par similitudes planes a été proposée [79, 20]. Elle est construite à partir des dérivées d'ordre 1 et 2 de la TFM standard. Outre les problèmes de robustesse liés au calcul numérique des dérivées de l'image, cette méthode ne peut pas s'étendre à la TFMA et se heurte aux problèmes de divergence décrits dans la section 1.3 page 19. Parallèlement à ces travaux, F. Ghorbel a proposé une nouvelle famille complète de descripteurs invariants tenant compte du prolongement analytique de la transformée de Fourier-Mellin [53, 55]. Cette famille est également intégrable, ce qui permet de définir une vraie distance entre formes.

Dans la suite de ce chapitre, nous proposons d'étendre cette famille par un procédé de normalisation de la TFMA qui assure simultanément l'invariance, la complétude et la sommabilité des descripteurs obtenus. Le principe de la méthode consiste à transformer une image dans une « position standard »² qui offre une description invariante complète de tous les objets équivalents au sens de la transformation géométrique [104]. Le procédé de normalisation permet également de mettre en évidence et de clarifier le rôle de deux paramètres de normalisation que l'on peut choisir parmi un large éventail de possibilités. Ainsi, sur la base de critères numériques, nous proposons de nouvelles familles qui s'avèrent particulièrement robustes et discriminantes. Avant de présenter la méthode en détail, il est nécessaire de préciser quelques propriétés de la TFMA.

3.2 Représentation des formes par la TFMA

Dans cette section, nous examinons quelques propriétés de la TFMA qui nous seront utiles par la suite. En particulier, nous définissons l'égalité de Parseval et le théorème du retard adaptés à notre cas, comme extension des résultats classiques concernant la transformée de Fourier (cf. section 1.1 page 6). Ces deux propriétés nous permettent ensuite de caractériser

²— Canonical frame.

une forme comme une classe d'équivalence regroupant l'ensemble des objets à niveaux de gris similaires au sens du groupe \mathcal{G} .

3.2.1 L'égalité de Parseval sur le groupe des similitudes

En premier lieu, précisons quelques notations. $\mathbf{L}^2(\mathcal{G})$ désigne l'espace vectoriel des fonctions f définies sur \mathcal{G} et de carré intégrables par rapport à la mesure invariante $d\theta \frac{dr}{r}$, *i.e.*

$$f \in \mathbf{L}^2(\mathcal{G}) \Rightarrow \|f\|_{\mathbf{L}^2(\mathcal{G})} = \left(\int_0^\infty \int_0^{2\pi} |f(r, \theta)|^2 d\theta \frac{dr}{r} \right)^{\frac{1}{2}} < \infty. \quad (3.1)$$

De manière équivalente, $\mathbf{L}^2(\hat{\mathcal{G}})$ désigne l'espace vectoriel des fonctions \mathcal{M}_f définies sur $\hat{\mathcal{G}}$ et de carré intégrables par rapport à la mesure invariante $\delta k \otimes dv$, *i.e.*

$$\mathcal{M}_f \in \mathbf{L}^2(\hat{\mathcal{G}}) \Rightarrow \|\mathcal{M}_f\|_{\mathbf{L}^2(\hat{\mathcal{G}})} = \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_f(k, v)|^2 dv \right)^{\frac{1}{2}} < \infty. \quad (3.2)$$

Il est important de noter que si une fonction est intégrable, elle n'est pas nécessairement de carré intégrable et vice versa. Il n'y a pas, en général, de relation d'inclusion entre les espaces \mathbf{L}^1 et \mathbf{L}^2 . $\mathbf{L}^2(\mathcal{G})$ et $\mathbf{L}^2(\hat{\mathcal{G}})$ sont des espaces métriques pour les distances respectivement définies par : $d_2(f, g) = \|f - g\|_{\mathbf{L}^2(\mathcal{G})}$ et $d_2(\mathcal{M}_f, \mathcal{M}_g) = \|\mathcal{M}_f - \mathcal{M}_g\|_{\mathbf{L}^2(\hat{\mathcal{G}})}$.

Le *Théorème de Plancherel*, appliqué au cas du groupe des similitudes planes vectorielles, précise que la TFM des fonctions intégrables et de carré intégrables sur \mathcal{G} (espace $\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$) est une isométrie pour les normes \mathbf{L}^2 . Nous pouvons ainsi définir l'*égalité de Parseval* adaptée à notre cas :

$$\|f_\sigma\|_{\mathbf{L}^2(\mathcal{G})} = \|\mathcal{M}_{f_\sigma}\|_{\mathbf{L}^2(\hat{\mathcal{G}})},$$

en tenant compte de la modification apportée aux objets. En développant cette expression, nous obtenons l'égalité suivante :

$$\int_0^\infty \int_0^{2\pi} r^{2\sigma} |f(r, \theta)|^2 d\theta \frac{dr}{r} = \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_{f_\sigma}(k, v)|^2 dv. \quad (3.3)$$

Le théorème de Plancherel précise également que la TFM est bijective et bicontinue (uniformément) sur ces espaces. Si \mathcal{O} désigne un objet à niveaux de gris, $f_\sigma \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ désigne la *représentation originale* de \mathcal{O} et, de manière équivalente, $\mathcal{M}_{f_\sigma} \in \mathbf{L}^2(\hat{\mathcal{G}})$ désigne sa représentation duale au sens de la transformée de Fourier-Mellin (on parlera de *représentation de Fourier-Mellin*).

3.2.2 Action du groupe des similitudes sur la TFMA

Nous rappelons que deux objets à niveaux de gris \mathcal{O}_1 et \mathcal{O}_2 ont *même forme* si et seulement si il existe une similitude $\mathcal{S}_{\alpha, \beta}$ avec $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{S}^1$ telle que $\mathcal{O}_2 = \mathcal{S}_{\alpha, \beta} \circ \mathcal{O}_1$, soit :

$$\forall (r, \theta) \in \mathcal{G}, \quad f^2(r, \theta) = f^1(\alpha r, \theta + \beta), \quad (3.4)$$

où les fonctions modifiées $f_\sigma^1, f_\sigma^2 \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ désignent les représentations originales³ de \mathcal{O}_1 et \mathcal{O}_2 .

L'action des similitudes planes vectorielles sur l'espace de Fourier-Mellin se traduit par :

$$\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma^2}(k, v) = \alpha^{-\sigma+iv} e^{ik\beta} \mathcal{M}_{f_\sigma^1}(k, v). \quad (3.5)$$

Preuve : $\forall(k, v) \in \hat{\mathcal{G}}$,

$$\begin{aligned} \mathcal{M}_{f_\sigma^2}(k, v) &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f^2(r, \theta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r}, \\ &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f^1(\alpha r, \theta + \beta) r^{\sigma-iv} e^{-ik\theta} d\theta \frac{dr}{r}, \\ &= \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f_\sigma^1(\alpha r, \theta + \beta) r^{-iv} e^{-ik\theta} d\theta \frac{dr}{r}. \end{aligned}$$

En effectuant un changement des variables d'intégration et, puisque la mesure du groupe est invariante, nous obtenons le résultat énoncé :

$$\begin{aligned} \mathcal{M}_{f_\sigma^2}(k, v) &= \alpha^{-\sigma+iv} e^{ik\beta} \frac{1}{2\pi} \int_0^\infty \int_0^{2\pi} f_\sigma^1(r, \theta) r^{-iv} e^{-ik\theta} d\theta \frac{dr}{r}, \\ &= \alpha^{-\sigma+iv} e^{ik\beta} \mathcal{M}_{f_\sigma^1}(k, v). \end{aligned}$$

L'équation (3.5) montre que la TFMA convertit l'action interne d'une similitude sur les coordonnées des objets en une action externe sur la phase et l'amplitude du spectre. Cette relation peut être vue comme le *théorème du retard* pour le groupe des similitudes planes.

3.2.3 Caractérisation des formes et interprétation

La notion intuitive de *forme d'un objet* est très souvent employée (ne serait-ce que dans les expressions « reconnaissance de formes » et « classification de formes ») alors qu'elle est rarement définie de manière précise. Elle fait référence au fait qu'un objet ne change pas fondamentalement lorsqu'il subit une transformation géométrique. Il conserve donc sa forme.

En s'appuyant sur les propriétés précédentes, nous présentons une caractérisation des formes à niveaux de gris dans l'espace de représentation de Fourier-Mellin. La démarche s'inspire des travaux réalisés dans le cadre des objets-contours décrits par leur coefficients de Fourier [53, 56].

L'équation (3.4) caractérise l'action d'une similitude sur la représentation originale des objets. En parcourant l'espace des transformations et en regroupant l'ensemble des objets

³— Il est clair que, si un objet est représenté par la fonction f_σ de l'espace $\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$, alors il en est de même pour toutes les fonctions représentant des objets de même forme.

\mathcal{O}_i ainsi obtenus, on construit la *forme d'un objet* f :

$$\begin{aligned} \mathbb{F} &= \left\{ \mathcal{S}_{\alpha,\beta} \circ \mathcal{O}, (\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{S}^1 \right\} \\ &= \left\{ \left(f(\alpha r, \theta + \beta) \right)_{(r,\theta) \in \mathcal{G}}, (\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{S}^1 \right\}. \end{aligned} \tag{3.6}$$

Une forme regroupe donc l'ensemble des objets similaires ne se différenciant que par leur taille et leur orientation dans l'image. C'est la caractérisation d'une forme dans le domaine de représentation original des objets.

En utilisant le théorème de Plancherel (3.3) et le théorème du retard (3.5), une forme se définit de manière strictement équivalente dans le domaine de représentation de Fourier-Mellin :

$$\mathbb{F} = \left\{ \left(\alpha^{-\sigma+iv} e^{ik\beta} \mathcal{M}_{f_\sigma}(k, v) \right)_{(k,v) \in \hat{\mathcal{G}}}, (\alpha, \beta) \in \mathcal{G} \right\} \tag{3.7}$$

avec $\mathcal{M}_{f_i} \in \mathbf{L}^2(\hat{\mathcal{G}})$.

La relation « être de la même forme » définit une relation d'équivalence [74, 56] et une forme est une classe d'équivalence sur les deux espaces de représentation $\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ et $\mathbf{L}^2(\hat{\mathcal{G}})$. Une forme est donc un élément des espaces quotients

$$\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G}) / \mathcal{G} \quad \text{et} \quad \mathbf{L}^2(\hat{\mathcal{G}}) / \mathcal{G},$$

pour la représentation originale et pour la représentation de Fourier-Mellin (cf. figure 3.2).

Les espaces de représentation sont constitués de l'union disjointes de toutes les formes [75]. Les formes ne constituent pas des ensembles compacts car le groupe $\mathbb{R}_+^* \times \mathbb{S}^1$ ne l'est pas. Par contre, chaque forme peut être décomposée en sous-partitions en considérant soit le sous-groupe des rotations \mathbb{S}^1 , soit le sous-groupe des dilatations \mathbb{R}_+^* du plan⁴. Les classes d'équivalence par rotation sont alors uniquement constituées d'objets de même taille et les formes deviennent des ensembles compacts. Nous serons amenés à utiliser ce résultat dans le chapitre 4.

Tout objet f d'une forme peut être considéré comme un représentant de sa classe. Ainsi, chaque objet f^i de la classe de f s'exprime en fonction de f et d'une similitude $\mathcal{S}_{\alpha_i, \beta_i}$ par la relation $f^i(r, \theta) = f(\alpha_i r, \theta + \beta_i)$, ou bien encore par $\mathcal{M}_{f_\sigma^i}(k, v) = \alpha_i^{-\sigma+iv} e^{ik\beta_i} \mathcal{M}_{f_\sigma}(k, v)$. Dans ce contexte, le problème de *l'estimation des paramètres de mouvement* entre deux objets consiste à mesurer la position relative de ces objets à l'intérieur d'une même classe, et la *reconnaissance de formes* consiste à déterminer si ces deux objets appartiennent à la même classe.

⁴— Ce sont deux cas de groupes non transitifs de transformations par rapport à \mathbb{R}^2 (cf. les fonctions circulaires harmoniques de l'image, paragraphe 1.2.1 page 15).

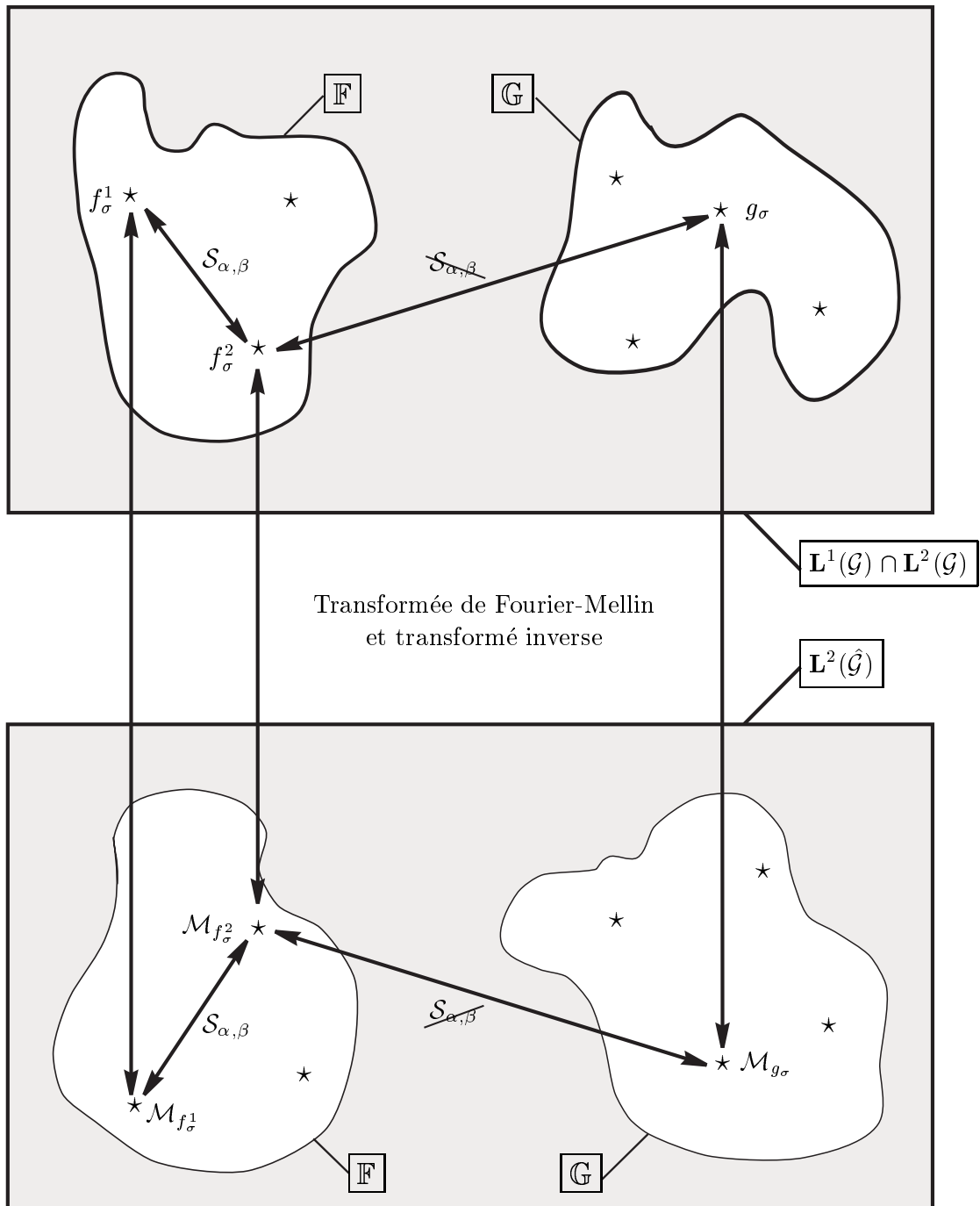


Figure 3.2 — Caractérisation des formes dans le domaine de représentation original (représentation dans $\mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ en haut) et dans le domaine transformé de Fourier-Mellin (représentation équivalente dans $\mathbf{L}^2(\hat{\mathcal{G}})$ en bas). \mathbb{F} et \mathbb{G} désignent deux formes distinctes, *i.e.* deux formes pour lesquelles il n'existe pas de similitude $\mathcal{S}_{\alpha,\beta}$ permettant de les relier entre-elles.

3.3 Extraction de nouvelles familles complètes de descripteurs invariants

Dans cette section, nous présentons un procédé original de normalisation de la transformée de Fourier-Mellin analytique qui assure simultanément l'invariance, la complétude et la sommabilité des descripteurs obtenus. Les familles ainsi extraites sont caractérisées par deux paramètres de normalisation que l'on peut choisir parmi un vaste éventail de possibilités. Dans ce contexte, une première famille est introduite en utilisant deux harmoniques de Fourier-Mellin particulières. Enfin, nous proposons une interprétation de la représentation d'une forme à l'aide d'un objet particulier que nous appelons *objet-prototype*.

3.3.1 Description du procédé de normalisation

3.3.1.1 Principe de normalisation de la TFMA

Considérons maintenant la famille de scalaires $\{\mathcal{I}_{h_\sigma}(k, v)\}_{(k,v) \in \hat{\mathcal{G}}}$ construite sur le modèle suivant :

$$\boxed{\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{h_\sigma}(k, v) = (\Gamma_{h_\sigma})^{-\sigma+iv} e^{ik(\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(k, v)} \quad (3.8)$$

où $h_\sigma \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ représente un objet. Γ_{h_σ} et Θ_{h_σ} sont deux nombres réels respectivement compatibles avec un facteur d'échelle et un angle de rotation, *i.e.* $\Gamma_{h_\sigma} \in \mathbb{R}_+^*$ et $\Theta_{h_\sigma} \in \mathbb{S}^1$. Recherchons à présent des conditions sur ces paramètres pour que la famille de scalaires soit invariante par rapport au groupe \mathcal{G} .

Supposons la famille (3.8) invariante. Par définition, pour deux objets f^1 et f^2 de même forme \mathbb{F} tels que $f^2(r, \theta) = f^1(\alpha r, \theta + \beta)$ et, pour tout couple $(k, v) \in \hat{\mathcal{G}}$, nous avons :

$$\begin{aligned} \mathcal{I}_{f_\sigma^1}(k, v) &= \mathcal{I}_{f_\sigma^2}(k, v) \\ (\Gamma_{f_\sigma^1})^{-\sigma+iv} e^{ik(\Theta_{f_\sigma^1})} \mathcal{M}_{f_\sigma^1}(k, v) &= (\Gamma_{f_\sigma^2})^{-\sigma+iv} e^{ik(\Theta_{f_\sigma^2})} \mathcal{M}_{f_\sigma^2}(k, v) \end{aligned}$$

À l'aide du théorème du retard, l'égalité précédente se simplifie :

$$\left(\frac{\Gamma_{f_\sigma^1}}{\Gamma_{f_\sigma^2}} \right)^{-\sigma+iv} e^{ik(\Theta_{f_\sigma^1} - \Theta_{f_\sigma^2})} = \alpha^{-\sigma+iv} e^{ik\beta}$$

et en identifiant les modules et arguments des deux membres de l'égalité, nous obtenons finalement les deux conditions suivantes :

$$\boxed{\begin{cases} \frac{\Gamma_{f_\sigma^1}}{\Gamma_{f_\sigma^2}} &= \alpha \\ \Theta_{f_\sigma^1} - \Theta_{f_\sigma^2} &= \beta \ (2\pi) \end{cases}} \quad (3.9)$$

La famille de descripteurs $\{\mathcal{I}_{h_\sigma}(k, v)\}_{(k,v) \in \hat{\mathcal{G}}}$ est invariante par rapport au groupe des similitudes planes vectorielles lorsque le rapport de $\Gamma_{f_\sigma^1}/\Gamma_{f_\sigma^2}$, et la différence $\Theta_{f_\sigma^1} - \Theta_{f_\sigma^2}$

des paramètres de deux objets de même forme fournissent, respectivement, le facteur d'homothétie et l'angle de rotation qui distinguent les deux objets. Réciproquement, on peut montrer que les hypothèses précédentes rendent la famille (3.8) invariante. Une telle famille est appelée *famille de descripteurs invariants* ou, plus simplement, famille d'invariants de l'objet h .

L'expression de la famille (3.9) n'est pas fortuite. Sa construction s'inspire de la structure du théorème du retard. Les paramètres Γ_{h_σ} et Θ_{h_σ} sont introduits pour compenser les actions distinctes d'une dilation et d'une rotation sur la transformée de Fourier-Mellin. C'est pourquoi nous les appellerons dorénavant *paramètres de normalisation* de la famille. Ce procédé de normalisation a récemment été utilisé pour les moments complexes (cf. paragraphe 1.2.2.1 page 16) définissant ainsi une famille complète de moments invariants par similitudes [40]. Celle-ci sera utilisée pour comparer les résultats obtenus avec les descripteurs de Fourier-Mellin que nous proposons par la suite.

Remarque : la famille précédente est invariante par rotation *et* facteur d'échelle. Comme les compensations de l'orientation et la taille sont effectuées indépendamment, il est possible de construire des familles uniquement invariantes sous l'action d'une seule transformation. Par exemple, la famille suivante est invariante par rotation :

$$\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{h_\sigma}(k, v) = e^{ik\Theta_{h_\sigma}} \mathcal{M}_{h_\sigma}(k, v), \quad (3.10)$$

avec

$$\Theta_{f_\sigma^1} - \Theta_{f_\sigma^2} = \beta(2\pi).$$

où f^1 et f^2 représentent des objets de même forme.

3.3.1.2 Unicité et sommabilité des invariants - Distance entre formes

Sans hypothèse supplémentaire, nous vérifions dans ce paragraphe que la famille (3.8) est également complète et intégrable sur l'espace $\mathbf{L}^2(\hat{\mathcal{G}})$. Ces propriétés sont à l'origine de la définition d'une vraie distance entre formes permettant de comparer des objets indépendamment des similitudes qu'ils ont subies. Auparavant, nous donnons un petit résultat sur la symétrie de la famille.

✓ Symétrie des familles proposées

Par extension de la symétrie de la TFM (2.6), on vérifie aisément que la famille possède une symétrie hermitienne :

$$\begin{aligned} \forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{h_\sigma}(-k, -v) &= (\Gamma_{h_\sigma})^{-\sigma-iv} e^{-ik(\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(-k, -v), \\ &= \overline{(\Gamma_{h_\sigma})^{-\sigma+iv} e^{ik(\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(k, v)}, \\ &= \overline{\mathcal{I}_{h_\sigma}(k, v)}. \end{aligned} \quad (3.11)$$

✓ **Unicité des descripteurs invariants**

En inversant l'expression de la famille, nous obtenons :

$$\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{h_\sigma}(k, v) = (\Gamma_{h_\sigma})^{\sigma-iv} e^{-ik(\Theta_{h_\sigma})} \mathcal{I}_{h_\sigma}(k, v). \quad (3.12)$$

Ainsi, nous pouvons retrouver la transformée de Fourier-Mellin d'un objet (et donc sa représentation originale par la TFMA inverse) à partir de ses descripteurs invariants, à une homothétie et une rotation près. Cette similitude est codée par les paramètres de normalisation Γ_{h_σ} et Θ_{h_σ} . La famille d'invariants est *complète* : elle identifie une forme de manière unique, et caractérise un objet à une similitude près représentant ses orientation et taille originales dans l'image. Toutes les familles invariantes construites sur le modèle de l'équation (3.8) sont donc complètes. Nous désignerons par $\{\mathcal{I}_{\mathbb{H}}\}_{(k,v) \in \mathbf{L}^2(\hat{\mathcal{G}})}$ les invariants calculés à partir de l'un quelconque des éléments h_σ d'une forme \mathbb{H} .

✓ **Sommabilité sur l'espace \mathbf{L}^2**

Nous rappelons que la transformée de Fourier-Mellin d'une fonction $h_\sigma \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ est intégrable sur l'espace $\mathbf{L}^2(\hat{\mathcal{G}})$, *i.e.* $\|\mathcal{M}_{h_\sigma}\|_{\mathbf{L}^2(\hat{\mathcal{G}})} < \infty$ (cf. paragraphe 3.2.1). Examinons maintenant la sommabilité de la famille d'invariants sur ce même espace :

$$\begin{aligned} \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{I}_{h_\sigma}(k, v)|^2 dv &= \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \left| (\Gamma_{h_\sigma})^{-\sigma+iv} e^{ik(\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(k, v) \right|^2 dv, \\ &= (\Gamma_{h_\sigma})^{-2\sigma} \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_{h_\sigma}(k, v)|^2 dv, \\ &= (\Gamma_{h_\sigma})^{-2\sigma} \|\mathcal{M}_{h_\sigma}\|_{\mathbf{L}^2(\hat{\mathcal{G}})}^2. \end{aligned} \quad (3.13)$$

Or, comme $(\Gamma_{h_\sigma})^{-2\sigma}$ est fini, la famille est intégrable sur l'espace $\mathbf{L}^2(\hat{\mathcal{G}})$ et sa norme est donnée par :

$$\|\mathcal{I}_{h_\sigma}\|_{\mathbf{L}^2(\hat{\mathcal{G}})} = (\Gamma_{h_\sigma})^{-\sigma} \|\mathcal{M}_{h_\sigma}\|_{\mathbf{L}^2(\hat{\mathcal{G}})} \quad (3.14)$$

✓ **Distance entre formes**

Étant donnée une famille complète et sommable de descripteurs invariants, il est possible de définir une distance euclidienne sur l'espace des formes $\mathbf{L}^2(\hat{\mathcal{G}})/\mathcal{G}$ héritée de la norme de l'équation (3.14) :

$$\begin{aligned} d_2(\mathbb{F}, \mathbb{G}) &= d_2(\mathcal{I}_{\mathbb{F}}, \mathcal{I}_{\mathbb{G}}) = d_2(\mathcal{I}_{f_\sigma}, \mathcal{I}_{g_\sigma}) \\ &= \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{I}_{f_\sigma}(k, v) - \mathcal{I}_{g_\sigma}(k, v)|^2 dv \right)^{\frac{1}{2}} \end{aligned} \quad (3.15)$$

où f_σ et g_σ sont deux représentants quelconques des formes \mathbb{F} et \mathbb{G} . Cette distance est appelée *distance invariante* [55].

Elle donne un critère quantitatif de classification et est utilisée dans la section suivante pour comparer des objets indépendamment de leur position, leur orientation et leur taille

relatives. Contrairement aux méthodes de classification basées sur des réseaux de neurone ou des méthodes statistiques, la distance invariante ne nécessite pas de phase d'apprentissage, ce qui la rend très souple d'utilisation. Du point de vue numérique, elle permet également de mesurer le degré d'isotropie des traitements et nous l'utiliserons pour tester les trois méthodes d'approximation de la TFMA sur des objets dont on sait, a priori, qu'ils ont même forme.

Finalement, l'ensemble infini de descripteurs caractérisé par les équations (3.8) et (3.9) constitue une *famille complète et intégrable de descripteurs invariants par similitude plane vectorielle*. Ces propriétés ont permis de définir une vraie distance entre formes pour la comparaison d'objets indépendamment des similitudes qu'ils ont subies. Il convient maintenant de déterminer de manière explicite des paramètres de normalisation qui remplissent les conditions énoncées en page 52.

3.3.2 Premier exemple de paramètres de normalisation

La description explicite d'une famille d'invariants consiste donc à proposer des paramètres de normalisation Γ_{h_σ} et Θ_{h_σ} vérifiant les conditions énoncées dans l'équation (3.9). Le paramètre Γ_{h_σ} est lié à la taille des objets alors que le paramètre Θ_{h_σ} est lié à leur orientation. Il n'est pas aisé, a priori, de fournir des exemples dans l'espace de représentation original. Par contre, un moyen simple et judicieux consiste à exploiter les propriétés de la représentation des objets dans l'espace de Fourier-Mellin. Nous fournissons ici un premier exemple de paramètres de normalisation, que nous appliquons ensuite au cas analytique du demi-disque introduit dans le chapitre 2.

3.3.2.1 Famille initiale

Un exemple simple consiste à utiliser les harmoniques $\mathcal{M}_{h_\sigma}(0, 0)$ et $\mathcal{M}_{h_\sigma}(1, 0)$ pour composer les paramètres de normalisation de la manière suivante :

$$\begin{cases} {}^{0,0}\Gamma_{h_\sigma} = \mathcal{M}_{h_\sigma}(0, 0)^{\frac{1}{\sigma}}, \\ {}^{1,0}\Theta_{h_\sigma} = -\arg(\mathcal{M}_{h_\sigma}(1, 0)), \end{cases} \quad (3.16)$$

en supposant $\mathcal{M}_{h_\sigma}(0, 0) \neq 0$. Les exposants placés à gauche de Γ et Θ indiquent les indices des harmoniques utilisées. On montre aisément que ces coefficients remplissent les conditions énoncées dans l'équation (3.9). Ces paramètres définissent donc une famille complète et intégrable de descripteurs invariants dont l'expression explicite est donnée par :

$$\boxed{\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{h_\sigma}(k, v) = \mathcal{M}_{h_\sigma}(0, 0)^{\frac{-\sigma+iv}{\sigma}} e^{-ik \arg(\mathcal{M}_{h_\sigma}(1, 0))} \mathcal{M}_{h_\sigma}(k, v)} \quad (3.17)$$

Cette famille est une extension de la famille introduite par F. Ghorbel [55] à toute valeur σ strictement positive. Nous l'appellerons *famille initiale* de coefficients $({}^{0,0}\Gamma_{h_\sigma}, {}^{1,0}\Theta_{h_\sigma})$ pour la distinguer de celles que nous introduirons par la suite.

Examinons maintenant le cas défavorable des fonctions f telles que $\mathcal{M}_{f_\sigma}(0, 0) = 0$. Pour de telles fonctions, la famille de descripteurs invariants n'est pas définie. Or, d'après la propriété

énoncée dans le paragraphe 2.1.2 page 27, la TFMA de f est identiquement nulle. D'après le théorème d'unicité de la TFM, la fonction f représente une image complètement noire ($f = 0$). C'est un cas dégénéré d'objet à niveaux de gris qui n'a que peu d'intérêt et ne sera plus considéré par la suite. Notons cependant que l'orbite de l'objet nul n'est constituée que d'un objet, l'objet nul lui-même.

Cet exemple constitue la première famille que nous utiliserons en reconnaissance de formes. Les paramètres de normalisation que nous avons choisis ne constituent qu'une possibilité parmi un large éventail de possibilités. D'autres coefficients, sélectionnés sur la base de critères de robustesse numérique, seront introduits dans la suite de ce chapitre. Auparavant, à titre d'exemple, nous présentons les invariants obtenus sur l'image d'un demi-disque.

3.3.2.2 Les descripteurs invariants de l'image d'un demi-disque

Dans ce paragraphe, nous complétons l'exemple de l'image d'un demi-disque f de rayon R , d'orientation δ et de niveau de gris constant L qui a été introduit dans la section 2.1 page 24 pour illustrer la transformée de Fourier-Mellin analytique.

À partir de l'expression analytique de sa TFMA (2.5), les paramètres de normalisation s'écrivent :

$$\begin{cases} {}^{0,0}\Gamma_{f_\sigma} = R \left(\frac{L}{2\sigma} \right)^{\frac{1}{\sigma}} \\ {}^{1,0}\Theta_{f_\sigma} = \delta + \frac{\pi}{2} \end{cases} \quad (3.18)$$

et, après quelques calculs, nous obtenons l'expression de la famille d'invariants de la « forme demi-disque » :

$$\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{f_\sigma}(k, v) = -L^{\frac{iv}{\sigma}} (2\sigma)^{\frac{\sigma}{\sigma-iv}} \begin{cases} \frac{1}{2} & \text{pour } k = 0 \\ \frac{1}{k\pi} i^{k-1} & \text{pour } k \text{ impair} \\ 0 & \text{ailleurs} \end{cases} \quad (3.19)$$

Le modules et phases de ces descripteurs sont illustrés à la figure 3.3. Il apparaît clairement que les invariants ne dépendent ni du rayon R ni de l'orientation δ du demi-disque. Ainsi, tous les demi-disques, indépendamment de leurs taille et orientation initiales, sont représentés par ces descripteurs de formes.

Comme le montre cette famille initiale, la transformée de Fourier-Mellin se suffit à elle-même pour extraire des descripteurs invariants puisque même les paramètres de normalisation sont directement extraits du spectre des objets. Avant de présenter des résultats numériques de reconnaissance de formes, nous proposons maintenant une interprétation du rôle des paramètres de normalisation pour la description des formes.

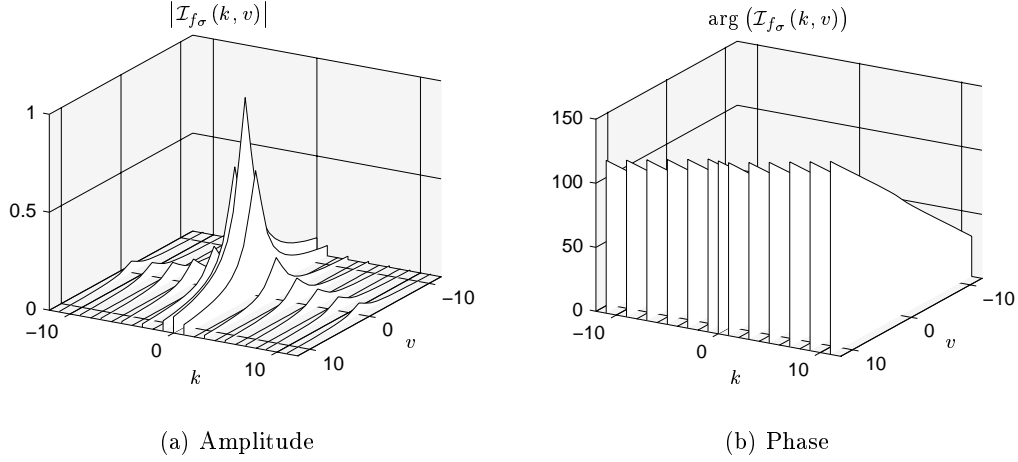


Figure 3.3 — Module et phase de quelques descripteurs invariants de faibles ordres [équation (3.19)] du demi-disque de la figure 2.1 page 26.

3.3.3 Interprétation des paramètres : objet canonique d'une forme

Il a été mentionné dans le paragraphe 3.2.3 page 49 que chaque objet peut être considéré comme un représentant de sa classe d'équivalence. Nous interprétons ici le procédé de normalisation comme le moyen d'extraire, de toute forme, un objet particulier que nous appellerons *objet canonique* de la forme. L'interprétation peut être suivie grâce au schéma 3.4 qui complète la figure 3.2 page 51.

Soit \mathbb{F} une forme contenant des objets \mathcal{O}_i représentés par les fonctions $f_\sigma^i \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$, ou, de manière équivalente, par leur transformée de Fourier-Mellin, *i.e.* $\mathcal{M}_{f_\sigma^i} \in \mathbf{L}^2(\hat{\mathcal{G}})$, avec $i \in \mathbb{N}$. Nous avons vu qu'une famille complète de descripteurs invariants caractérise une forme de manière unique et $\{\mathcal{I}_{\mathbb{F}}(k, v)\}_{(k,v) \in \hat{\mathcal{G}}}$ désigne l'ensemble des descripteurs invariants obtenus par l'un quelconque des objets f_σ^i . Comme la famille est également intégrable, l'ensemble des descripteurs invariants peut être vu comme la transformée de Fourier-Mellin d'un objet f^p défini de la manière suivante :

$$\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma^p}(k, v) = \mathcal{I}_{\mathbb{F}}(k, v). \quad (3.20)$$

Il est clair que l'objet f^p est un élément de la forme \mathbb{F} . La particularité de cet objet provient du fait que ses descripteurs invariants $\{\mathcal{I}_{f_\sigma^p}(k, v)\}$ sont identiques aux descripteurs invariants de la forme \mathbb{F} . Ses paramètres de normalisation ont donc les valeurs suivantes : $\Gamma_{f_\sigma^p} = 1$ et $\Theta_{f_\sigma^p} = 0$. Nous savons qu'il existe une similitude vectorielle permettant de relier l'objet f_σ^p à tout autre objet f_σ^i de \mathbb{F} . Précisons maintenant la position relative de cet objet à l'intérieur de sa classe.

En utilisant l'équation (3.20) et l'expression des descripteurs invariants de l'objet f^i , la TFMA de f^p s'écrit :

$$\forall (k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma^p}(k, v) = (\Gamma_{f_\sigma^i})^{-\sigma+iv} e^{ik(\Theta_{f_\sigma^i})} \mathcal{M}_{f_\sigma^i}(k, v) \quad (3.21)$$

D'après le théorème du retard (3.5), la similitude vectorielle qui distingue les deux objets est donnée par les paramètres $(\Gamma_{f_\sigma^i}, \Theta_{f_\sigma^i})$. Dans le domaine de représentation originale, cette

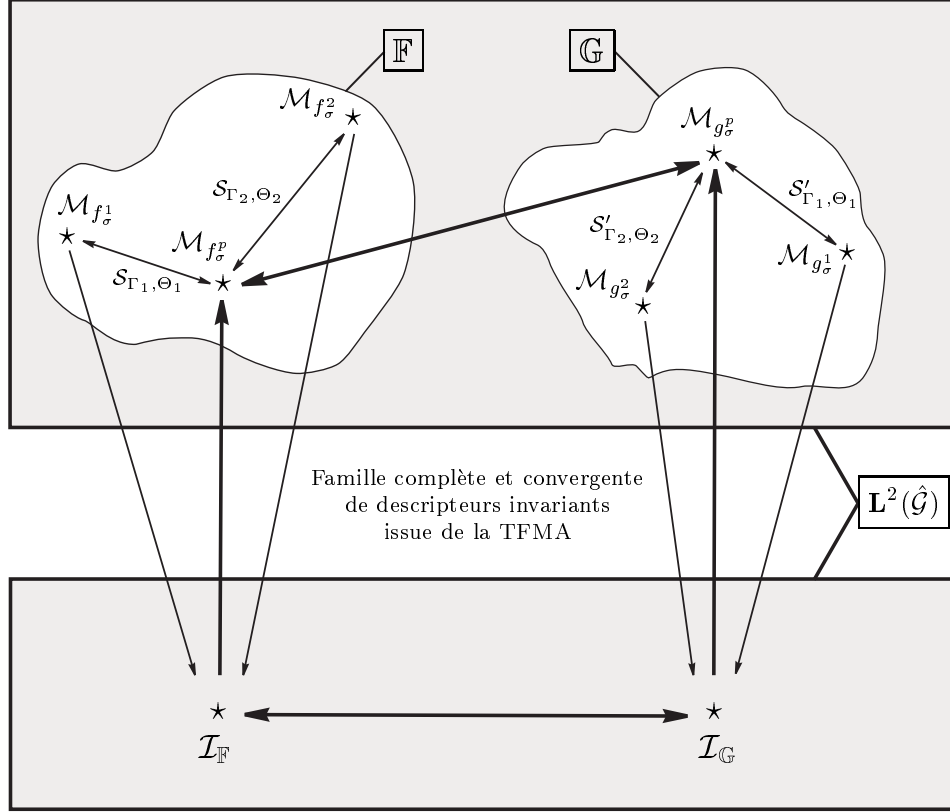


Figure 3.4 — Interprétation du rôle des paramètres de normalisation et caractérisation de l'objet prototype d'une forme. L'objet canonique f_σ^p est le représentant de la forme \mathbb{F} relativement à la famille d'invariants. Les notations $\mathcal{S}_{\Gamma_1, \Theta_1}$, $\mathcal{S}_{\Gamma_2, \Theta_2}$, $\mathcal{S}'_{\Gamma_1, \Theta_1}$, $\mathcal{S}'_{\Gamma_2, \Theta_2}$ désignent respectivement les similitudes $\mathcal{S}_{\Gamma_{f_\sigma^1}, \Theta_{f_\sigma^1}}$, $\mathcal{S}_{\Gamma_{f_\sigma^2}, \Theta_{f_\sigma^2}}$, $\mathcal{S}_{\Gamma_{g_\sigma^1}, \Theta_{g_\sigma^1}}$, $\mathcal{S}_{\Gamma_{g_\sigma^2}, \Theta_{g_\sigma^2}}$.

relation s'écrit :

$$\forall (r, \theta) \in \mathcal{G}, \quad f^p(r, \theta) = f^i(\Gamma_{f_\sigma^i} r, \Theta_{f_\sigma^i} + \theta) \quad (3.22)$$

soit,

$$\mathcal{O}_p = \mathcal{S}_{\Gamma_{f_\sigma^i}, \Theta_{f_\sigma^i}} \circ \mathcal{O}_i.$$

Cet objet particulier est le représentant de la forme \mathbb{F} relativement à la famille d'invariants \mathcal{I} . C'est l'*objet prototype ou canonique* de la forme. Les formes étant deux à deux disjointes, les objets prototypes d'objets différents sont nécessairement distincts et la distance invariante (3.15) est la distance entre les objets canoniques des formes : $d_2(\mathcal{I}_{\mathbb{F}}, \mathcal{I}_{\mathbb{G}}) = d_2(\mathcal{I}_{f_\sigma^p}, \mathcal{I}_{g_\sigma^p})$.

Ainsi, pour une famille donnée de descripteurs invariants, les paramètres de normalisation d'un objet sont les paramètres de la similitude qui distingue cet objet de l'objet prototype. De cette façon, on construit un système de coordonnées abstrait sur les formes : un élément $\mathcal{M}_{f_\sigma^1}$ a une « coordonnée de forme $\mathcal{I}_{\mathbb{F}}$ » qui nous renseigne sur sa classe, et une « coordonnée géométrique $\mathcal{S}_{\Gamma_{f_\sigma^1}, \Theta_{f_\sigma^1}}$ » qui nous renseigne sur sa position relative à l'intérieur de la classe.

Le procédé de normalisation que nous avons décrit et interprété dans cette section clarifie le rôle des paramètres de normalisation pour l'extraction des descripteurs de formes invariants. Il offre également la possibilité d'extraire de nouvelles familles choisies sur la base de critères numériques tels que la robustesse vis-à-vis des approximations numériques et des faibles variations de formes. Ces aspects sont analysés dans la suite de ce chapitre.

3.4 Extension des paramètres et représentation numérique des formes

Toutes les familles complètes et intégrables de descripteurs que l'on peut extraire grâce à l'équation (3.8) ont les mêmes propriétés théoriques et fournissent une représentation unique des formes qui interdit toute confusion entre les objets. Du point de vue numérique, les descripteurs extraits des objets ne sont jamais parfaitement invariants. Ceci est notamment dû à la représentation des images sur une grille cartésienne qui n'est pas isotrope quant aux rotations et aux dilatations, ainsi qu'à l'approximation de la TFMA. Cependant l'invariance effective d'une famille dépend également, et en grande partie, de l'estimation de ses deux facteurs de normalisation qui, pour deux objets de même forme, doivent fournir une estimation fiable des paramètres de mouvement. Dans cette section, nous étudions donc l'influence numérique du choix de ces coefficients. Nous aboutissons à la définition d'une nouvelle famille robuste de descripteurs de formes dont les paramètres de normalisation combinent plusieurs harmoniques de Fourier-Mellin. Les résultats numériques présentés ici sont issus de l'approximation rapide de la transformée de Fourier-Mellin (TFMA-F). Cette étude est complétée dans la section suivante par des expériences de reconnaissance et de discrimination de formes.

3.4.1 Robustesse des paramètres de normalisation

Pour vérifier la robustesse d'une famille, il convient de s'assurer que ses paramètres de normalisation fournissent une approximation robuste $(\hat{\alpha}, \hat{\beta})$ des paramètres de mouvement qui distinguent deux objets de même forme, selon les conditions énoncées par l'équation (3.9).

Le paramètre de normalisation du facteur d'échelle ${}^{0,0}\Gamma_{h_\sigma}$ de la famille initiale peut être très simplement étendu à toute harmonique du spectre de Fourier-Mellin puisque l'amplitude de la TFM ne dépend que de la taille des objets. Ainsi les paramètres suivants sont des alternatives au paramètre initial :

$${}^{i,j}\Gamma_{h_\sigma} = \left| \mathcal{M}_{h_\sigma}(i, j) \right|^{\frac{1}{\sigma}}, \quad (3.23)$$

avec $i, j \in \mathbb{Z}$ ⁽⁵⁾ et $\mathcal{M}_{h_\sigma}(i, j) \neq 0$ ⁽⁶⁾.

⁵— Bien que toute valeur réelle de j puisse convenir, nous nous limitons aux entiers naturels par cohérence avec la représentation numérique du spectre de Fourier-Mellin présentée dans le chapitre 2. L'harmonique utilisée pour la normalisation est directement accessible depuis le spectre, sans calcul supplémentaire.

⁶— Il faut noter que, dans le cas marginal de quelques images présentant des axes de symétrie par rotation, certaines harmoniques peuvent être nulles (cf. section 4.2). Il suffit alors de prendre la première harmonique non nulle pour normaliser la taille des objets.

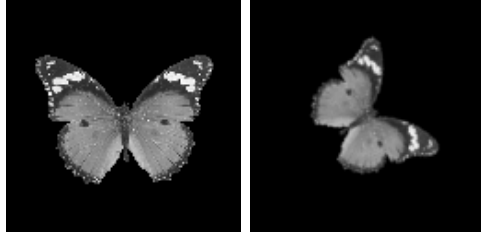


Figure 3.5 — Deux images du même papillon utilisées pour tester la fiabilité des paramètres de normalisation ($\hat{\alpha} = 0,75$ et $\hat{\beta} = -60^\circ$).

Le paramètre de normalisation de l'angle de rotation ${}^{1,0}\Theta_{h_\sigma}$ de la famille initiale peut également être étendu aux harmoniques de l'axe $v = 0$ de la manière suivante :

$${}^{n,0}\Theta_{h_\sigma} = -\frac{1}{n} \arg \left(\mathcal{M}_{h_\sigma}(n, 0) \right),$$

avec $n \in \mathbb{Z}^*$. Cette solution n'est pas numériquement satisfaisante à cause du terme $\frac{1}{n}$ qui fournit une estimation de l'angle de rotation modulo $\frac{2\pi}{n}$. On préférera utiliser la solution équivalente suivante :

$${}^{n,0}\Theta_{h_\sigma} = -\arg \left(\mathcal{M}_{h_\sigma}(n+1, 0) \cdot \mathcal{M}_{h_\sigma}(-n, 0) \right). \quad (3.24)$$

L'erreur commise sur les harmoniques de Fourier-Mellin se répercute sur le calcul des coefficients de normalisation et, pour deux objets f et g de même forme, nous n'obtenons qu'une estimation des paramètres de la similitude entre les deux objets :

$$\begin{cases} \frac{{}^{i,j}\widehat{\Gamma}_{f_\sigma}}{{}^{i,j}\widehat{\Gamma}_{g_\sigma}} & = \widehat{\alpha}_{i,j} \\ {}^{n,0}\widehat{\Theta}_{f_\sigma} - {}^{n,0}\widehat{\Theta}_{g_\sigma} & = \widehat{\beta}_{n,0} \end{cases}$$

En général, l'estimation des paramètres de mouvement se détériore au fur et à mesure que l'ordre de l'harmonique utilisée augmente, surtout pour le paramètre de normalisation de la taille des objets. Il s'avère qu'une légère erreur sur l'estimation du facteur d'échelle à un impact relativement important sur le calcul de la distance invariante comme le souligne la figure 3.6. Dans cette expérience, nous avons calculé les descripteurs invariants de la famille initiale (3.17) pour les deux papillons de la figure 3.5. Nous avons introduit une légère erreur sur les coefficients ${}^{0,0}\widehat{\Gamma}_{g_\sigma}$ et ${}^{1,0}\widehat{\Theta}_{g_\sigma}$ de telle manière que les estimations $\widehat{\alpha}_{0,0}$ et $\widehat{\beta}_{1,0}$ varient autour des valeurs réelles de la similitude qui distingue les deux images. Les deux courbes représentent la distance invariante (par approximation de l'intégrale (3.15), page 54) entre les descripteurs invariants obtenus. Celles-ci montrent que lorsqu'on s'éloigne des paramètres réels de la similitude, la distance invariante augmente sensiblement. Dans le cas du facteur d'échelle, on constate qu'une erreur de 10 % augmente la distance d'un facteur 10. Le coefficient de normalisation de l'orientation des objets influe moins sur la distance.

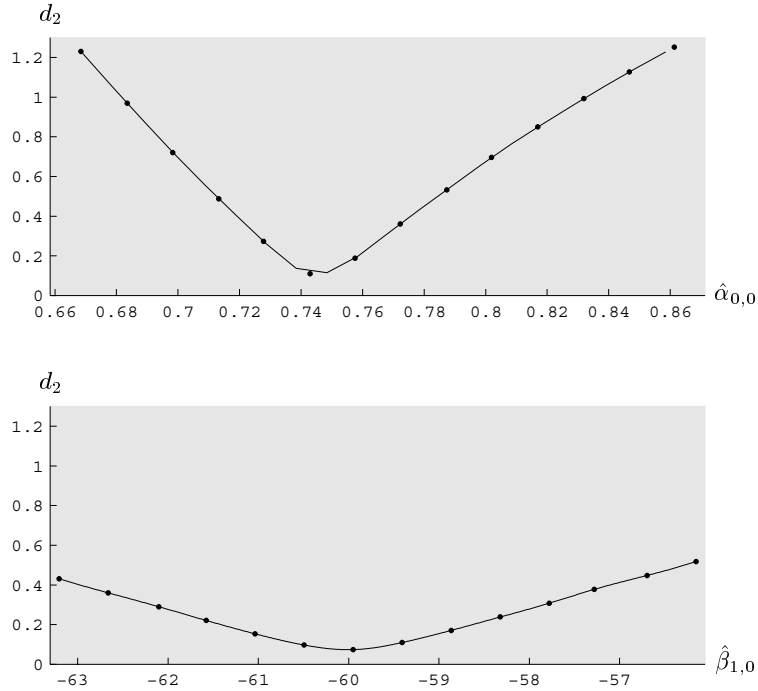


Figure 3.6 — Influence de l'erreur d'estimation des paramètres de normalisation des deux papillons de la figure 3.5 sur le calcul de la distance invariante (cf. texte). En haut : paramètre de normalisation du facteur d'échelle; en bas : paramètre de normalisation de l'angle de rotation.

3.4.2 Extension de la famille initiale

Sans connaissance a priori sur la qualité de l'estimation de la TFMA, une solution consiste à associer plusieurs harmoniques de Fourier-Mellin. Ainsi, en combinant M paramètres de normalisation de la taille pour différents couples (i, j) [équation (3.23)], et N paramètres de normalisation de l'orientation pour différents couples $(n, 0)$ [équation (3.24)] sous les formes suivantes :

$$\begin{cases} {}^M\Gamma_{h_\sigma} &= \left({}^{i_1, j_1}\Gamma_{h_\sigma} \cdot {}^{i_2, j_2}\Gamma_{h_\sigma} \cdot {}^{i_3, j_3}\Gamma_{h_\sigma} \dots \right)^{\frac{1}{M}} \\ {}^N\Theta_{h_\sigma} &= \frac{1}{N} \left({}^{n_1, 0}\Theta_{h_\sigma} + {}^{n_2, 0}\Theta_{h_\sigma} + {}^{n_3, 0}\Theta_{h_\sigma} + \dots \right), \end{cases} \quad (3.25)$$

nous obtenons une nouvelle famille complète et intégrable de descripteurs invariants.

Les paramètres de mouvement $\hat{\alpha}_M$ et $\hat{\beta}_N$ estimés à l'aide de l'équation (3.25) apparaissent comme des « moyennes » des paramètres $\hat{\alpha}_{i,j}$ et $\hat{\beta}_{n,0}$. Il est en général préférable d'utiliser un petit nombre d'harmoniques de faibles ordres car celles-ci sont moins sujettes aux erreurs d'approximation.

Pour les expériences conduites dans la section suivante, nous analyserons et comparerons

- la *famille initiale* $\{\mathcal{I}(k, v)\}$ dont l'expression est rappelée ici :

$$\boxed{\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{I}_{h_\sigma}(k, v) = ({}^{0,0}\Gamma_{h_\sigma})^{-\sigma+iv} e^{ik({}^{1,0}\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(k, v)}, \quad (3.26)$$

- et la *famille étendue* $\{\mathcal{J}(k, v)\}$ construite grâce aux paramètres de l'équation (3.25) :

$$\boxed{\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{J}_{h_\sigma}(k, v) = ({}^M\Gamma_{h_\sigma})^{-\sigma+iv} e^{ik({}^N\Theta_{h_\sigma})} \mathcal{M}_{h_\sigma}(k, v)}. \quad (3.27)$$

Examinons auparavant la représentation numériques des objets canoniques et l'influence de la troncature des deux familles d'invariants sur la description des formes.

3.4.3 Représentation numérique des formes et objets canoniques

Nous avons vu dans le paragraphe 3.3.3 page 57 qu'une famille complète et intégrable de descripteurs invariants permet de définir précisément un objet prototype représentant la forme d'une image quelconque h . L'objet prototype est caractérisé par la famille de descripteurs invariants, et est donc lié au choix des paramètres de normalisation (cf. figure 3.7). Les familles initiale et étendue fournissent donc deux objets prototypes différents que l'on peut exprimer relativement à l'objet h par [cf. équation (3.22)] :

$$\begin{cases} \mathcal{O}_p &= \mathcal{S}_{({}^{0,0}\Gamma_{h_\sigma}, {}^{1,0}\Theta_{h_\sigma})} \circ \mathcal{O}_h \\ \mathcal{O}'_p &= \mathcal{S}_{({}^M\Theta_{h_\sigma}, {}^N\Theta_{h_\sigma})} \circ \mathcal{O}_h \end{cases}$$

Les deux objets canoniques se distinguent par leur taille et leur orientation dans l'image selon la similitude $\mathcal{S}_{\tau,\omega}$ de paramètres :

$$\mathcal{O}_p = \mathcal{S}_{\tau,\omega} \circ \mathcal{O}'_p \quad \text{avec} \quad \begin{cases} \tau &= \frac{{}^{0,0}\Gamma_{h_\sigma}}{{}^M\Gamma_{h_\sigma}} \\ \omega &= {}^{1,0}\Theta_{h_\sigma} - {}^N\Theta_{h_\sigma}. \end{cases}$$

Du point de vue numérique, l'objet prototype d'une forme peut être obtenu en calculant la TFMA inverse des descripteurs invariants de l'objet h comme le montre la figure 3.7. En général, les modules des invariants sont nettement plus faible que les modules des TFMA d'objets et l'objet canonique est de très petite taille. Ainsi, les deux objets prototypes ont été agrandis d'un même facteur 100. Les objets présentent des orientations identiques (angle 0° dans le repère cartésien), donc ${}^{1,0}\Theta_{h_\sigma} = {}^N\Theta_{h_\sigma}$. Ce n'est pas toujours le cas.

Pour la reconstruction des objets prototypes de la figure 3.7, nous avons utilisé un nombre très important de descripteurs invariants ($S_{63,63} = 8065$). La propriété de complétude est « quasiment atteinte » ce qui donne une reconstruction presque parfaite des objets prototypes. Ainsi, les descripteurs invariants contiennent presque la totalité de l'information sur la forme des objets. Cependant, il n'est pas envisageable, en général, de représenter plusieurs centaines d'images par autant de descripteurs, à cause de la quantité de mémoire nécessaire

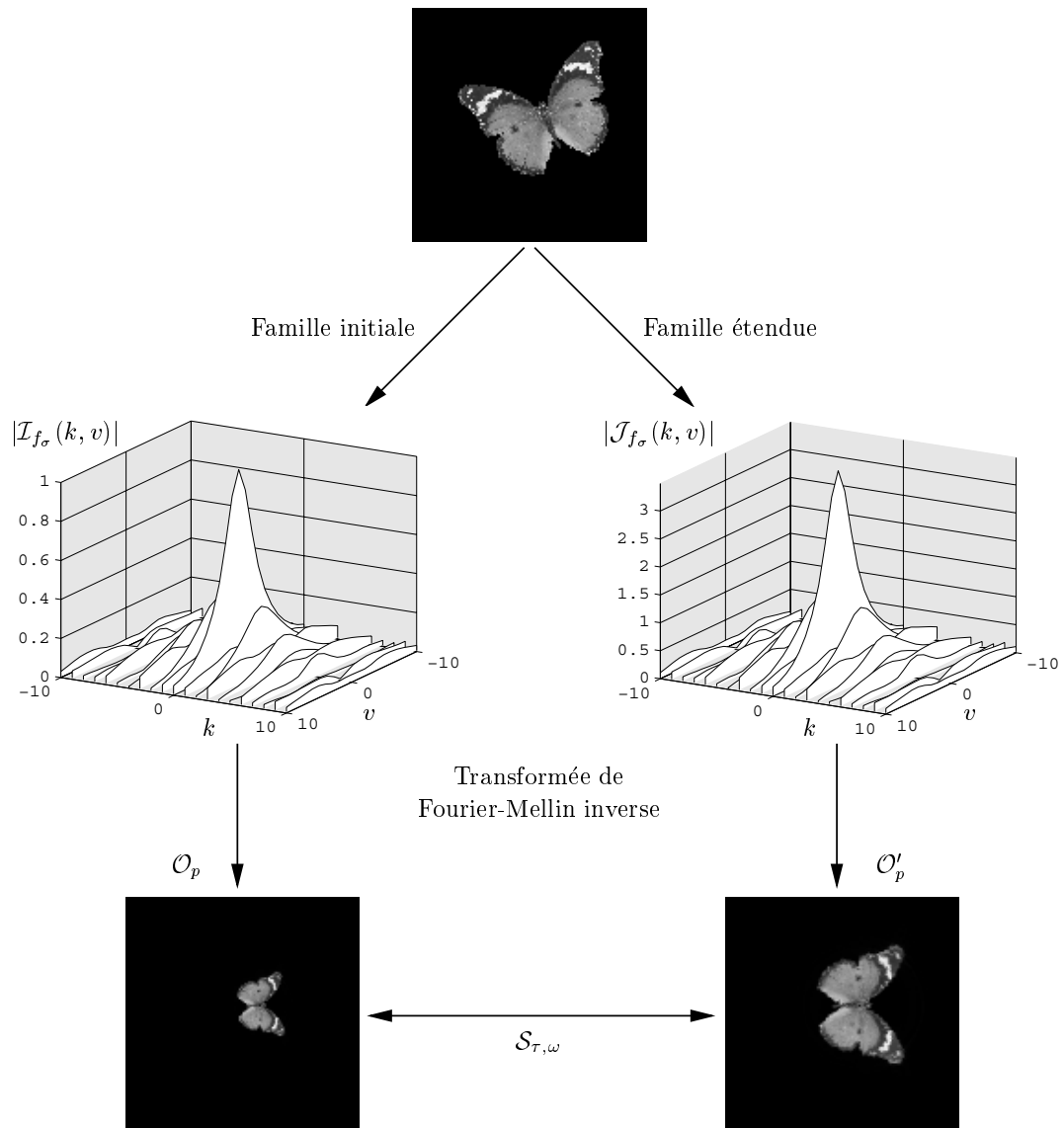


Figure 3.7 — Reconstruction numérique des objets prototypes de l'image d'un papillon pour les familles initiale et étendue de descripteurs invariants. En haut : l'image originale ; au milieu : amplitude des descripteurs invariants ; en bas : reconstruction numérique des objets canoniques (algorithme TFMA-F) à partir d'un ensemble fini de $S_{63,63} = 8065$ descripteurs invariants. Les objets prototypes ont été agrandis d'un facteur 100.

au stockage des invariants mais aussi des temps de calcul pour leur estimation. Dans les applications pratiques, il est donc nécessaire de réduire la représentation invariante des images. Il apparaît donc important d'étudier l'influence de la troncature des descripteurs invariants pour la reconnaissance et la discrimination d'objets.

3.5 Reconnaissance de formes et approximations de la TFMA

Cette section rapporte et analyse les résultats d'expériences en reconnaissance et discrimination de formes à partir des deux familles de descripteurs invariants par similitude plane que nous avons proposées dans ce chapitre. Pour chacune des trois méthodes d'approximation de la transformée de Fourier-Mellin analytique (chapitre 2), nous étudions l'influence de ne considérer qu'un ensemble fini et discret d'invariants pour la représentation numérique des formes, et en particulier les conséquences en terme de discrimination d'objets. La distance entre formes est ainsi utilisée pour :

- Mesurer l'erreur globale commise sur les invariants numériques lorsque les objets sont de même forme – Quelle méthode d'approximation de la TFM fournit la meilleur estimation de la transformée de Fourier-Mellin d'un objet ?
- Estimer le pouvoir de discrimination des deux familles lorsque les objets sont de formes différentes – Quelles sont les conséquences de la suppression des invariants d'ordre élevé sur la discrimination des objets ?
- Étudier la robustesse des invariants vis-à-vis des faibles distorsions de formes⁷.

Auparavant, nous commençons par présenter les images de test et les expériences réalisées dans cette section.

3.5.1 Images de test et présentation des résultats

Les images utilisées pour les expériences de cette section sont regroupées dans la figure 3.8. Il s'agit de quelques papillons sélectionnés dans la base présentée en annexe C.1. Les images (a), (b), (c) représentent le même papillon avec différentes tailles et orientations. L'image (d) résulte d'une transformation perspective du papillon (a) et, par conséquent, l'objet n'a pas la même forme que les trois précédents. Les images (e) et (f) représentent le même papillon à une différence d'échelle importante. Le papillon (g) a un aspect sensiblement similaire aux deux précédents et sera utilisé pour tester la robustesse des descripteurs. Enfin, le papillon de l'image (h) présente une forme et une texture nettement différentes des autres et permettra de tester la discrimination des descripteurs invariants. Il convient de noter que reconnaître de telles images est relativement délicat étant donné que leur aspect général présente des caractéristiques de formes fortement similaires, la principale différence se situant au niveau de la texture des ailes.

Pour toutes ces images, nous avons estimé le spectre de Fourier-Mellin $\widehat{\mathcal{M}}_{f_\sigma}$ à l'aide des trois méthodes d'approximation décrites dans le chapitre 2 (avec $\sigma = 0,5$). À partir des spectres numériques, nous avons calculé les descripteurs invariants pour la famille initiale $\widehat{\mathcal{I}}_{f_\sigma}$ [équation (3.26)] et la famille étendue $\widehat{\mathcal{J}}_{f_\sigma}$ (équation (3.27) avec $N=2$ et $M=3$). Enfin, la distance invariante entre les ensembles finis et discrets de descripteurs est calculée en substituant les intégrales de l'équation (3.15) page 54 par des sommes :

$$d_2^{K,V} \left(\widehat{\mathcal{I}}_{f_\sigma}, \widehat{\mathcal{I}}_{g_\sigma} \right) = \left(\sum_{v=-V}^V \sum_{k=-K}^K \left| \widehat{\mathcal{I}}_{f_\sigma}(k, v) - \widehat{\mathcal{I}}_{g_\sigma}(k, v) \right|^2 dv \right)^{\frac{1}{2}} \quad (3.28)$$

⁷ — Une *distorsion* est, pour notre modèle de transformation géométrique, soit une déformation locale de l'image (bruit, variations naturelles des formes, ...), soit une transformation globale ne se résumant pas à une similitude plane, *i.e.* un étirement ou une transformation projective, ...

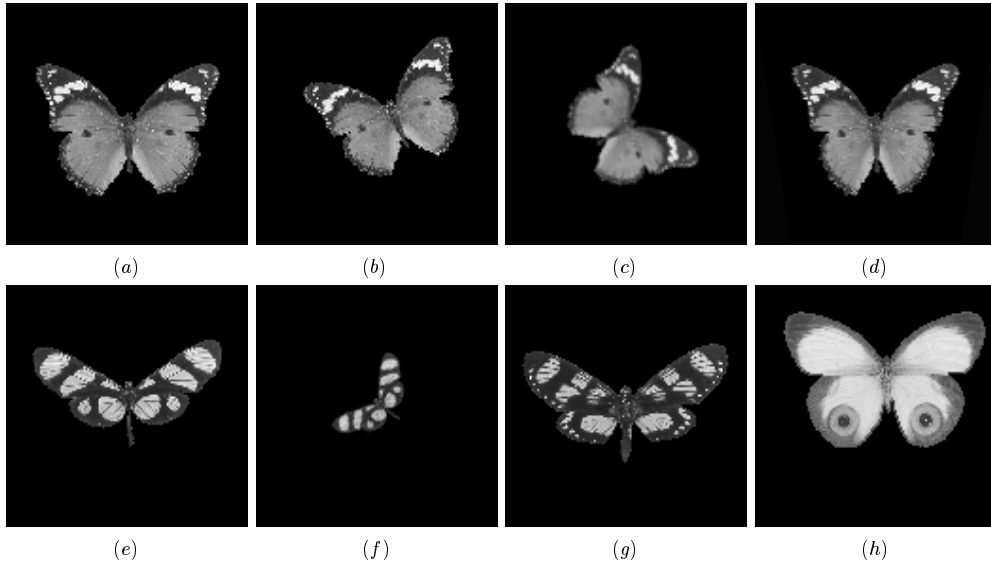


Figure 3.8 — Huit images de test utilisées pour tester la robustesse et la discrimination des familles de descripteurs invariants (128×128 pixels). (a) *Graphium Agamemnon*; (b) image (a) avec $\alpha = 0,9$ et $\beta = -20^\circ$; (c) image (a) avec $\alpha = 0,75$ et $\beta = 60^\circ$; (d) image (a) transformée par étirement; (e) *Methona Confusa*; (f) image (e) avec $\alpha = 0,5$ et $\beta = -50^\circ$; (g) *Gazera Linus*; (h) *Taenaris Macrops*.

avec K et V strictement positifs.

Les résultats de cette section se présentent sous la forme de courbes dressant l'évolution de la distance invariante et du rapport de discrimination (ratio entre les distances de paires d'objets différents et paires d'objets similaires) en fonction du nombre de descripteurs utilisés pour les calculer⁸. Les colonnes de gauche et droite présentent les résultats obtenus respectivement avec la famille initiale et avec la famille étendue. Chaque graphe reporte les résultats obtenus pour les trois méthodes d'approximation de la TFMA. Notons enfin que, pour des commodités d'écriture, la distance invariante entre les papillons (a) et (g) par exemple sera notée $d_2(a, g)$.

3.5.2 Distance entre des objets similaires

La figure 3.9 illustre l'évolution de la distance invariante entre des papillons de même forme (images (a), (b) et (c) d'une part, et images (e) et (f) d'autre part) pour les deux familles de descripteurs invariants. En haut et à gauche de chaque graphe, nous avons superposé les estimations $(\hat{\alpha}_D, \hat{\beta}_D)$, $(\hat{\alpha}_F, \hat{\beta}_F)$ et $(\hat{\alpha}_C, \hat{\beta}_C)$ des paramètres de mouvement obtenus grâce aux paramètres de normalisation des deux familles.

Comme attendu, les distances ne sont pas nulles et augmentent lorsque le nombre de descripteurs invariants s'accroît. Ainsi, les courbes traduisent l'erreur additionnelle commise en incluant de plus en plus d'invariants dans le calcul de la distance. Les différents algorithmes d'estimation de la TFMA ne fournissent pas les mêmes résultats. Sur l'ensemble des

⁸— Nous avons volontairement restreint l'étude au cas où $K = V$ pour la clarté des résultats ($S_{K,K} = (K+1)^2 + K^2$).

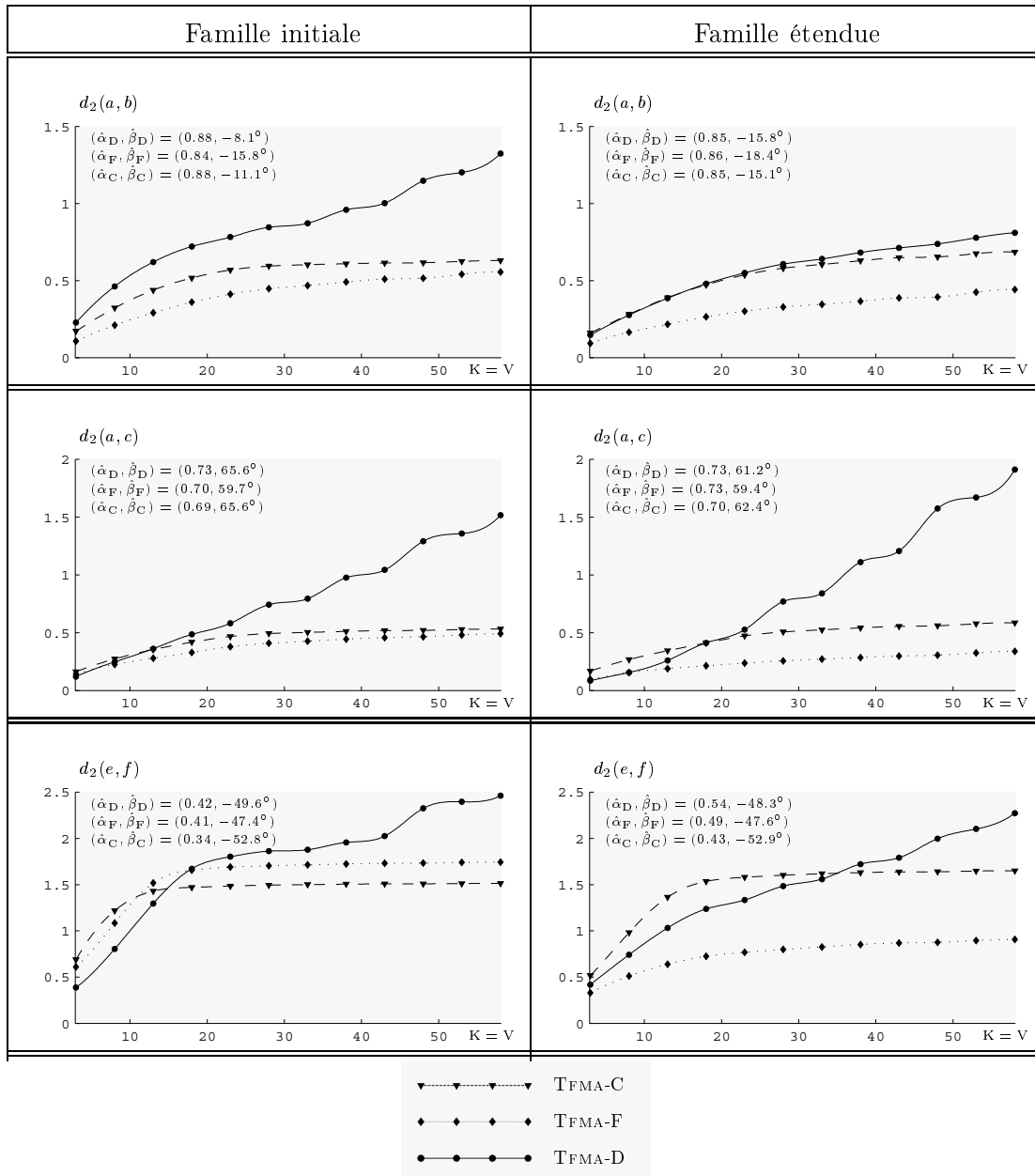


Figure 3.9 — Évolution de la distance invariante entre des objets de même forme en fonction du nombre d'invariants. En haut : papillons (a) et (b) ; au milieu : papillons (a) et (c) ; en bas : papillons (e) et (f).

graphes, on remarque que la distance calculée à partir de l'approximation directe (algorithme TFMA-D) s'accroît plus rapidement que celles obtenues à partir des deux autres méthodes d'approximation. Les invariants d'ordres élevés ont une contribution importante sur la valeur de la distance numérique. Ces quelques exemples semblent suggérer que l'algorithme TFMA-D fournit une estimation peu fiable des harmoniques hautes fréquences.

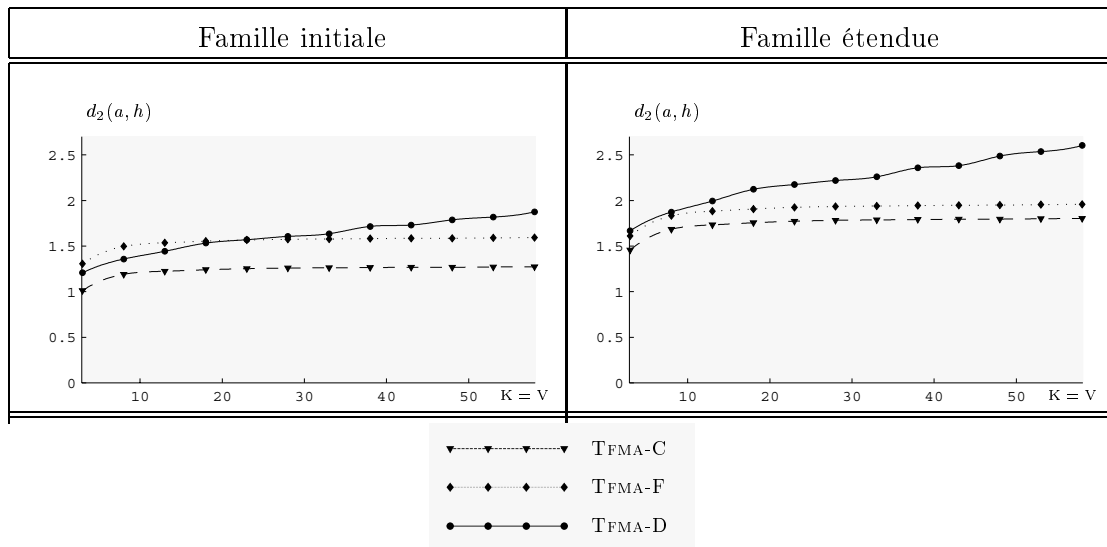


Figure 3.10 — Évolution de la distance invariante entre des objets de formes différentes [papillons (a) et (h)] en fonction du nombre d'invariants.

Famille initiale : les approximations rapide (TFMA-F) et cartésienne (TFMA-C) ont un comportement relativement similaire. Les distances ont tendance à se stabiliser au delà d'un certain rang. Pour les papillons (a), (b) et (c), la distance calculée avec l'algorithme TFMA-F est légèrement inférieure à celle obtenue par l'algorithme TFMA-C. Pour les papillons (e) et (f), nous obtenons l'inverse au-delà du rang $K = V = 12$.

Famille étendue : sur les trois exemples, il apparaît clairement que la distance calculée à partir de l'approximation TFMA-F est sensiblement plus faible que celle obtenue par l'approximation TFMA-C. Les harmoniques de faibles ordres servant à définir les paramètres de normalisation de la famille étendue sont mieux estimées par l'algorithme TFMA-F que par l'algorithme TFMA-C. En conséquence, nous obtenons une meilleure estimation des paramètres de mouvement entre les objets, et les descripteurs invariants sont plus robustes aux approximations numériques.

3.5.3 Évaluation de la discrimination des deux familles

Pour classer des objets selon leur ressemblance, il est important que les descripteurs invariants puissent discriminer des objets de même forme et des objets de formes différentes. Il est donc primordial que la distance entre des objets de formes différentes soit sensiblement plus élevée que la distance entre des objets de même forme. La figure 3.10 montre l'évolution de la distance invariante entre des papillons d'aspects différents —images (a) et (h)— en fonction du nombre de descripteurs invariants. Il est intéressant de noter que la distance obtenue avec la famille étendue est légèrement plus importante que celle obtenue avec la famille initiale. L'approximation TFMA-C fournit les distances les plus faibles.

Pour illustrer la discrimination des deux familles, les deux graphes de la figure 3.11 montrent le ratio entre les distances invariantes $d_2(a, h)$ et $d_2(a, b)$ d'une part, et entre les distances invariantes $d_2(a, h)$ et $d_2(a, c)$ d'autre part. Notons que ce rapport n'a de sens que d'un point de vue numérique. La discrimination entre des objets est importante lorsque le

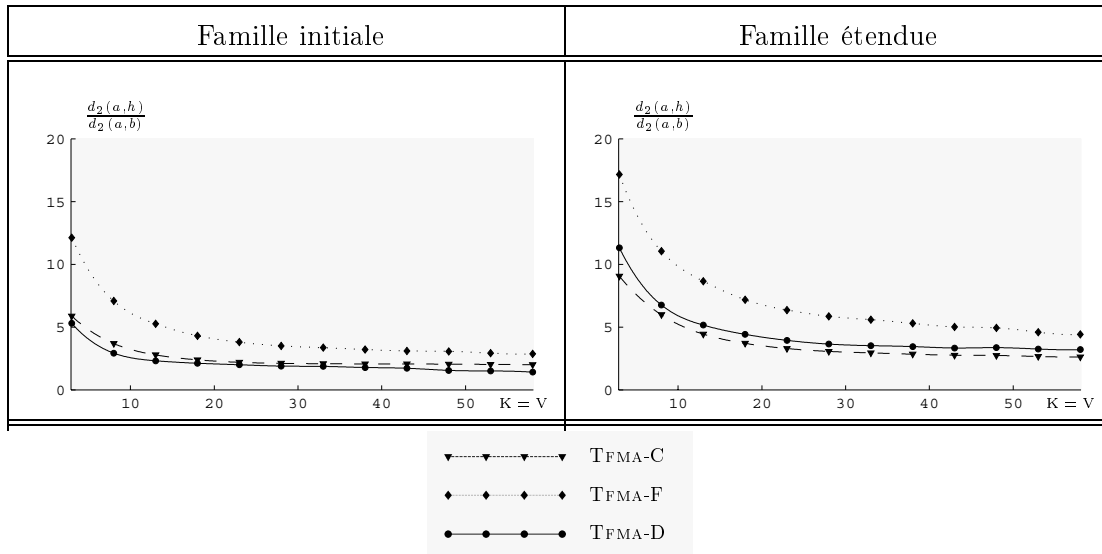


Figure 3.11 — Évaluation de la discrimination numérique des familles initiale et étendue de descripteurs invariants. Évolution du ratio de discrimination entre les papillons de formes différentes (a,h) et les papillons de même forme (a,b).

ratio est élevé. Par contre, si celui-ci est proche de 1, alors les distances invariantes sont presque identiques et il devient impossible de discerner entre des objets similaires et des objets de formes différentes. Le ratio procure ainsi une mesure du potentiel de discrimination des deux familles et de la robustesse des différentes méthodes d'approximation de la TFMA.

Les deux courbes reportées dans la figure 3.11 montrent que le ratio de discrimination diminue au fur et à mesure que le nombre de descripteurs invariants augmente. Il faut cependant noter que, pour les trois méthodes d'approximation de la TFMA, ce ratio est plus élevé dans le cas de la famille étendue que dans le cas de la famille initiale et reste discriminant même pour des ensembles de descripteurs invariants importants. Il apparaît clairement sur les graphes que l'approximation rapide de la TFMA fournit des résultats de discrimination nettement supérieurs à ceux des deux autres méthodes.

La figure 3.12 à la page ci-contre illustre la représentation numérique de la forme des papillons (a) et (g) par leur objet prototype (cf. paragraphe 3.4.3), lorsque celui-ci est reconstruit à partir d'un ensemble fini et discret d'invariants (famille étendue, Algorithme TFMA-F). Pour des ensembles de descripteurs invariants relativement larges, on distingue très nettement les formes des objets. Par contre, pour des ensembles de dimensions $S_{15,15}$ et $S_{7,7}$, le prototype ne représente visuellement qu'une signature partielle de la forme de l'objet. Il nous est plus difficile de distinguer les deux représentations. Or c'est pour ces dimensions que les familles invariantes sont les plus discriminantes. Nous pouvons interpréter ce résultat en soulignant que l'erreur commise sur les descripteurs invariants d'ordres élevés prend le pas sur l'information discriminante supplémentaire qu'ils contiennent.

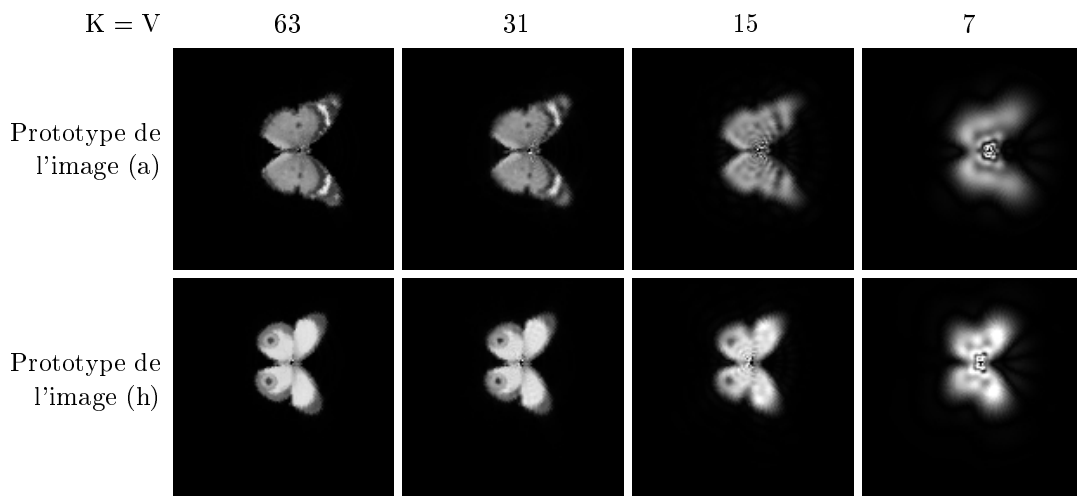


Figure 3.12 — Reconstruction des objets prototypes (algorithme TFMA-F) de deux papillons de formes très différentes (a) et (h) pour des ensembles discrets et finis $(S_{K,V})$ de descripteurs invariants issus de la famille étendue.

Les résultats des expériences montrent que pour discriminer des objets d'aspects fortement différents, il est préférable de n'utiliser qu'un faible nombre de descripteurs invariants. Qu'en est-il des objets dont les formes sont proches sans être tout à fait similaires ?

3.5.4 Robustesse des descripteurs pour des formes presque similaires

Dans ce paragraphe, nous étudions le comportement des descripteurs invariants lorsque les objets à comparer sont différents mais de formes très proches. Pour cela, nous analysons les résultats de discrimination entre les papillons (a) et (d) d'une part, et les papillons (e) et (g) d'autre part.

Résultats concernant les images (a) et (d) :

L'image (d) a été obtenue en appliquant un étirement horizontal à l'image (a). Les objets n'ont donc pas la même forme au sens du groupe des similitudes planes et il est intéressant d'analyser le comportement des deux familles de descripteurs invariants pour ces images.

Les deux graphes situés en haut de la figure 3.13 présentent l'évolution de la distance invariante pour les trois méthodes d'approximation de la TFMA. Elle est sensiblement inférieure à celle obtenue pour des papillons de formes franchement différentes (à comparer avec la figure 3.10 page 67). Par contre, la distance invariante entre les images (a) et (d) n'est que très légèrement supérieure à celle obtenue entre les papillons (a) et (b), comme le montre le rapport entre les distances $d_2(a, d)$ et $d_2(a, b)$.

Famille initiale : le ratio de discrimination tend vers 1, quelque soit la méthode d'approximation de la TFMA utilisée. Ainsi, pour un nombre important de descripteurs invariants, la discrimination entre les images (a), (b) et (d) devient critique. L'approximation directe de la TFMA montre même un ratio inférieur à 1, ce qui signifie que, au sens de l'algorithme TFMA-D, les images (a) et (h) ont une forme plus semblable que les images (a) et (b) !

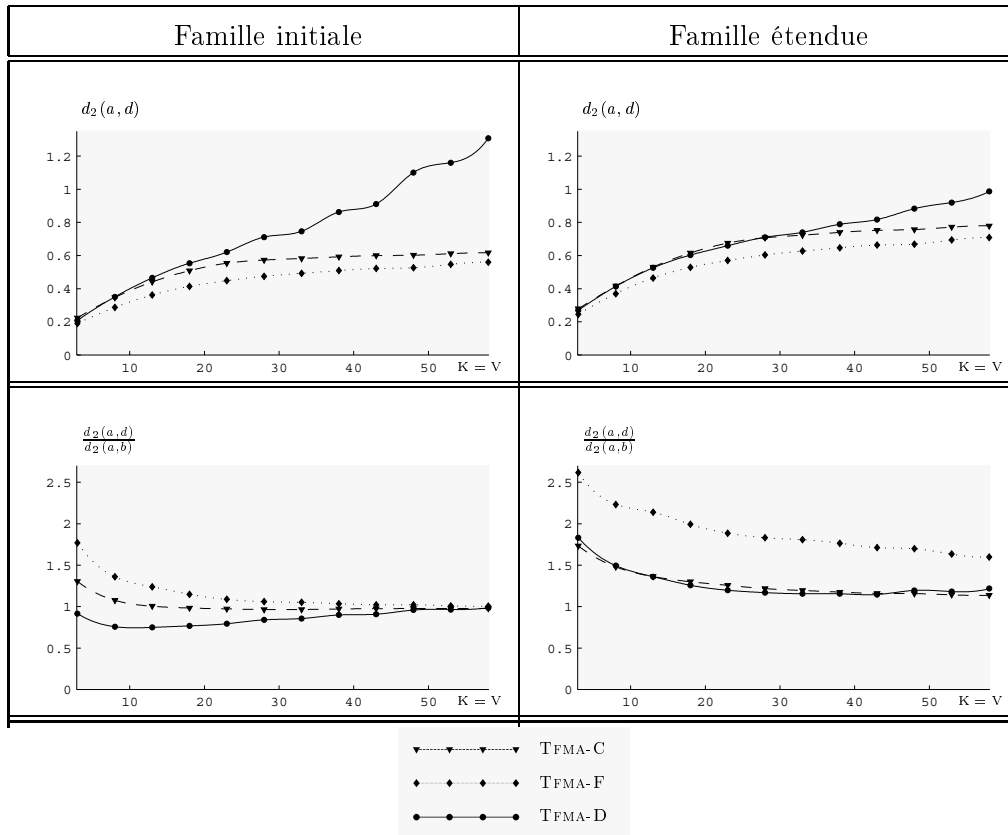


Figure 3.13 — Évaluation de la discrimination numérique des familles initiales et étendues entre deux papillons de formes proches mais différentes (étirement horizontal). En haut : distance invariante entre les papillons (a) et (d) ; en bas : rapport de discrimination entre les papillons (a,d) et les papillons (a,b).

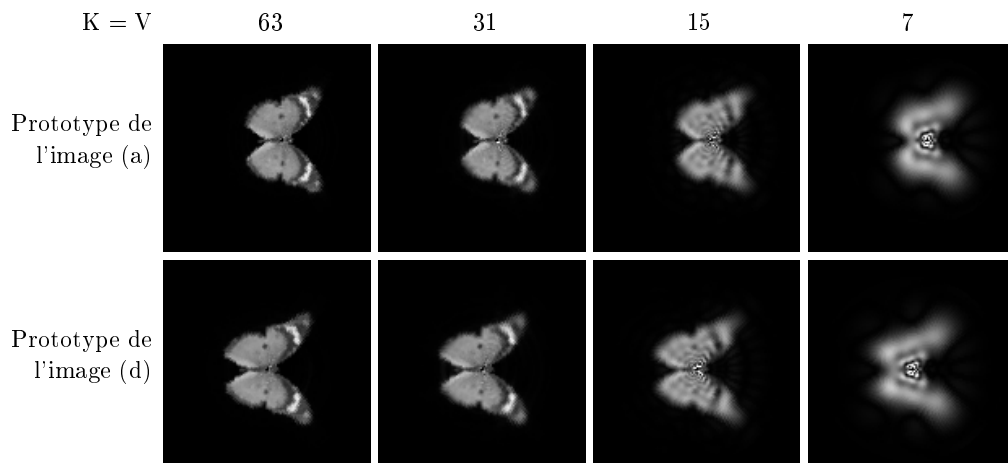


Figure 3.14 — Reconstruction des objets prototypes (algorithme TFMA-F) de deux papillons de formes proches mais différentes (a) et (d) (étirement horizontal) pour des ensembles discrets et finis ($S_{K,V}$) de descripteurs invariants issus de la famille étendue.

Famille étendue : on constate que le ratio de discrimination, même s'il reste faible, ne tend pas vers 1. Il est donc toujours possible de dissocier les images avec un nombre important de descripteurs invariants. L'algorithme TFMA-F se distingue des deux autres algorithmes par un ratio sensiblement plus élevé. La famille étendue de descripteurs invariants reflète mieux l'ordre de proximité visuelle entre les trois images.

Résultats concernant les images (e) et (g) :

Dans ce second exemple, nous utilisons les images (e) et (g) de papillons présentant de faibles différences à la surface de leurs ailes. Ces différences proviennent de variations naturelles et il est intéressant d'analyser le comportement des deux familles de descripteurs invariants en fonction des déformations locales des images (cf. figure 3.15).

Les distances invariantes entre les papillons (e) et (g) présentent des profils analogues pour les trois méthodes d'approximation de la TFMA et restent relativement faibles. Le ratio de discrimination en bas de la figure 3.15 permet de comparer les résultats avec ceux obtenus entre les papillons de même forme (e) et (f) (cf. figure 3.9 page 66).

Famille initiale : le ratio de discrimination tend rapidement vers 1, quelque soit la méthode d'approximation de la TFMA utilisée. Ainsi, pour un nombre important de descripteurs invariants, la discrimination entre les images (e), (f) et (g) devient critique. La distance invariante obtenue par l'approximation rapide de la TFMA montre même un résultat de l'ordre de 1, quelle que soit le nombre de descripteurs invariants utilisés. Ainsi, il est impossible de distinguer les papillons.

Famille étendue : le ratio est nettement plus important que dans le cas précédent et n'atteint jamais la valeur critique de 1. Bien que les papillons (e) et (f) présentent de fortes différences de tailles et d'orientation, il est toujours possible de discriminer les trois images. Il est également intéressant de constater que le ratio de discrimination présente un maximum, non plus pour les ensembles de descripteurs les plus faibles, mais pour des ensembles de taille $S_{10,10} = 221$. Contrairement au cas d'objets fortement différents, les descripteurs jusqu'à cet ordre apportent une information pertinente supplémentaire pour la discrimination des objets présentant de faibles déformations locales.

Les résultats présentés dans cette section nous permettent de tirer quelques conclusions tant sur les familles de descripteurs invariants que sur les méthodes d'approximation du prolongement analytique de la transformée de Fourier-Mellin.

À partir des différents tests effectués, il apparaît clairement que la famille étendue est plus robuste que la famille initiale aux approximations et aux problèmes d'isotropie de la grille cartésienne vis-à-vis des similitudes. Elle s'avère également plus discriminante et permet de discerner entre des objets de même forme et des objets de formes proches. Le gain peut être directement attribué aux paramètres de normalisation de la famille étendue qui, d'un point de vue numérique, réalisent mieux les conditions d'invariance (meilleure estimation des paramètres de mouvement lorsque les objets ont même forme).

Ces expériences nous ont également permis de tester et de comparer les trois méthodes

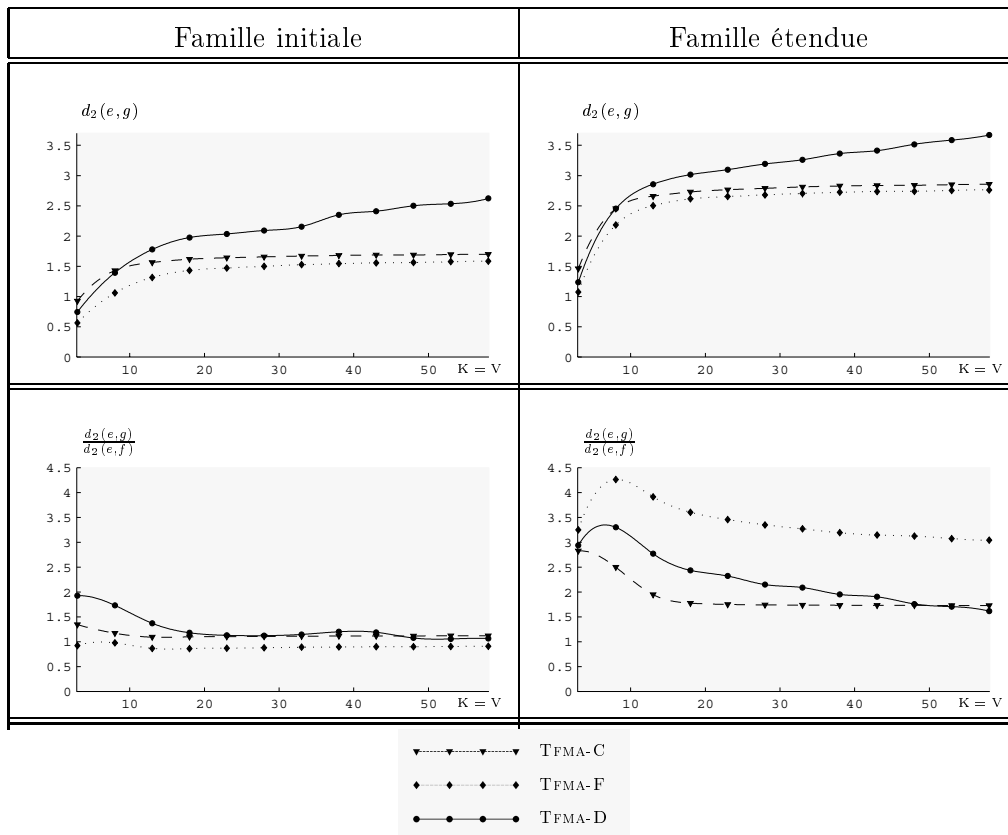


Figure 3.15 — Évaluation de la discrimination numérique des familles initiales et étendues entre deux papillons de formes proches mais différentes (faibles variations locales de texture). En haut : distance invariante entre les papillons (e) et (g) ; en bas : rapport de discrimination entre les papillons (e,g) et les papillons (e,f).

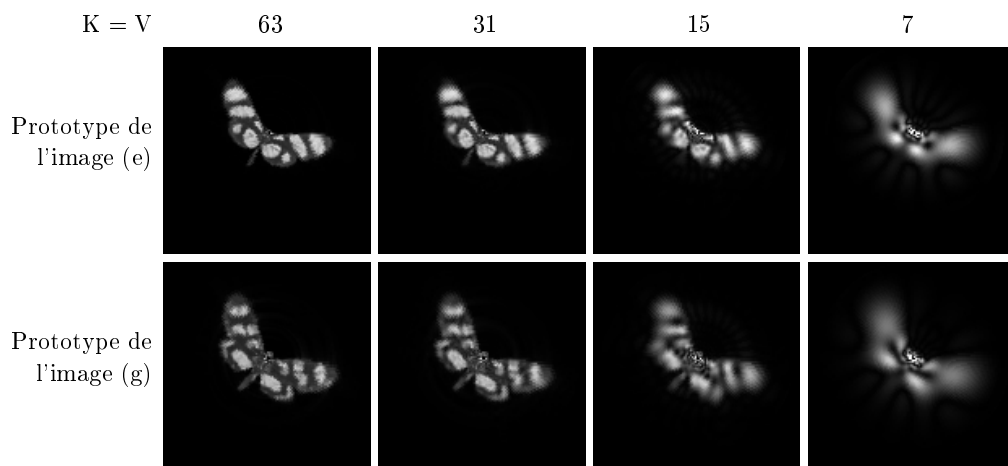


Figure 3.16 — Reconstruction des objets prototypes (algorithme TFMA-F) de deux papillons de formes proches mais différentes (e) et (f) (variations naturelles de la texture des ailes) pour des ensembles discrets et finis ($S_{K,V}$) de descripteurs invariants issus de la famille étendue.

d'estimation de la TFMA. De manière générale, il faut noter le bon comportement des trois approximations vis-à-vis de l'invariance par similitudes, confirmant ainsi que les algorithmes proposés dans le chapitre 2 fournissent une approximation fiable de la transformée de Fourier-Mellin analytique. Néanmoins, les trois algorithmes présentent certaines différences. Sur les exemples, l'approximation directe (TFMA-D) fournit une estimation peu robuste des harmoniques d'ordres élevés. Cette remarque confirme les résultats obtenus sur la reconstruction des images. Les approximations cartésienne (TFMA-C) et rapide (TFMA-F) offrent un comportement similaire. Néanmoins, la méthode d'intégration numérique de l'algorithme TFMA-C n'est pas adaptée aux similitudes, ce qui se traduit par des résultats de discrimination relativement médiocres, mais toujours discriminants, comparables à ceux obtenus par l'approximation directe. Il apparaît clairement que l'approximation TFMA-F fournit le meilleur compromis entre invariance et discrimination, même pour des ensembles de descripteurs importants.

3.6 Conclusion du chapitre

Dans ce chapitre, nous avons proposé et analysé de nouvelles familles de descripteurs de formes invariants par similitude, extraits depuis le prolongement analytique de la transformée de Fourier-Mellin. Comme ces familles sont également complètes et intégrales, une distance sur l'espace des formes a été utilisée pour mesurer le degré de similarité des objets indépendamment de leur pose, de leur orientation et de leur taille.

Le procédé original de normalisation de la TFMA qui a permis d'extraire ces familles met en évidence et clarifie le rôle de deux paramètres de normalisation que l'on peut choisir parmi un large éventail de possibilités. Bien que toutes les familles construites sur ce modèle aient les mêmes propriétés théoriques, l'invariance numérique effective d'une famille est inhérente à la qualité de l'estimation de ses paramètres de normalisation. Nous avons montré et vérifié de manière expérimentale que ceux-ci sont d'autant plus robustes qu'ils fournissent une estimation fiable des paramètres de mouvement pour des objets de même forme. Ainsi, sur la base de critères numériques, de nouveaux paramètres qui améliorent sensiblement la robustesse des descripteurs ont été introduits (famille étendue).

Les expériences ont également permis de constater que l'approximation rapide de la TFMA fournit des résultats d'invariance et de discrimination légèrement supérieurs à ceux obtenus avec les approximations directe et cartésienne (algorithmes TFMA-D et TFMA-C). Compte tenu de ses bons résultats de reconstruction, l'algorithme TFMA-F apparaît comme l'estimation la plus fiable des trois.

Dans le chapitre 5, nous utiliserons les familles complètes de descripteurs invariants pour la recherche d'images en vue d'indexer des bases d'objets par le contenu. Auparavant, nous proposons un algorithme original d'appariement d'objets à niveaux de gris à partir de la transformée de Fourier-Mellin analytique des images.

4 Estimation du Mouvement Global entre Objets

Le chapitre précédent a mis en évidence l'importance de l'estimation des paramètres de mouvement pour la construction des familles complètes de descripteurs invariants par similitudes. Dans ce chapitre, nous proposons une méthode d'appariement des images à niveaux de gris utilisant la TFMA, indépendamment des translations, rotations et facteurs d'échelle. L'algorithme n'a pas recours à l'extraction préalable de primitives dans les images. Toute l'information disponible sur la forme des objets est prise en compte, ce qui rend la méthode particulièrement robuste.

L'algorithme de correspondance que nous proposons se base sur la distance euclidienne et consiste à minimiser une fonction de deux variables de mouvement (ρ, ψ) . Le minimum de cette fonction fournit, d'une part, une estimation de la similitude entre deux objets quelconques, et d'autre part, un écart entre formes. Cette mesure de similarité est importante puisqu'elle permet de déterminer si les paramètres de mouvement estimés sont significatifs (cf. section 4.1).

Dans la section 4.2, nous abordons le cas particulier du groupe des rotations du plan. Dans ce cas, nous montrons que le minimum de la restriction de la fonction d'appariement est très précisément la distance de Hausdorff exprimée sur l'espace des formes à niveaux de gris. À titre d'application, nous proposons un algorithme pour détecter et estimer l'ensemble des axes de symétrie par rotation d'une image.

Enfin, la section 4.3 résume les résultats d'expériences en estimation de mouvement et en discrimination de formes. Ces expériences permettent également de tester, une fois de plus, la fiabilité des trois méthodes d'approximation de la TFMA quant aux problèmes d'isotropie de la grille cartésienne par rapport aux similitudes.

4.1 Appariement d'objets et estimation des paramètres de mouvement

Les techniques d'appariement de primitives se basent généralement sur l'extraction et la localisation de certaines zones caractéristiques de l'image comme des points particuliers, des segments, des coins... que l'on extrait depuis les contours de l'image (images structurées comme des vues aériennes d'une zone urbaine). Ensuite, une étape de mise en correspondance, que l'on peut assimiler à une phase de reconnaissance de formes, consiste à appairer les primitives des images de façon combinatoire par des techniques d'alignement [67], de corrélation, ou bien encore en utilisant la distance de hausdorff [65]. Chaque appariement fournit une estimation des paramètres de mouvement entre les images. Beaucoup des difficultés rencontrées par ces méthodes sont dues à leur donnée d'entrée : l'extraction des primitives et leur mise en correspondance sont cause de nombreuses erreurs (faux appariements). Ainsi, les algorithmes ont généralement recours à un système de vote qui détermine les correspondances auxquelles on apporte le plus de crédibilité. Celui-ci est basé sur la présence d'une zone d'accumulation dans l'espace des paramètres qui détermine le mouvement le plus fréquemment détecté.

Dans le cas d'images texturées, et non structurées comme précédemment, C. SCHMID [110, 111] propose d'utiliser des points d'intérêts extraits directement depuis l'information à niveaux de gris de l'image. Le système global de mise en correspondance reste essentiellement identique et permet d'appairer des images indépendamment des rotations. La méthode ne prend pas directement en compte le facteur d'échelle.

Dans le cas des objets à niveaux de gris, les axes principaux permettent d'estimer les paramètres de translation, de rotation et de facteur d'échelle entre deux images. Une solution similaire consiste à utiliser la relation entre certains moments géométriques des images [125]. Ces méthodes sont comparables à l'estimation fournie par les paramètres de normalisation (cf. chapitre précédent). Elles prennent en compte l'ensemble de l'information disponible sur les objets, ce qui permet de s'affranchir de l'étape d'extraction de primitives. Néanmoins, ces méthodes ne permettent pas de tester si les objets ont même forme. Or, estimer des paramètres de mouvement n'a de sens qu'entre des objets de même forme, et ces algorithmes supposent qu'une phase de reconnaissance des objets a été préalablement réalisée.

La méthode que nous présentons dans la suite de cette section réalise une correspondance dense des objets. Toute l'information contenue dans la forme et la texture des objets participe à l'estimation des paramètres de mouvement, ce qui rend la méthode particulièrement robuste. L'algorithme consiste à minimiser une fonction d'appariement qui s'inspire de la distance euclidienne entre les objets. Le minimum de cette fonction et ses coordonnées fournissent respectivement une mesure de ressemblance entre les objets et les paramètres de mouvement estimés (paragraphe 4.1.1). Ainsi, nous pourrions mesurer la pertinence des paramètres de mouvement estimés.

Une des principales difficultés de l'algorithme est l'aspect combinatoire de la recherche dans l'espace des paramètres. Pour cela, nous utilisons un algorithme d'optimisation déterministe que nous initialisons par une estimation préalable des paramètres de mouvement qui assure une convergence rapide pour des objets de même forme (paragraphe 4.1.2).

4.1.1 Principe de la méthode d'appariement

Un algorithme élémentaire pour estimer les paramètres de mouvement distinguant deux objets \mathcal{O}_1 et \mathcal{O}_2 de même forme \mathbb{F} consiste à appliquer une similitude de paramètres (ρ, ψ) à \mathcal{O}_2 et à confronter l'image obtenue avec \mathcal{O}_1 . Si les deux images se superposent, alors le couple (ρ, ψ) fournit les paramètres de la transformation entre les objets, sinon on réitère l'opération pour toutes les similitudes possibles. L'espace de recherche est déterminé par le groupe \mathcal{G} . Outre le côté systématique et laborieux de la méthode, une des principales difficultés réside donc dans le choix d'une mesure de similarité entre objets qui permet d'une part, de déterminer si les objets ont même forme, c'est-à-dire si le calcul des paramètres de mouvement a un sens, et d'autre part, de donner une estimation robuste de la pose relative des images.

Pour cela, nous proposons une méthode qui s'inspire de la distance euclidienne entre objets et consiste à minimiser une fonction bi-dimensionnelle dépendant des paramètres de mouvement. À titre d'exemple, nous présentons également la surface d'appariement obtenue dans le cas de demi-disques de taille et d'orientation différentes.

4.1.1.1 Fonction d'appariement

Nous rappelons que la distance euclidienne quadratique entre deux fonctions f_σ et $g_\sigma \in \mathbf{L}^1(\mathcal{G}) \cap \mathbf{L}^2(\mathcal{G})$ est définie par l'équation suivante :

$$d_2(f_\sigma, g_\sigma) = \|f_\sigma - g_\sigma\|_{\mathbf{L}^2(\mathcal{G})} = \left(\int_0^\infty \int_0^{2\pi} |f_\sigma(r, \theta) - g_\sigma(r, \theta)|^2 d\theta \frac{dr}{r} \right)^{\frac{1}{2}}. \quad (4.1)$$

Cette distance est nulle pour des objets strictement identiques. Dans le cas d'objets de même forme, celle-ci n'est pas nulle et ne permet pas de comparer directement des objets de taille et d'orientation différentes. Pour tenir compte des transformations géométriques que peuvent subir les objets, nous proposons de modifier l'équation (4.1) afin d'obtenir une fonction dépendant des paramètres $(\rho, \psi) \in \mathbb{R}_+^* \times \mathbb{S}^1$:

$$\begin{aligned} \mathbf{E}_{f_\sigma, g_\sigma}(\rho, \psi) &= \|f_\sigma(r, \theta) - g_\sigma(\rho r, \theta + \psi)\|_{\mathbf{L}^2(\mathcal{G})}, \\ &= \left(\int_0^\infty \int_0^{2\pi} r^{2\sigma} |f(r, \theta) - g(\rho r, \theta + \psi)|^2 d\theta \frac{dr}{r} \right)^{\frac{1}{2}}. \end{aligned} \quad (4.2)$$

Lorsque les paramètres (ρ, ψ) parcourent l'ensemble des transformations possibles, nous dressons une « surface d'appariement » entre les deux objets. Cette fonction mesure, en tout point de l'espace des paramètres, la similarité entre les deux objets et nous l'appellerons *fonction d'appariement*.

En utilisant l'identité de Parseval pour le groupe des similitudes planes (cf. paragraphe 3.2.1 page 48), la fonction de l'équation (4.2) peut s'écrire de manière strictement

équivalente sur l'espace de représentation de Fourier-Mellin $\mathbf{L}^2(\hat{\mathcal{G}})$:

$$\begin{aligned} \mathbf{E}_{f_\sigma, g_\sigma}(\rho, \psi) &= \left\| \mathcal{M}_{f_\sigma(r, \theta)} - \mathcal{M}_{g_\sigma(\rho r, \theta + \psi)} \right\|_{\mathbf{L}^2(\hat{\mathcal{G}})}, \\ &= \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \left| \mathcal{M}_{f_\sigma(r, \theta)}(k, v) - \mathcal{M}_{g_\sigma(\rho r, \theta + \psi)}(k, v) \right|^2 dv \right)^{\frac{1}{2}}, \end{aligned}$$

soit, en utilisant le théorème du retard pour la fonction g_σ (3.5), nous obtenons :

$$\boxed{\begin{aligned} \forall (\rho, \psi) \in \mathbb{R}_+^* \times \mathbb{S}^1, \\ \mathbf{E}_{f_\sigma, g_\sigma}(\rho, \psi) = \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \left| \mathcal{M}_{f_\sigma}(k, v) - \rho^{-\sigma + iv} e^{ik\psi} \mathcal{M}_{g_\sigma}(k, v) \right|^2 dv \right)^{\frac{1}{2}} \end{aligned}} \quad (4.3)$$

D'un point de vue numérique, le calcul de la fonction d'appariement sur l'espace de représentation de Fourier-Mellin est plus rapide et plus robuste que le calcul direct sur l'espace original. En effet, pour estimer la fonction sous la forme de l'équation (4.2), il est nécessaire d'interpoler les niveaux de gris de l'image g pour tous les points de coordonnées polaires $(\rho r, \theta + \psi)$, ce qui engendrerait une source d'erreur importante. Par contre, dans le cas de l'équation (4.3), il s'agit d'une simple multiplication du spectre de g par le nombre complexe $\rho^{-\sigma + iv} e^{ik\psi}$, sans erreur numérique supplémentaire. De plus, cette formulation nous permettra par la suite d'estimer les paramètres de mouvement à partir de spectres discrets et d'étendue finie calculés à l'aide des différentes approximations numériques de la TFMA.

4.1.1.2 Mesure de correspondance et estimation des paramètres de mouvement

La fonction d'appariement ne s'annule jamais exactement pour des objets de même forme et, estimer les paramètres de mouvement revient donc à rechercher un couple $(\hat{\alpha}_{\mathbf{E}}, \hat{\beta}_{\mathbf{E}})$ qui minimise l'équation (4.3) :

$$\boxed{(\hat{\alpha}_{\mathbf{E}}, \hat{\beta}_{\mathbf{E}}) = \arg \left(\underbrace{\min_{\psi \in \mathbb{S}^1, \rho \in \mathbb{R}_+^*} \hat{\mathbf{E}}_{f_\sigma, g_\sigma}(\rho, \psi)}_{\hat{\mathbf{E}}_{f_\sigma, g_\sigma}^{\min}} \right)} \quad (4.4)$$

$\mathbf{E}_{f_\sigma, g_\sigma}^{\min}$ n'est pas une distance entre formes car le groupe \mathcal{G} n'est pas compact¹. En effet si le minimum vaut 0, alors il n'est pas possible de conclure que les objets ont même forme. Par contre, c'est un *écart entre formes*. L'ordre dans lequel les fonctions f_σ et g_σ sont appariées n'a pas d'importance : $\hat{\mathbf{E}}_{f_\sigma, g_\sigma}^{\min} = \hat{\mathbf{E}}_{g_\sigma, f_\sigma}^{\min}$, et les paramètres de mouvement estimés sont conjugués. L'écart donne une mesure de correspondance entre les objets : plus celui-ci est faible, plus les objets sont similaires au sens de cette mesure. La valeur \mathbf{E}^{\min} peut donc être utilisée pour

¹— Par contre, en restreignant les transformations admissibles au groupe compact des rotations du plan, $\mathbf{E}_{f_\sigma, g_\sigma}^{\min}$ est très précisément la distance de Hausdorff sur l'espace des formes (cf. section 4.2 page 84).

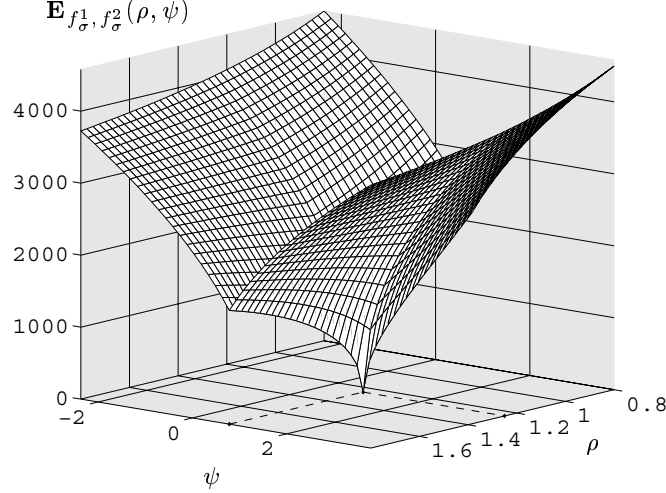


Figure 4.1 — Fonction d'appariement entre deux images de demi-disques définis respectivement par $R_1 = 40$, $\delta_1 = 0$ et $R_2 = 50$, $\delta_2 = \frac{\pi}{4}$. La fonction s'annule au point de coordonnées $(\alpha, \beta) = (1, 25; \frac{\pi}{4})$ qui correspond au paramètre de la similitude qui distingue les deux demi-disques.

classer les objets selon leur ordre de similarité, indépendamment de l'orientation et de la taille des objets.

À titre d'exemple, nous présentons maintenant le résultat du calcul de la fonction $\mathbf{E}_{f_{\sigma}, g_{\sigma}}(\rho, \psi)$ avec des images de demi-disques de tailles et d'orientations différentes.

4.1.1.3 Exemple de deux demi-disques

Considérons deux demi-disques f^1 et f^2 de niveaux de gris constant L , de tailles respectives R_1 et R_2 et d'orientations respectives δ_1 et δ_2 . À partir de l'expression de leur TFMA [cf. équation (2.5)], la fonction d'appariement entre les deux demi-disques se calcule de manière analytique, et nous obtenons : $\forall(\rho, \psi) \in \mathbb{R}_+^* \times \mathbb{S}^1$,

$$\mathbf{E}_{f^1, f^2}(\rho, \psi) = \frac{L^2 \pi}{2\sigma} * \begin{cases} R_1^{2\sigma} + \left(\frac{R_2}{\rho}\right)^{2\sigma} \left(\frac{2}{\pi} |\psi + \delta_1 - \delta_2| - 1\right) & \text{pour } \rho \geq \frac{R_2}{R_1} \\ \left(\frac{R_2}{\rho}\right)^{2\sigma} + R_1^{2\sigma} \left(\frac{2}{\pi} |\psi + \delta_1 - \delta_2| - 1\right) & \text{pour } \rho < \frac{R_2}{R_1} \end{cases} \quad (4.5)$$

La figure 4.1 représente la fonction (4.5) pour deux demi-disques définis par les paramètres $(R_1 = 40, \delta_1 = 0)$ et $(R_2 = 50, \delta_2 = \frac{\pi}{4})$. La fonction s'annule au point de coordonnées $\left(\frac{R_2}{R_1} = 1, 25; \delta_2 - \delta_1 = \frac{\pi}{4}\right)$ qui définissent les paramètres de la similitude qu'il faut appliquer à f^1 pour obtenir f^2 .

Examinons maintenant le cas de surfaces d'appariement numériques obtenues à partir de spectres discrets.

4.1.2 Optimisation de la fonction $\mathbf{E}(\rho, \psi)$

Dans ce paragraphe, nous nous intéressons aux aspects numériques du calcul de la fonction $\mathbf{E}(\rho, \psi)$ et nous montrons quelques résultats préliminaires d'expériences d'appariement et d'estimation des paramètres de mouvement. Notamment, nous illustrons les surfaces des fonctions d'appariement entre des objets de même forme et entre des objets de formes différentes. Ensuite, pour s'affranchir de la recherche systématique du minimum global de la fonction sur l'espace des paramètres, nous utilisons un algorithme d'optimisation permettant d'accélérer la recherche de l'écart et d'améliorer l'estimation des paramètres de mouvement. Pour toutes les expériences reportées ici, les spectres sont estimés grâce à l'algorithme TFMA-F.

4.1.2.1 Appariement d'objets similaires et d'objets différents

La fonction $\mathbf{E}(\rho, \psi)$ entre des images représentées par leur spectre numérique est estimée en substituant les intégrales de l'équation (4.3) page 78 par des sommes :

$$\widehat{\mathbf{E}}_{f_\sigma, g_\sigma}^{K, V}(\rho, \psi) = \left(\sum_{v=-V}^V \sum_{k=-K}^K \left| \widehat{\mathcal{M}}_{f_\sigma}(k, v) - \rho^{-\sigma+iv} e^{ik\psi} \widehat{\mathcal{M}}_{g_\sigma}(k, v) \right|^2 dv \right)^{\frac{1}{2}}, \quad (4.6)$$

avec K et V non nuls.

La figure 4.2 illustre la fonction d'appariement obtenue entre les images (a) et (b) d'une part [figure (d)], et les images (a) et (c), d'autre part [figure (e)]. Les surfaces sont représentées par échantillonnage des variables ρ et ψ sur une grille régulière.

Cas d'objets de même forme : la figure (d) présente un puits très marqué dont la valeur minimum $\widehat{\mathbf{E}}_{a,b}^{\min} = 144$ est atteinte au point de coordonnées $\hat{\alpha}_{\mathbf{E}} = 0,75$ et $\hat{\beta}_{\mathbf{E}} = 1,0$ rad. Comme attendu, le minimum de la fonction n'est pas exactement nul. Les coordonnées et la valeur du minimum dépendent de la grille d'échantillonnage de la surface et ne représentent qu'une valeur approximative du résultat.

Cas d'objets de formes différentes : la figure (e) ne présente pas de minimum remarquable dans le domaine de recherche que nous avons limité à $\rho \in [0, 7; 3, 0]$ pour le facteur d'échelle. Le minimum de la fonction d'appariement est atteint au point de coordonnées $\hat{\alpha}_{\mathbf{E}} = 3,00$ et $\hat{\beta}_{\mathbf{E}} = 2,80$ rad et vaut $\widehat{\mathbf{E}}_{a,c}^{\min} = 1215$. L'écart est très nettement supérieur à celui obtenu dans le cas d'objets de même forme.

4.1.2.2 Estimation de l'écart entre formes

Pour s'affranchir de l'étape de recherche systématique dans l'espace des paramètres, nous avons mis en œuvre un algorithme d'optimisation fondé sur la méthode de Powell [100, chap. 10]. Cette méthode ne fait pas appel au calcul de dérivées. À partir d'un point de la surface

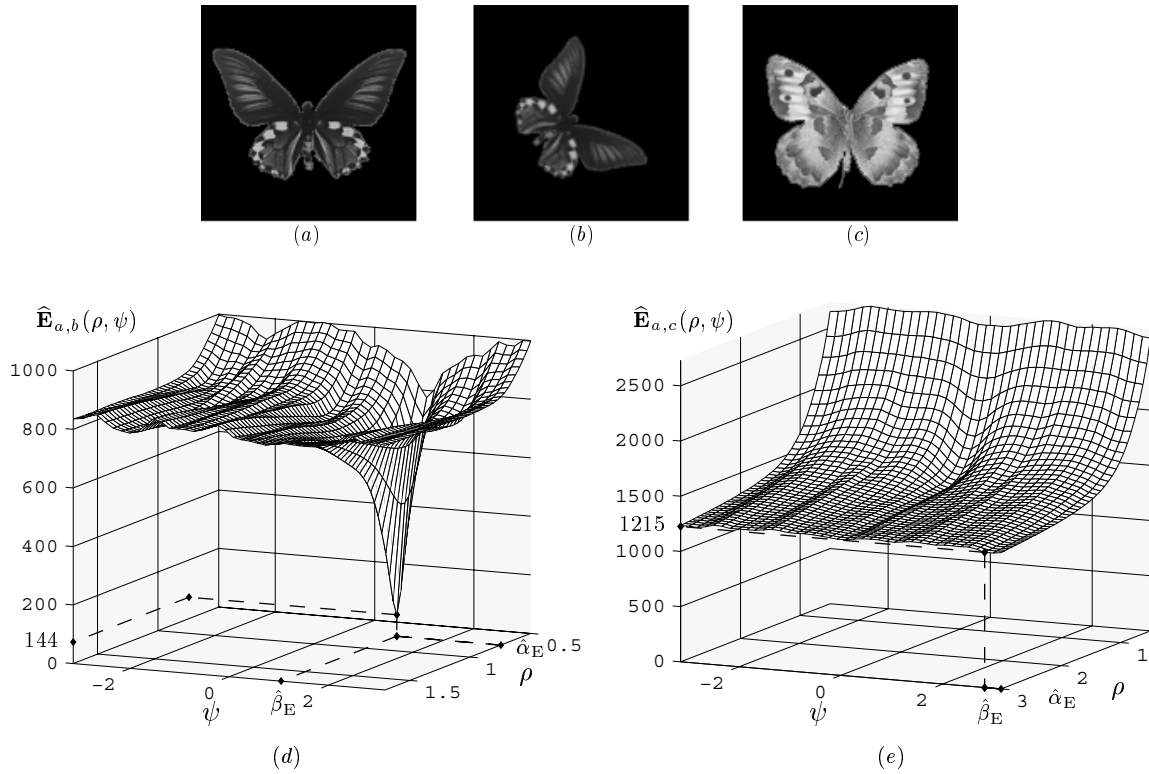


Figure 4.2 — Illustration de la fonction d'appariement entre deux papillons similaires (a et b) et deux papillons de formes différentes (a et c). (a) image du papillon de référence; (b) image (a) avec $\alpha = 0,75$ et $\beta = 60^\circ$ soit $1,047$ rad; (c) image d'un papillon de forme différente; (d) $\hat{E}_{a,b}^{\min} = 144$, $\hat{\alpha}_E = 0,75$ et $\hat{\beta}_E = 1,0$ rad; (e) $\hat{E}_{a,c}^{\min} = 1215$, $\hat{\alpha}_E = 3,00$ et $\hat{\beta}_E = 2,8$ rad.

(appelé point de gradient), l'algorithme consiste à déterminer la « meilleure direction » permettant de descendre rapidement le long de la pente en évaluant la fonction d'appariement en différents endroits (algorithme de Brent [21]). Cette direction définit un nouveau point de gradient et l'algorithme est réitéré jusqu'à la convergence vers un minimum (fixée par un certain seuil de tolérance).

Si cet algorithme simple converge, en général, rapidement vers un minimum, il est par contre sensible à la présence de minima locaux. Pour éviter une estimation erronée des paramètres de mouvement, il est donc important d'accorder un soin particulier au choix du point qui initialise l'algorithme d'optimisation. Celui-ci conditionne, d'une part, la convergence vers un minimum local ou global et, d'autre part, la rapidité avec laquelle ce minimum est atteint. Pour initialiser l'algorithme d'optimisation nous utilisons l'estimation de mouvement fournit par les paramètres de normalisation introduits dans le chapitre précédent. Nous avons en effet pu juger que ceux-ci offrent une estimation robuste des paramètres de la similitude qui distinguent deux objets de même forme.

Le tableau 4.1 présente les itérations de l'algorithme d'optimisation pour les papillons (a)

et (b) de la figure 4.2. Le seuil de tolérance est fixé à 10^{-4} . Pour chaque point de gradient, nous avons reporté les estimations du facteur d'échelle $\hat{\rho}$ et de l'angle de rotation $\hat{\psi}$ ainsi que la valeur $\widehat{\mathbf{E}}(\hat{\rho}, \hat{\psi})$.

Exemple 1 : l'initialisation de l'algorithme provient des paramètres de normalisation de la famille initiale (3.16) :

$$\left. \begin{aligned} \hat{\rho} &= \frac{{}^{0,0}\Gamma_b}{{}^{0,0}\Gamma_a} = 0,684 \\ \hat{\psi} &= {}^{1,0}\Theta_b - {}^{1,0}\Theta_a = 1,03 \text{ rad} \end{aligned} \right\} \Rightarrow \widehat{\mathbf{E}}(\hat{\rho}, \hat{\psi}) = 260,23.$$

Exemple 2 : l'initialisation de l'algorithme provient des paramètres de normalisation de la famille étendue (3.25) :

$$\left. \begin{aligned} \hat{\rho} &= \frac{{}^M\Gamma_b}{{}^M\Gamma_a} = 0,716 \\ \hat{\psi} &= {}^N\Theta_b - {}^N\Theta_a = 1,04 \text{ rad} \end{aligned} \right\} \Rightarrow \widehat{\mathbf{E}}(\hat{\rho}, \hat{\psi}) = 144,87.$$

Itérations		1	2	3	4	5
Exemple 1	N	1	18	30	45	67
	$\hat{\rho}$	0,684	0,761	0,49	0,749	0,749 = $\hat{\alpha}_{\mathbf{E}}$
	$\hat{\psi}$	1,0364	1,0452	1,0455	1,0460	1,0460 = $\hat{\beta}_{\mathbf{E}}$
	$\widehat{\mathbf{E}}(\hat{\rho}, \hat{\psi})$	260,229	85,444	72,743	72,726	72,724 = $\widehat{\mathbf{E}}_{a,b}^{\min}$
Exemple 2	N	1	13	24	38	58
	$\hat{\rho}$	0,716	0,753	0,748	0,749	0,749 = $\hat{\alpha}_{\mathbf{E}}$
	$\hat{\psi}$	1,0447	1,0463	1,0461	1,0460	1,0460 = $\hat{\beta}_{\mathbf{E}}$
	$\widehat{\mathbf{E}}(\hat{\rho}, \hat{\psi})$	144,871	74,892	72,725	72,724	72,724 = $\widehat{\mathbf{E}}_{a,b}^{\min}$

Tableau 4.1 — Itérations de l'algorithme d'optimisation pour l'estimation des paramètres de mouvement entre des objets de même forme. L'exemple 1 correspond à une initialisation par les paramètres de normalisation de la famille initiale alors que l'exemple 2 correspond à une initialisation par les paramètres de la familles étendue. N représente le nombre cumulé de fois où la fonction $\mathbf{E}_{f_{\sigma}, g_{\sigma}}(\rho, \psi)$ a été estimée.

Pour ces deux exemples, l'algorithme converge en 5 itérations vers des solutions identiques : $\hat{\alpha}_{\mathbf{E}} = 0,749$, $\hat{\beta}_{\mathbf{E}} = 1,046$ et $\widehat{\mathbf{E}}_{a,b}^{\min} = 72,7$. Il est intéressant de noter que le nombre d'évaluations de la fonction d'appariement (qui prend une place prépondérante dans le coût global de l'optimisation) est légèrement plus faible dans le cas de l'exemple 2. L'algorithme d'optimisation converge plus rapidement lorsqu'il est initialisé à l'aide des paramètres de normalisation de la famille étendue. La figure 4.3 illustre le chemin parcouru par l'algorithme d'optimisation à la surface de la fonction d'appariement. La localisation du minimum, donc l'estimation des paramètres de mouvement, gagne légèrement en précision. Ce qui est plus intéressant encore, c'est la diminution de la valeur de l'écart (d'un facteur 3,5 pour l'exemple 1, et d'un facteur 2 pour l'exemple 2).

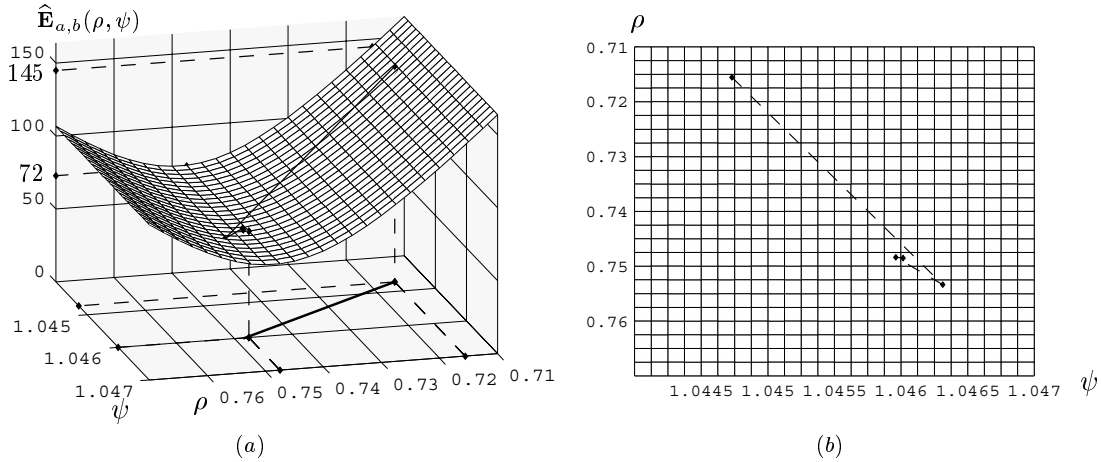


Figure 4.3 — Chemin parcouru par l’algorithme d’optimisation à la surface de la fonction d’appariement calculée pour deux objets de même forme. (a) Zoom sur le puits de la figure 4.2 et parcours de l’algorithme d’optimisation ; (b) Chemin parcouru par l’algorithme d’optimisation dans l’espace des paramètres depuis l’estimation initiale (fournie par les paramètres de normalisation de la famille étendue) jusqu’au minimum estimé.

Le tableau 4.2 rend compte de quelques itérations de l’algorithme d’optimisation dans le cas d’objets de formes différentes (images (a) et (c) de la figure 4.2). L’algorithme converge, pour un seuil de tolérance de 10^{-4} , au bout de 51 itérations (soit 660 calculs de l’équation (4.6)) pour finalement aboutir à la solution suivante : $\hat{\alpha}_{\mathbf{E}} = 311, 1$, $\hat{\beta}_{\mathbf{E}} = -1, 86$ rad et $\hat{\mathbf{E}}_{a,b}^{\min} = 822$.

Itérations		1	11	21	31	41	51
Exemple 2	N	1	128	271	402	535	660
	$\hat{\rho}$	7, 7	83, 0	126, 2	194, 4	231, 6	311, 1 = $\hat{\alpha}_{\mathbf{E}}$
	$\hat{\psi}$	0, 13	3, 06	-1, 75	-1, 92	-1, 74	-1, 86 = $\hat{\beta}_{\mathbf{E}}$
	$\hat{\mathbf{E}}(\hat{\rho}, \hat{\psi})$	1073	852	835	826	824	822 = $\hat{\mathbf{E}}_{a,b}^{\min}$

Tableau 4.2 — Quelques itérations de l’algorithme d’optimisation entre des objets de formes différentes. L’algorithme est initialisé à l’aide des paramètres de normalisation de la famille étendue.

L’écart entre formes est bien plus important dans le cas d’objets de formes différentes que dans le cas d’objets similaires. Le facteur d’échelle estimé entre les deux images dépasse largement les valeurs acceptables. Il faut également noter que l’algorithme met sensiblement plus de temps pour extraire un minimum entre des objets de formes différentes puisqu’il faut, dans ce cas précis, 660 évaluations de la fonction d’appariement (contre 60 dans le cas d’objets de même forme).

Avant de poursuivre les expériences, examinons le cas particulier de la fonction d’appariement lorsque le mouvement est restreint au groupe des rotations du plan.

4.2 Sous-groupe des rotations vectorielles et distance de Hausdorff

Dans cette section, nous nous intéressons au cas particulier où le mouvement apparent entre les objets se résume aux seules rotations du plan : $\mathcal{G} = \mathbb{S}^1$. Dans ce cas, nous montrons que le minimum de la restriction de la fonction $\mathbf{E}(\rho, \psi)$ à \mathbb{S}^1 :

$$\mathbf{E}_{f_\sigma, g_\sigma}(\psi) = \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_{f_\sigma}(k, v) - e^{ik\psi} \mathcal{M}_{g_\sigma}(k, v)|^2 dv \right)^{\frac{1}{2}}, \quad (4.7)$$

est très précisément la distance de Hausdorff sur l'espace de représentation de Fourier-Mellin. Nous commençons par un bref rappel sur la distance de Hausdorff et son utilisation pour l'appariement d'images. Le résultat est ensuite appliqué à la détection des axes de symétrie par rotation des images.

4.2.1 Rappels sur la distance de Hausdorff

Soient \mathcal{A} et \mathcal{B} deux sous-ensembles fermés et bornés d'un espace métrique (\mathbf{X}, d) . La distance de Hausdorff est donnée par [64] :

$$\Delta(\mathcal{A}, \mathcal{B}) = \max(\delta(\mathcal{A}, \mathcal{B}), \delta(\mathcal{B}, \mathcal{A})) \quad (4.8)$$

où $\delta(\mathcal{A}, \mathcal{B})$ désigne la « distance de Hausdorff dirigée » de \mathcal{A} vers \mathcal{B} ⁽²⁾ :

$$\delta(\mathcal{A}, \mathcal{B}) = \max_{a \in \mathcal{A}} \min_{b \in \mathcal{B}} d(a, b)$$

La distance de Hausdorff dirigée de \mathcal{A} vers \mathcal{B} identifie l'élément a de \mathcal{A} le plus éloigné de \mathcal{B} et calcul la distance de a à l'élément de \mathcal{B} le plus proche. $\delta(\mathcal{A}, \mathcal{B})$ sera d'autant plus faible que tous les éléments de \mathcal{B} seront proches de certains éléments de \mathcal{A} (et inversement pour $\delta(\mathcal{B}, \mathcal{A})$). Notons finalement par ζ la distance de Hausdorff dirigée entre un élément a et un ensemble \mathcal{B} :

$$\zeta(a, \mathcal{B}) = \min_{b \in \mathcal{B}} d(a, b).$$

Typiquement, la distance de Hausdorff est utilisée pour apparier un modèle M dans une image I dont on a, au préalable, extrait certains points caractéristiques. Les modèles et images sont alors représentés par des ensembles de points de $\mathbf{X} = \mathbb{R}^2$ et la distance de Hausdorff est utilisée pour mesurer l'étendue, au sens de la norme euclidienne d , avec laquelle chaque point du modèle est proche d'un point de l'image et vice-versa. Par extension de la méthode, certains travaux utilisent des attributs tels que des segments de droite [133]. La fiabilité de la distance de Hausdorff pour la mise en correspondance des primitives dépend essentiellement de la robustesse avec laquelle ces primitives ont été extraites depuis les images.

La distance de Hausdorff mesure la ressemblance entre deux ensembles qui sont à des positions fixes l'un par rapport à l'autre. Pour localiser un modèle dans une image qui a

²— δ n'est pas une distance au sens mathématique du terme car elle n'est pas symétrique en général.

subit une transformation géométrique, D.P. Huttenlocher et al [65] utilise la distance de Hausdorff comme une fonction de la pose relative des ensembles. Cependant, cette méthode est particulièrement coûteuse et plusieurs stratégies de recherche ont été proposées par la suite pour réduire la complexité des algorithmes. Celles-ci reposent généralement sur la partition et la réduction de l'espace des paramètres par des approches probabiliste [94] ou multirésolution [66, 33]. Le lecteur intéressé par ces différents aspects peut se référer au livre de W. Rucklidge [105].

Dans le paragraphe suivant, nous utilisons la distance de Hausdorff, non pas sur des ensembles de points, mais sur l'espace des formes à niveaux de gris. Ce résultat est une extension du travail réalisé pour les contours [56, 30] au cas des images à niveaux de gris.

4.2.2 Expression de la distance de Hausdorff sur l'espace des formes

Par restriction au groupe \mathbb{S}^1 des rotations vectorielles du plan, une forme est constituée de l'ensemble des objets semblables à une rotation près (cf. section 3.2 page 47), soit, dans le domaine de représentation de Fourier-Mellin :

$$\mathbb{F} = \left\{ \left(e^{ik\beta} \mathcal{M}_{f_\sigma}(k, v) \right)_{(k,v) \in \hat{\mathcal{G}}}, \beta \in \mathbb{S}^1 \right\}, \quad \mathcal{M}_{f_\sigma} \in \mathbf{L}^2(\hat{\mathcal{G}}). \quad (4.9)$$

Les classes d'équivalence par rotation sont uniquement constituées d'objets de même taille et les formes deviennent des ensembles fermés et bornés puisque \mathbb{S}^1 est compact. L'espace des formes est alors l'espace quotient $\mathbf{L}^2(\hat{\mathcal{G}})/\mathbb{S}^1$, espace métrique dont la distance naturelle est donnée par la distance de Hausdorff [41]. Montrons à présent que la distance de Hausdorff sur cet espace est précisément le minimum de la fonction $\mathbf{E}(\psi)$ [équation (4.7)].

Soient \mathbb{F} et \mathbb{G} deux formes. La distance de Hausdorff dirigée de \mathbb{F} vers \mathbb{G} s'écrit alors :

$$\begin{aligned} \delta(\mathbb{F}, \mathbb{G}) &= \max_{\varphi \in \mathbb{S}^1} \min_{\phi \in \mathbb{S}^1} \left\| e^{ik\varphi} \mathcal{M}_{f_\sigma}(k, v) - e^{ik\phi} \mathcal{M}_{g_\sigma}(k, v) \right\|_{\mathbf{L}^2(\hat{\mathcal{G}})} \\ &= \max_{\varphi \in \mathbb{S}^1} \min_{\phi \in \mathbb{S}^1} \left\| \mathcal{M}_{f_\sigma}(k, v) - e^{ik(\phi-\varphi)} \mathcal{M}_{g_\sigma}(k, v) \right\|_{\mathbf{L}^2(\hat{\mathcal{G}})} \end{aligned}$$

En effectuant le changement de variable $\psi = \phi - \varphi$, nous obtenons :

$$\begin{aligned} \delta(\mathbb{F}, \mathbb{G}) &= \min_{\psi \in \mathbb{S}^1} \left\| \mathcal{M}_{f_\sigma}(k, v) - e^{ik\psi} \mathcal{M}_{g_\sigma}(k, v) \right\|_{\mathbf{L}^2(\hat{\mathcal{G}})} \\ &= \zeta(f_\sigma, \mathbb{G}) \end{aligned}$$

Le calcul de la distance de Hausdorff dirigée entre deux formes se réduit au calcul de la distance dirigée entre une forme et un représentant de la seconde forme. Comme de plus,

$\zeta(f_\sigma, \mathbb{G}) = \zeta(g_\sigma, \mathbb{F})$, nous obtenons finalement :

$$\begin{aligned} \Delta(\mathbb{F}, \mathbb{G}) &= \delta(\mathbb{F}, \mathbb{G}) = \delta(\mathbb{G}, \mathbb{F}) = \zeta(f_\sigma, \mathbb{G}) = \zeta(g_\sigma, \mathbb{F}) \\ &= \min_{\psi \in \mathbb{S}^1} \left(\int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} \left| \mathcal{M}_{f_\sigma}(k, v) - e^{ik\psi} \mathcal{M}_{g_\sigma}(k, v) \right|^2 dv \right)^{\frac{1}{2}} \end{aligned} \quad (4.10)$$

Ainsi, le calcul de la distance de Hausdorff entre formes est précisément le minimum de la fonction $\mathbf{E}(\psi)$ de l'équation (4.7), ce qui conclut la démonstration.

4.2.3 Application à la détection des symétries d'une image par rotation

Une application intéressante de la distance de Hausdorff sur l'espace des formes concerne la détection des axes de symétries d'une image par rotation. La symétrie des formes est une des primitives de base des objets et est utilisée, par exemple, en robotique pour la reconnaissance, l'inspection et la manipulation d'objets [19]. Une image plane est dite symétrique d'ordre P si elle est invariante pour toutes les rotations de $2\pi/P$ radians par rapport à son barycentre :

$$\forall (r, \theta) \in \mathbb{R}_+^* \times \mathbb{S}^1, \quad f(r, \theta) = f\left(r, \theta + \frac{2\pi l}{P}\right),$$

avec $l = 0, 1, \dots, P - 1$. La méthode que nous proposons consiste à détecter et estimer les P angles définissant les axes de symétrie à l'aide de la distance de hausdorff calculée entre une image et elle-même.

La figure 4.4 illustre les fonctions d'appariement $\mathbf{E}_{f_\sigma, f_\sigma}(\rho, \psi)$ de trois images symétriques³. La première image est une image binaire présentant 9 axes de symétrie. La seconde est un cercle possédant une « infinité de symétries » (cas marginal). Enfin, la troisième représente une image à niveaux de gris caractérisée par trois axes de symétrie. Les résultats nous amènent à plusieurs commentaires, d'une part, sur la forme particulière des spectres des objets et, d'autre part, sur les fonctions d'appariement. Nous proposons ensuite un algorithme permettant de détecter tous les axes de symétrie d'une image.

4.2.3.1 La TFMA des objets symétriques

Les transformées de Fourier-Mellin analytique des trois objets présentent des particularités dues à la symétrie des images : les spectres ont un bon nombre de composantes nulles. Cela provient des fonctions harmoniques circulaires (cf. paragraphe 1.2.1 page 15) intervenant dans l'expression la TFMA [113]. En effet, dans le cas d'une image présentant P axes de symétrie, les FHC s'écrivent :

$$\forall (r, \theta) \in \mathcal{G}, \quad f(r, \theta) = \sum_{m=-\infty}^{\infty} f_{mP}(r) e^{imP\theta},$$

où

$$f_{mP}(r) = \frac{P}{2\pi} \int_0^{\frac{2\pi}{P}} f(r, \theta) e^{-imP\theta} d\theta.$$

³— La distance de hausdorff entre une image et elle-même est la courbe telle que $\rho = 1$.

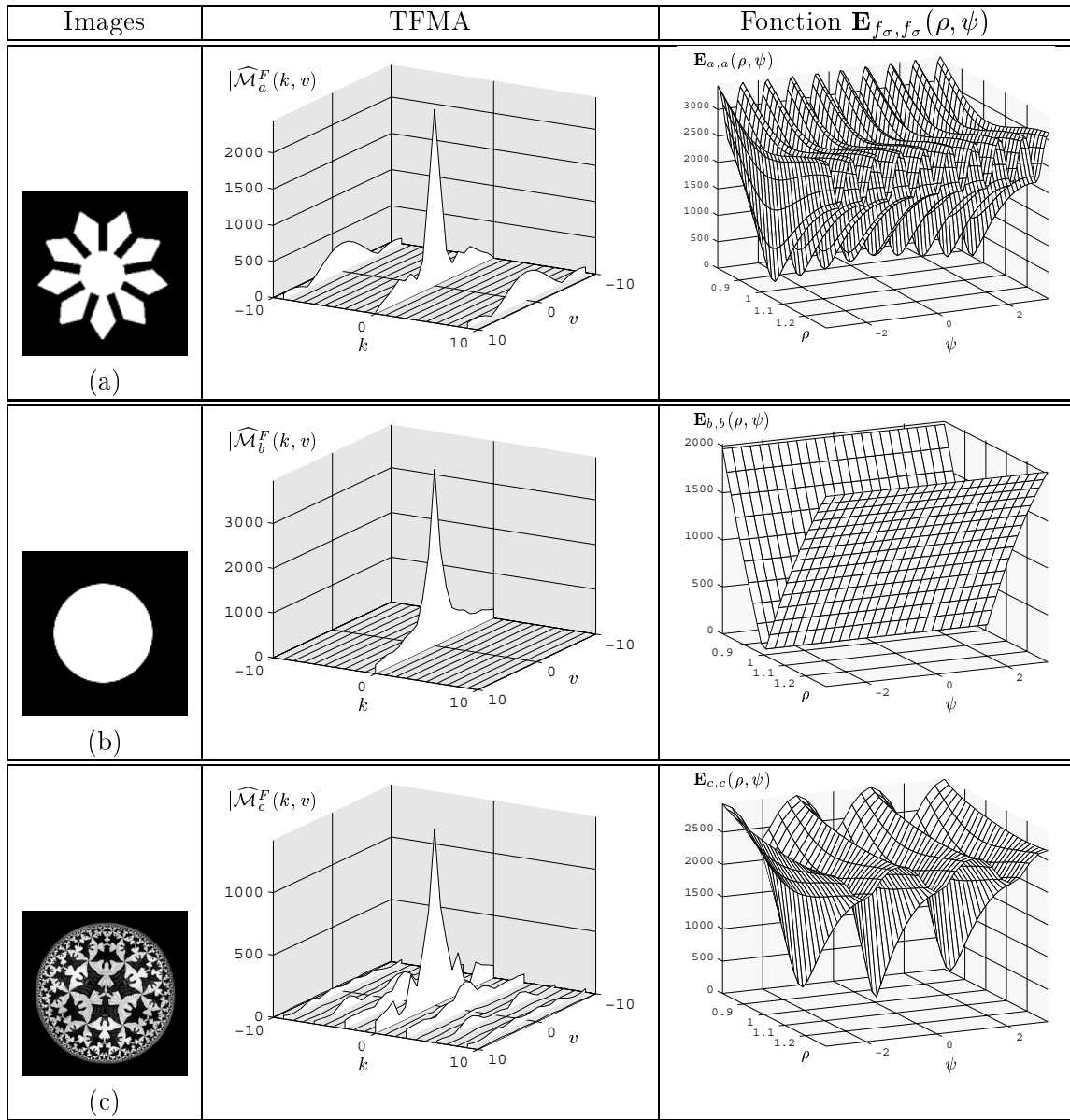


Figure 4.4 — Transformées de Fourier-Mellin analytique (colonne centrale) et fonctions d'appariement (colonne de droite) de trois objets plans possédant des symétries par rotation (colonne de gauche). Les spectres des objets ont été estimés grâce à l'algorithme TFMA-F.

Ainsi, les fonctions harmoniques circulaires $f_k(r)$ d'une image présentant P axes de symétrie par rotation sont nulles pour tout k non multiple de P ⁽⁴⁾. Comme la TFMA peut

⁴— Un résultat similaire a été présenté dans [29] pour les objets-contours décrits par leurs coefficients de Fourier.

s'exprimer en fonction des FHC de l'image :

$$\forall(k, v) \in \hat{\mathcal{G}}, \quad \mathcal{M}_{f_\sigma}(k, v) = \frac{1}{2\pi} \int_0^\infty f_k(r) r^{\sigma-iv} \frac{dr}{r},$$

il vient immédiatement que les harmoniques $\mathcal{M}_{f_\sigma}(k, v)$ sont nulles pour tout k non multiple du nombre d'axes de symétrie dans l'image.

Ceci justifie les résultats numériques obtenus sur les images (a) et (c) présentant respectivement 9 et 3 axes de symétrie par rotation. Dans le cas marginal de l'image (b), le cercle possède une « infinité d'axes de symétrie » et seules les harmoniques pour $k = 0$ sont non nulles. Ce résultat peut être confirmé de manière analytique en développant des calculs semblables à ceux présentés pour les demi-disques.

Remarques : Il existe des images symétriques dont les FHC sont nulles même pour certaines valeurs de k multiples de P (cf. exemples dans [113]). Inversement, certaines images ne possédant pas de symétrie peuvent avoir leurs FHC nulles pour certaines valeurs de k (cf. exemple du demi-disque où les harmoniques de Fourier-Mellin sont nulles pour k pair, voir paragraphe 2.1.2 page 25).

4.2.3.2 Algorithme de détection des axes de symétrie d'une image

Les fonctions d'appariement des trois objets présentent autant de minima qu'il existe d'axes de symétrie dans les images. Pour le cas particulier de l'image d'un disque, tous les angles sont possibles et la fonction d'appariement est minimale pour tout $\psi \in \mathbb{S}^1$. Étant donné l'anisotropie de la grille cartésienne par rotation et les approximations faites lors du calcul des spectres des objets, les minima ne sont pas tous nuls et n'ont pas exactement la même valeur. Néanmoins, ceux-ci se distinguent clairement du reste de la surface, ce qui permet d'envisager une détection robuste.

L'algorithme d'optimisation présenté dans la section précédente ne permet de détecter qu'un seul axe de symétrie⁵. Pour détecter l'ensemble des axes d'une image, nous présentons maintenant un algorithme simple construit à partir de la distance de Hausdorff. Lorsque les fonctions f et g représentent le même objet, l'équation (4.7) se simplifie de la manière suivante :

$$\mathbf{E}_{f_\sigma, f_\sigma}(\psi) = 2 \int_{-\infty}^{+\infty} \sum_{k \in \mathbb{Z}} |\mathcal{M}_{f_\sigma}(k, v)|^2 (1 - \cos(k\psi)) dv \quad (4.11)$$

Minimiser cette quantité (et donc calculer la distance de Hausdorff) revient à chercher l'ensemble des zéros de la dérivée de $\mathbf{E}_{f_\sigma, f_\sigma}$ par rapport à ψ

$$\sum_{k \in \mathbb{Z}} k \sin(k\psi) \int_{-\infty}^{+\infty} |\mathcal{M}_{f_\sigma}(k, v)|^2 dv = 0,$$

par une méthode numérique, et sélectionner les zéros qui minimisent l'équation (4.11). Les angles de rotation ainsi détectés correspondent aux axes de symétrie de l'image. Cet algorithme garantit la détection de tous les axes de symétrie d'un objet à niveaux de gris.

⁵— Cet axe correspond en général à l'angle $\psi = 0$.

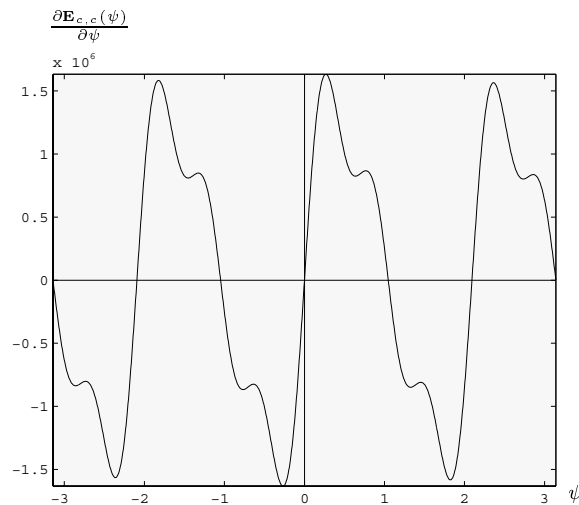


Figure 4.5 — Dérivée de la fonction $\mathbf{E}_{c,c}(\psi)$ calculée sur l'image (c) de la figure 4.4. La dérivée s'annule six fois sur l'intervalle $[0; 2\pi[$.

Pour illustrer la méthode, nous avons appliqué cet algorithme à l'image (c) de la figure 4.4. La figure 4.5 représente le calcul numérique de la dérivée de $\mathbf{E}_{c,c}$ par rapport à ψ . La fonction est impaire⁶ et s'annule six fois dans l'intervalle $[0; 2\pi[$. Pour chacune des valeurs qui annulent la dérivée, le tableau suivant reporte la valeurs de la distance de Hausdorff :

ψ (en deg)	-180	-120	-60	0	60	120
$\widehat{\mathbf{E}}_{c,c}(\psi)$	1420.5	241.4	1424.5	0	1424.5	241.4

Parmi les six zéros, trois d'entre eux présentent une distance de Hausdorff sensiblement plus faible comparée aux autres zéros. Les angles $\psi = -120^\circ, 0^\circ$ et 120° caractérisent les 3 axes de symétrie de l'image (c).

4.3 Approximations de la TFMA et estimation des paramètres de mouvement

Cette section rapporte les résultats d'expériences en discrimination de formes et estimation de mouvement en utilisant la méthode d'appariement présentée dans ce chapitre. Pour chacune des trois méthodes d'approximation de la transformée de Fourier-Mellin, nous étudions l'influence de considérer des spectres finis et discrets, et en particulier les conséquences en terme de reconnaissance d'objets et de robustesse de des paramètres de mouvement puis que ceux-ci dépendent du nombre d'harmoniques utilisées pour calculer $\mathbf{E}(\rho, \psi)$.

Pour ces expériences, nous utilisons les mêmes images de test que celles qui ont permis d'analyser les familles complètes de descripteurs invariants du chapitre précédent. Nous les

⁶— Cette propriété provient de la symétrie hermitienne de la TFMA et de la propriété de parité de la fonction sinus.

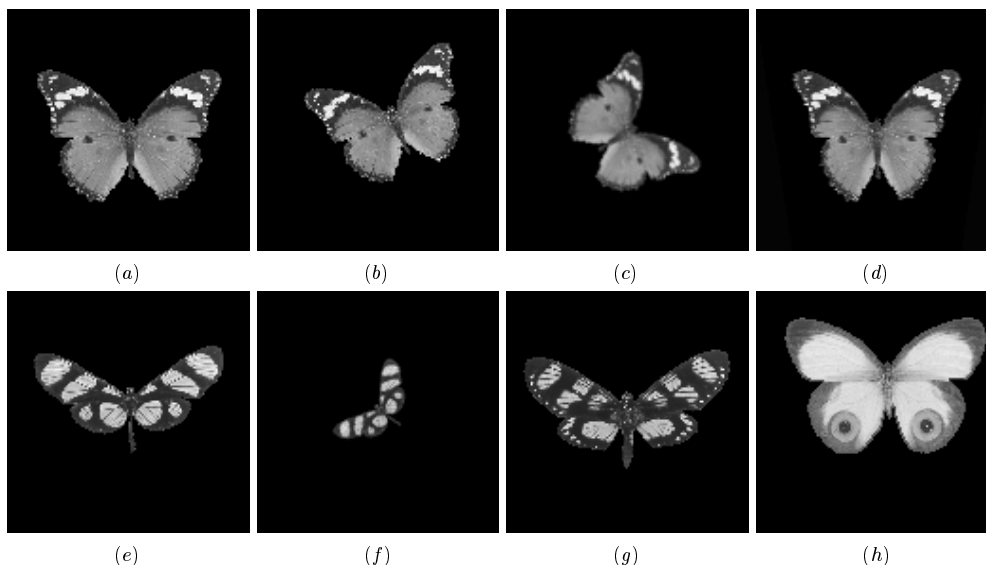


Figure 4.6 — Huit images de test utilisées pour tester la robustesse et la discrimination de l'écart entre formes (128×128 pixels). (a) *Graphium Agamemnon*; (b) image (a) avec $\alpha = 0,9$ et $\beta = -20^\circ$; (c) image (a) avec $\alpha = 0,75$ et $\beta = 60^\circ$; (d) image (a) transformée par étirement; (e) *Methona Confusa*; (f) image (e) avec $\alpha = 0,5$ et $\beta = -50^\circ$; (g) *Gazera Linus*; (h) *Taenaris Macrops*.

rappelons dans la figure 4.6 pour faciliter la lecture et l'interprétation des résultats de cette section. Ainsi, nous pourrons comparer les deux méthodes pour la classification des images

4.3.1 Estimation de mouvement entre des objets similaires

La figure 4.7 illustre l'évolution de l'écart entre formes et des paramètres de mouvement estimés entre des papillons des images (a) et (b) (colonne de gauche), et les papillons des images (e) et (f) (colonne de droite).

Écart entre formes : Dans les deux exemples, l'écart entre formes augmente au fur et à mesure que les dimensions des spectres deviennent plus importantes et traduit l'erreur commise sur les harmoniques de Fourier-Mellin. L'écart calculé à partir de l'approximation directe (algorithme TFMA-D) a tendance à s'accroître plus rapidement que les écarts obtenus à partir des deux autres méthodes d'approximation. Ces deux dernières ont un comportement relativement similaire, l'algorithme TFMA-F fournissant le plus petit écart.

Estimation des paramètres de mouvement : La seconde et la troisième ligne du tableau représentent respectivement l'évolution de l'estimation du facteur d'échelle et de l'angle de rotation en fonction des dimensions des spectres utilisées pour calculer l'écart. Pour chaque graphe, nous avons superposé l'estimation des paramètres de mouvement obtenu par les paramètres de normalisation de la famille étendue (estimations qui initialisent l'algorithme d'optimisation, cf. paragraphe 4.1.2 page 80). L'algorithme TFMA-F montre une remarquable stabilité pour l'estimation des paramètres de mouvement. Ceux-ci restent quasiment inchangés quel que soit les dimensions des spectres. L'approximation cartésienne de la TFMA se montre

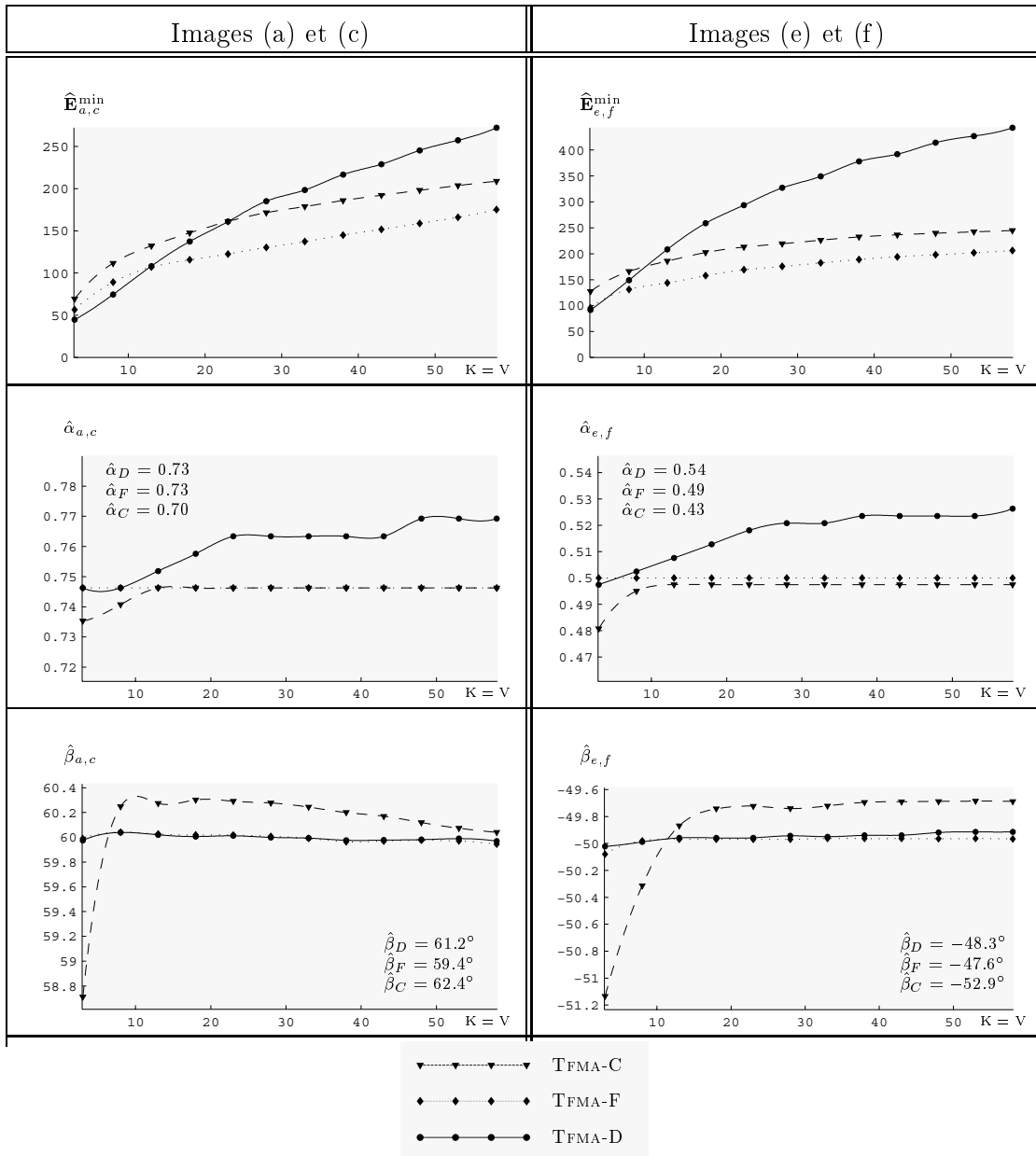


Figure 4.7 — Évolution de l'écart et des paramètres de mouvement estimés entre des objets de même forme pour les trois méthodes d'approximation de la TFMA. En haut : écart entre objets; au milieu : estimations du facteur d'échelle; en bas : estimations de l'angle de rotation.

ici plus robuste aux changements d'échelle qu'aux changements d'orientation, contrairement à l'approximation directe qui montre des résultats inverses.

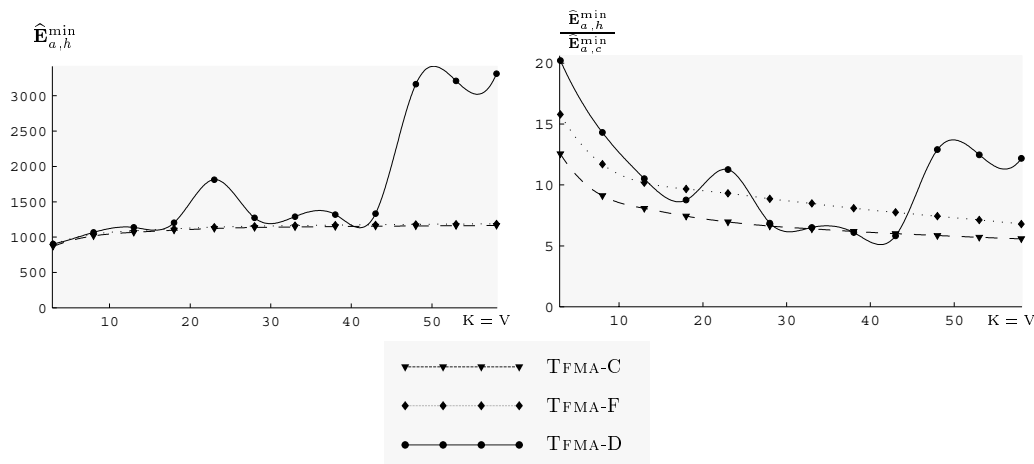


Figure 4.8 — Évaluation de la discrimination de l'écart pour les trois méthodes d'approximation de la TFMA. À gauche : évolution de l'écart entre les papillons (a) et (h) de formes différentes ; à droite : évolution du ratio de discrimination entre les papillons (a) et (h) et les papillons (a) et (b).

4.3.2 Évaluation de la discrimination de l'écart

Comme nous l'avons signalé, l'écart peut être utilisé pour classer des objets indépendamment des similitudes qu'ils ont subits. Il est donc intéressant d'examiner la discrimination de l'écart entre des objets de même forme et des objets de formes différentes. De manière similaire à ce que nous fait pour les familles de descripteurs invariants, la discrimination est évaluée en calculant le ratio des écarts entre des objets de formes différentes et entre des objets de même forme. La figure 4.8 illustre les résultats obtenus entre les papillons (a) et (h) d'une part, et les papillons (a) et (c) d'autre part.

Le graphe de gauche dresse l'évolution de l'écart entre les papillons (a) et (h). Les approximations TFMA-C et TFMA-F fournissent un écart relativement constant, quel que soit les dimensions des spectres ($\hat{\mathbf{E}}_{a,h}^{\min} \simeq 1000$). Par contre, l'approximation TFMA-D a un comportement erratique au-delà du rang $V = K = 17$ et l'écart varie entre 1000 et 3500.

Le graphe de droite dresse l'évolution du ratio de discrimination obtenu en divisant les résultats du graphe de gauche par les résultats obtenus pour les papillons (a) et (c) (cf. figure 4.7). En ce qui concerne l'approximation directe, le comportement de l'écart entre des objets de formes différentes se répercute sur le ratio de discrimination. Bien que relativement important et donc discriminant, le ratio évolue de manière très irrégulière, ce qui est gênant en vue d'une application de classification d'objets. L'approximation TFMA-F s'avère légèrement plus discriminante que l'algorithme TFMA-C. Le gain en discrimination de l'approximation rapide de la TFMA provient essentiellement de la valeur plus faible de l'écart pour des objets de même forme. Pour ces deux approximations, la discrimination est comparable aux résultats obtenus avec la familles étendue (cf. figure 3.11 page 68).

Examinons maintenant le cas d'objets présentant de faibles différences de formes.

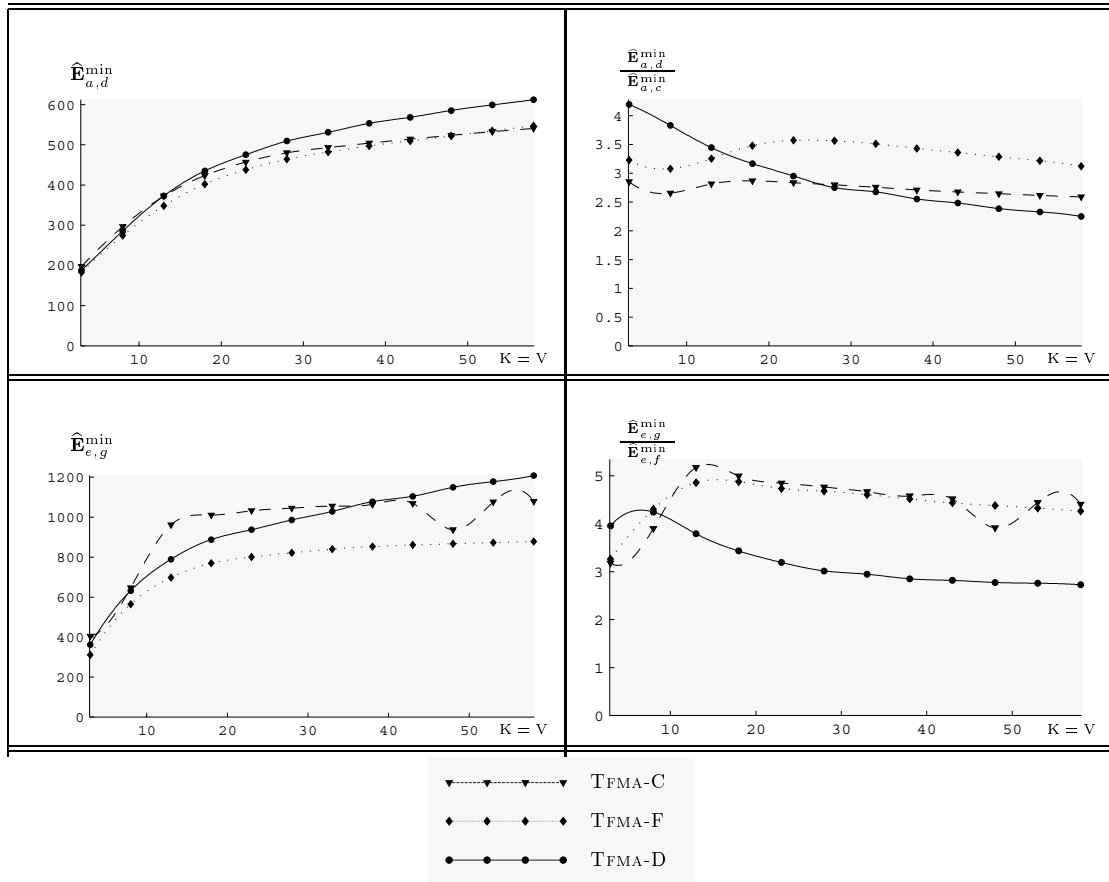


Figure 4.9 — Évolution de l'écart (colonne de gauche) et du ratio de discrimination (colonne de droite) entre des objets de formes presque identiques (papillons (a) et (d) d'une part, et papillons (e) et (g), d'autre part) pour les trois méthodes d'approximation de la TFMA.

4.3.3 Robustesse de l'écart aux faibles déformations

Dans le but d'examiner la robustesse de l'écart pour des objets de forme presque identiques, nous reprenons les exemples des images (a) et (d) (étirement horizontal) et des images (e) et (g) (faible différence de formes). La figure 4.9 dresse l'évolution de l'écart et du ratio de discrimination dans ces deux cas. Les écarts sont très légèrement supérieurs aux résultats obtenus entre des objets de même forme et sensiblement plus faibles que dans le cas d'objets de formes différentes. Les ratios de discrimination sont tout a fait satisfaisants car ils permettent de discriminer des objets de même forme d'objets de formes légèrement différentes, quel que soient les dimensions des spectres et les méthodes pour les estimer.

4.4 Conclusion du chapitre

Dans ce chapitre, nous avons proposé un nouvel algorithme d'appariement d'images à niveaux de gris qui prend en compte la totalité de l'information sur l'objet, contrairement aux méthodes basées sur l'extraction préalable de primitives. Le choix d'une mesure d'apparie-

ment appropriée est primordiale puisqu'elle permet de décider si les objets ont même forme, c'est-à-dire si les paramètres de mouvement estimés ont un sens. Dans le cadre de la représentation des formes proposée dans ce mémoire, nous avons utilisé une mesure de similarité qui s'inspire de la distance euclidienne sur l'espace de représentation de Fourier-Mellin. Cette mesure est un écart entre formes pour le groupe des similitudes et la distance de Hausdorff sur l'espace des formes lorsqu'on restreint le mouvement des objets aux rotations du plan. Une autre mesure de similarité pertinente est donnée par la *corrélation sur le groupe des similitudes planes*, comme extension de la corrélation classique sur \mathbb{R} . Ce travail est actuellement en cours d'évaluation et de comparaison avec la méthode présentée ici [18].

La méthode s'avère particulièrement robuste, tant pour l'estimation des paramètres de mouvement que pour la reconnaissance et la discrimination de formes. Les expériences montrent des résultats comparables à ceux obtenus par la famille étendue de descripteurs invariants. En ce qui concerne les méthodes d'approximation de la TFMA, les conclusions des expériences présentées dans ce chapitre sont essentiellement les mêmes que celles que nous avons tiré dans le chapitre précédent. Nous avons de plus mis en évidence que l'approximation directe est relativement sensible aux changements de taille des objets, alors que l'approximation cartésienne est plus sensible aux changements d'orientation. L'algorithme TFMA-F se montre d'une remarquable stabilité quel que soit les dimensions des spectres utilisées pour estimer l'écart.

L'écart s'avère donc être un bon candidat pour classer des images indépendamment des similitudes qu'elles ont subies. Cependant, les temps de calcul peuvent être relativement importants lorsque les deux objets à comparer ont des formes distinctes. En effet, dans ce cas, l'algorithme d'optimisation oscille entre des minima locaux, multipliant les évaluations de la fonction d'appariement. Ainsi, le coût de classification de bases d'images très hétérogènes peut devenir relativement important.

Application à la Recherche d'Images en vue d'Indexation de Bases par le Contenu

Ce dernier chapitre illustre l'application de la représentation des formes par la TFMA pour l'indexation et la recherche d'images dans des bases d'objets à niveaux de gris. Pour cela, nous utiliserons quelques uns des résultats présentés dans ce mémoire, en particulier, l'approximation de la TFMA (chapitre 2) et la représentation invariante des objets à niveaux de gris (chapitre 3).

L'indexation de bases d'images est un nouveau champ d'application particulièrement actif pour les méthodes de description des formes. Elle consiste à trouver des méthodes efficaces et rapides pour rechercher des images similaires à une image requête parmi de grandes bases d'objets (section 5.1). La similarité est estimée par rapport au contenu de l'image, contrairement aux méthodes plus traditionnelles utilisant des attributs textuels, et se base sur la texture, la couleur, la forme et l'organisation spatiale des images. Dans cette application, il ne s'agit pas tant de déterminer si l'objet requête a exactement un correspondant dans la base de modèles, que de trouver les objets les plus similaires au sens d'une certaine mesure de ressemblance. Il est donc important de vérifier si cette mesure correspond à notre notion visuelle de ressemblance.

Dans la section 5.2, nous présentons les résultats expérimentaux de recherche d'objets à niveaux de gris à l'aide de la représentation des formes par les familles complètes et intégrables de descripteurs invariants introduites dans la chapitre 3. La similarité entre les objets est estimée grâce à la distance sur l'espace des invariants. Les résultats sont illustrés par des expériences menées sur deux bases d'objets réels (une base originale de papillons et la base d'objets de l'université de Columbia). Les résultats de recherche sont comparés avec ceux obtenus, d'une part, avec les invariants de Hu et, d'autre part, avec une nouvelle famille complète de descripteurs invariants issus de la représentation des images par les moments complexes.

Enfin, la conclusion fait le bilan des résultats et dresse quelques perspectives permettant de compléter et d'enrichir la plateforme de tests proposée.

5.1 Principes de la recherche d'images par le contenu

L'essor des dispositifs d'acquisition (caméras numériques, scanner, IRM), des capacités de stockage (mémoire de masse), des systèmes de transmission et de diffusion (réseau Internet, réseaux spécialisés à hauts-débits) a favorisé la création et le développement de bases d'images volumineuses et de natures très diverses. Ceci pose le problème de l'archivage, de l'exploitation et de la consultation de ces bases d'images. Les techniques traditionnelles d'indexation manuelle de ces bases (essentiellement sous forme de mots-clés) sont confrontées à deux problèmes majeurs. Le premier concerne la quantité de travail nécessaire au processus d'annotation textuel. Le second est lié à la richesse du contenu sémantique des images et à la subjectivité de la description.

Pour s'affranchir de ces problèmes, l'accès aux bases par le contenu visuel des images est devenu, depuis une dizaine d'années, un sujet de recherche particulièrement dynamique. Au lieu d'une annotation manuelle par des attributs textuels, les images sont indexées automatiquement par leur propre contenu, par le biais de caractéristiques traduisant la couleur, la forme et la texture des images. La recherche dans ce domaine est particulièrement active et plusieurs systèmes ont été construits. Parmi les moteurs de recherche les plus connus, nous pouvons citer : QBIC [92, 49]¹, VIRAGE [13, 62]², RETRIEVALWARE [42]³ et SURFIMAGE [91]⁴. Les efforts dans ce domaine ont fait l'œuvre de nombreux numéros spéciaux dans les journaux dédiés à ce thème (cf. [1, 4], par exemple) et plusieurs états de l'art ont été présentés [7, 107].

5.1.1 Architecture d'un moteur de recherche

L'architecture typique d'un système d'indexation est schématisée par la figure 5.1. Dans ce modèle, on distingue deux étapes : la phase *off-line* pour les pré-traitements, l'organisation et la gestion de la base d'images de référence (sous-système d'indexation de la base), et la phase de consultation *on-line* pour les traitements propres à chaque nouvelle requête de l'utilisateur (sous-système de recherche). Les deux sous-systèmes ont en commun les deux traitements suivants :

- L'extraction de descripteurs caractéristiques des images, que ce soit une image de la base durant la phase *off-line* ou une image requête proposée par l'utilisateur lors de la phase *on-line* (cf. paragraphe 5.1.2).
- La construction des index à partir des descripteurs. Le rôle principal des index est d'accélérer la recherche des images similaires à la requête durant la phase de consultation (cf. paragraphe 5.1.3).

Pendant la consultation de la base, l'utilisateur sélectionne une image par le biais d'une interface graphique. Les index de la requête sont confrontés aux index des images de références. Finalement, le système sélectionne et présente à l'utilisateur les « images les plus similaires » à la requête.

Dans la suite de cette section, nous reprenons brièvement les étapes d'extraction et d'indexation des descripteurs de formes car elles sont au cœur des systèmes de recherche par le

¹— <http://www.qbic.almaden.ibm.com>

²— <http://www.virage.com/index.html>

³— <http://www.excalib.com/products/rw/rw.html>

⁴— <http://www-rocq.inria.fr/cgi-bin/imedia/surfimage.cgi>

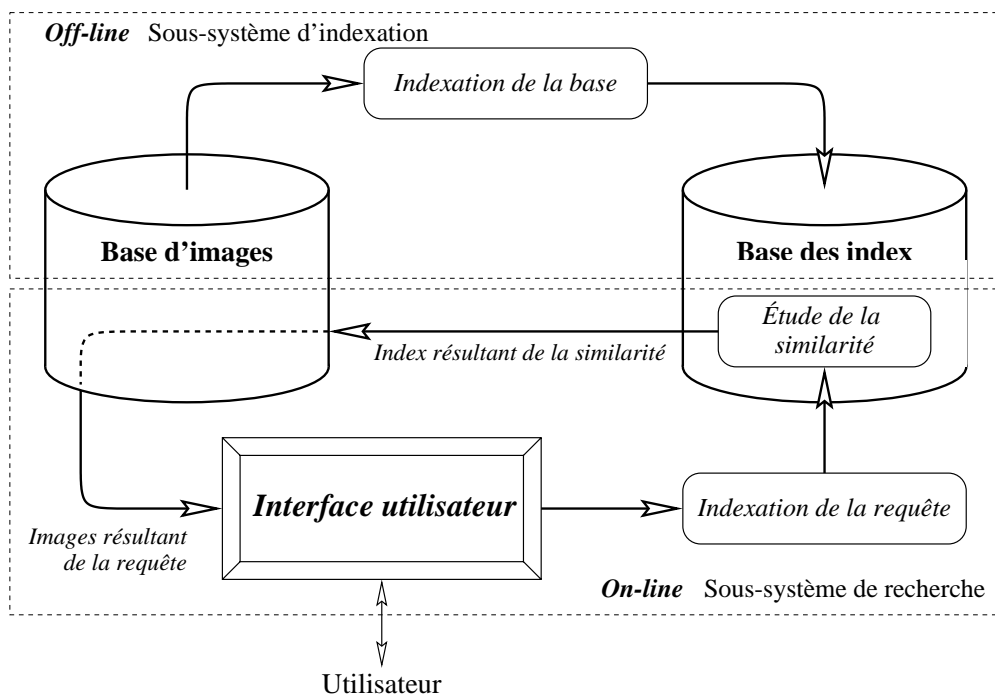


Figure 5.1 — Architecture générique d'un système d'indexation de bases d'images.

contenu. Nous terminons cette section en présentant les méthodes faisant intervenir l'utilisateur dans le processus de recherche pour s'adapter à la subjectivité de ses requêtes.

5.1.2 Description du contenu des images

Un des problèmes centraux de la recherche d'images par le contenu est de choisir une représentation pertinente des images donnant naissance à des primitives visuelles significatives et fiables qui traduisent le contenu sémantique de la base. En général, ces primitives sont regroupées en 3 classes de descripteurs :

- **La couleur :** L'information chromatique est généralement représentée par des caractéristiques (locales et globales) des distributions des couleurs. Les principaux descripteurs sont l'histogramme [120] et les moments des couleurs [119].
- **La texture :** Elle représente une zone de l'image qui possède une certaine homogénéité sans être de niveau de gris constant. Les descripteurs sont nombreux, bénéficiant des résultats de recherche en analyse d'images depuis une trentaine d'années. Nous pouvons citer les matrices de co-occurrence [63] et les propriétés visuelles des textures (notamment la rugosité, la régularité et le contraste) [122]. D'autres méthodes plus récentes sont basées sur les ondelettes, les fractales et les champs de Markov [93].
- **La forme :** Les descripteurs de formes sont soit orientés contour, soit orientés région et sont locaux ou globaux. Pour le cas global, la plupart des descripteurs utilisés dans les systèmes de recherche a été présentée dans le chapitre 1. Les systèmes de recherche utilisent essentiellement des descripteurs de Fourier pour les objets-contours et des moments dans le cas d'objets à niveaux de gris. Pour la description locale des contours, on utilise principalement des primitives géométriques comme des points d'intersection, d'inflexion,

de courbures maximales ou minimales, des segments... [27, 61, 81]. Pour la description locale des objets à niveaux de gris, une solution consiste à calculer des descripteurs à partir de certains points d'intérêts directement extraits de l'image à niveaux de gris [110].

Les descripteurs que nous venons de présenter sont, pour la plupart, invariants par rapport à certaines transformations géométriques, permettant de retrouver les modèles les plus semblables indépendamment de leur pose. Les descripteurs sont souvent ordonnés sous la forme d'un vecteur de signatures et la similarité entre les objets est évaluée en calculant la différence entre les signatures. Le principe de la recherche par le contenu consiste alors à comparer les descripteurs de l'image requête avec les descripteurs de chacun des modèles de la base. Les modèles sont ensuite triés selon leur ordre décroissant de similarité.

Pour s'affranchir d'une comparaison exhaustive avec tous les modèles, les systèmes ont souvent recours à une phase d'indexation.

5.1.3 Indexation des bases d'images

La phase d'indexation contient les opérations nécessaires pour organiser les descripteurs de manière à accéder rapidement aux données. Indexer une image consiste à calculer une clé depuis ses descripteurs par le biais d'une fonction d'indexation. Normalement, des clés similaires correspondent à des images similaires alors que des clés différentes correspondent à des images différentes. La complexité de la comparaison des images est alors réduite à la comparaison de leur clé d'indexation. Comparer deux clés donne une mesure de similarité entre les images mais ne permet pas d'établir une correspondance bijective avec les descripteurs des images [90].

Pour s'affranchir du parcours exhaustif de toute la base d'images, une solution consiste à stocker et organiser les index des images de la base de modèles dans des tables de hachage [73] ou dans des structures arborescentes [24], en associant à chaque index les modèles de la base correspondant. Le principe de ces méthodes est de créer une partition de l'espace des paramètres en classes, regroupant ainsi les modèles dans des régions. Ces structures forment des dictionnaires permettant une recherche accélérée grâce au classement des index. Lors de la phase d'interrogation, cela élimine les parcours inutiles pour aboutir directement à une solution [16].

En général, le nombre de clés est très nettement inférieur au nombre de descripteurs, la fonction de hachage n'est pas injective et les index peuvent correspondre à plusieurs images. Le problème est alors de définir une fonction uniforme pour organiser le dictionnaire de façon optimale par rapport aux descripteurs de formes (minimisant les temps de réponse lors de la phase de recherche). Il n'existe pas de fonction universelle et il est nécessaire de faire un apprentissage sur un échantillon représentatif de la base de modèles. Cela suppose que les descripteurs aient de bonnes propriétés statistiques.

5.1.4 Coopération des descripteurs

Dans le but d'organiser la description du contenu d'une image, le groupe MPEG a débuté un nouveau thème de travail appelé MPEG7, « Multimedia Content Description Interface » [2], dont l'objectif est de spécifier un ensemble de descripteurs pouvant décrire différents contenus

multimédia, et notamment les images [3, 43].

L'extraction de primitives pertinentes est un problème qui ne connaît pas de solution dans le cas générale. D'une part, parce qu'un unique type de descripteurs ne permet pas de caractériser aussi bien une base d'empreintes digitales qu'une base de visages ou une base de tableaux de maîtres par exemple. Le type de descripteur à inclure dans la représentation des images dépend de la nature de la base et de l'application qui en est faite. D'autre part, il semble qu'il n'existe pas d'attributs qui puissent modéliser une base selon tous les points de vues, pour des raisons multiples qui peuvent être illustrées par la subjectivité de la requête de l'utilisateur.

Les moteurs de recherche tentent de combiner des primitives variées pour améliorer la discrimination et la classification. Par conséquent, le nombre de descripteurs de formes par image est de plus en plus élevé. Un des problèmes est alors de déterminer le poids approprié pour chaque type de descripteur. La plupart des moteurs font donc également intervenir l'utilisateur par le moyen d'une interface interactive qui l'associe à l'interprétation des résultats en permettant, par exemple, la sélection de plusieurs requêtes simultanées. Le système recherche alors les modèles les plus semblables à toutes les requêtes présentées. Actuellement, pour améliorer les performances de recherche et pour s'adapter à la subjectivité de la requête de l'utilisateur, les efforts portent sur la conception de systèmes d'apprentissage dynamique qui modifient les poids des descripteurs au fur et à mesure des requêtes lancées par l'utilisateur⁵.

Dans la section suivante, nous mettons l'accent sur l'apport des descripteurs de formes issus de la représentation des images à niveaux de gris par la TFMA pour la recherche et l'indexation de bases d'objets par le contenu.

5.2 Résultats expérimentaux de recherche par le contenu

Dans cette section, nous utilisons les familles complètes et intégrables de descripteurs invariants par similitudes introduites dans le chapitre 3 pour la recherche d'objets à niveaux de gris parmi de grandes bases d'images. Les expériences ont mis en évidence que la famille étendue fournit une représentation numérique des formes robuste et discriminante, deux critères importants pour la recherche d'images par le contenu. Nous proposons ici de vérifier que la distance entre formes, qui provient des propriétés d'unicité et de sommabilité des descripteurs invariants, offre une mesure de similarité intéressante permettant de retrouver des images similaires à une requête.

Nous décrivons d'abord l'algorithme de recherche qui a été mis en œuvre ainsi que le principe de fonctionnement de la plateforme de test et l'interface graphique qui permet à l'utilisateur d'interroger la base. Ensuite, nous présentons quelques résultats de recherche sur deux bases de test. Nous comparons notamment les résultats obtenus avec deux ensembles de descripteurs invariants issus des moments de l'image.

⁵— Exemples : moteurs de recherche *FOUREYES* et *PHOTOBOOK* [87, 98] (<http://www.media.mit.edu/~tpminka/photobook/foureyes/>), *MARS* [108] (<http://jadzia.ifp.uiuc.edu:8000>) et *PICHUNTER* [28].

5.2.1 Algorithme de recherche et bases de test

Le principe de l'algorithme de recherche est relativement simple et profite des caractéristiques de la famille de descripteurs invariants. Théoriquement, l'unicité de la représentation des objets se traduit par un nombre infini de descripteurs invariants, ce qui peut apparaître comme contradictoire avec les contraintes liées aux traitements numériques. Cependant, cette propriété nous donne la possibilité d'extraire un nombre de descripteurs de forme aussi important que souhaité. En s'appuyant sur la propriété de sommabilité, une stratégie systématique d'extraction consiste à sélectionner les descripteurs centraux et à élargir l'ensemble ainsi formé en augmentant l'ordre des descripteurs. Ainsi, des ensembles suffisamment importants de descripteurs n'indexent qu'un seul objet de la base de modèles. Le nombre de descripteurs peut donc être ajusté en fonction de la nature de la base de modèles.

Algorithme de recherche :

En se référant au schéma présentant l'architecture générique d'un système d'indexation de bases d'images (cf. figure 5.1 page 97), la phase d'indexation consiste à calculer et à stocker en mémoire un nombre relativement important de descripteurs (quelques centaines) pour toutes les images de la base de modèles. Pour la phase de recherche, la distance entre formes (cf. équation 3.15 page 54) est utilisée directement pour mesurer la similarité entre une requête et les objets de la base, indépendamment de leur position, de leur orientation et de leur taille. Ainsi, la recherche d'un objet inconnu g parmi un ensemble de N objets de référence (la base de modèles) f_i est réalisée à travers ces différentes étapes :

- D'abord, on extrait les descripteurs invariants de g ,
- Puis, on estime la distance entre les invariants de g et les invariants de tous les modèles f_i ,
- Enfin, on sélectionne les M modèles les plus similaires à la requête, c'est-à-dire ceux qui présentent la distance la plus faible avec g .

La recherche dans la base de modèles est exhaustive. La mise en place de structures arborescentes permettant d'accélérer la sélection des modèles les plus similaires à la requête reste une perspective importante pour le développement de cette application, comme nous le signalerons en conclusion. L'avantage d'indexer directement les images par leurs descripteurs invariants est que la méthode ne nécessite pas de phase préalable d'apprentissage sur la base de modèles. Ainsi, on peut ajouter ou supprimer des modèles sans avoir à reconstruire tous les index. La mise à jour d'une base est de ce fait très rapide puisqu'il suffit d'y ajouter ou supprimer les descripteurs des nouveaux modèles, sans avoir à reconstruire l'ensemble des index.

Présentation des deux bases de test :

La première base d'objets est une *base d'images de papillons*. Elle est constituée de 98 images représentant 70 papillons différents ; nous y avons déjà eu recours à plusieurs reprises dans les chapitres précédents (cf. annexe C.1 page 121). Les papillons présentent une forte homogénéité de forme et la principale différence se situe dans la texture des ailes. Des descripteurs basés uniquement sur le contour extérieur des papillons seraient sans doute insuffisants pour les discriminer. Cette base se montre donc particulièrement adaptée pour tester l'invariance et la discrimination des descripteurs.

La seconde base utilisée pour nos expériences est la *base d'objets de l'université de Columbia*. Elle est constituée de 1440 images représentant 20 objets 3D pris dans 72 positions différentes (par rotation de 5° autour d'un axe vertical)⁶. Quelques images des 20 objets sont présentées en annexe C.2 page 122. La principale difficulté de cette base réside dans le fait que le mouvement apparent des objets fausse sensiblement le modèle de transformation géométrique que nous considérons ici. Cette base permet de tester la stabilité et la robustesse de nos descripteurs invariants aux déformations de formes.

Les différentes étapes de recherche ont été implémentées dans une plateforme de test. Le paragraphe suivant en décrit le principe de fonctionnement.

5.2.2 Principe du système de recherche et interface utilisateur

La plateforme de test que nous avons développée repose sur une architecture de type client-serveur et est accessible depuis le WEB. La figure 5.2 présente l'aspect de l'interface utilisateur alors que le principe de fonctionnement est schématisé dans la figure 5.3.

Interface utilisateur :

L'interface graphique est composée de deux parties (figure 5.2) :

- La partie supérieure comprend les trois boutons et la fenêtre de résultats. Une requête peut être lancée soit en désignant une des images présentées à l'écran, soit en sélectionnant directement le nom de l'image dans la liste. Deux autres options sont également possibles : la première permet de naviguer dans la base de modèles en proposant une sélection d'images ; la seconde permet de lancer une requête à partir d'une image sélectionnée de manière aléatoire par le système⁷).
- La partie inférieure qui comprend les champs de sélection des paramètres permet à l'utilisateur :
 - de sélectionner une base de modèles et de modifier la présentation des résultats à l'écran,
 - de modifier les paramètres qui ont une influence sur l'algorithme de recherche, notamment la famille de descripteurs invariants, la méthode d'approximation de la TFMA et le nombre de descripteurs invariants utilisés pour estimer la similarité.

Principe du système de recherche :

Le principe de fonctionnement du système de recherche est le suivant (cf. figure 5.3) :

- L'utilisateur émet une requête en sélectionnant une image à partir de l'interface graphique. La requête est caractérisée par un certain nombre de paramètres parmi lesquels le nom de l'image, la base de modèles, la famille de descripteurs invariants, le nombre de descripteurs invariants, etc. Elle permet également de sélectionner la méthode d'approximation de la TFMA utilisée pour calculer les descripteurs invariants.
- Le serveur lance un programme CGI⁸ (langage C) qui effectue la recherche parmi la

⁶— Cette base sert souvent de référence pour comparer les méthodes d'indexation et de recherche par le contenu. Elle est accessible à l'adresse <http://www.media.mit.edu/~tpminka/photobook/foureyes/>.

⁷— Ceci est surtout utile pour effectuer de nombreux tests.

⁸— CGI pour *Common Gateway Interface*

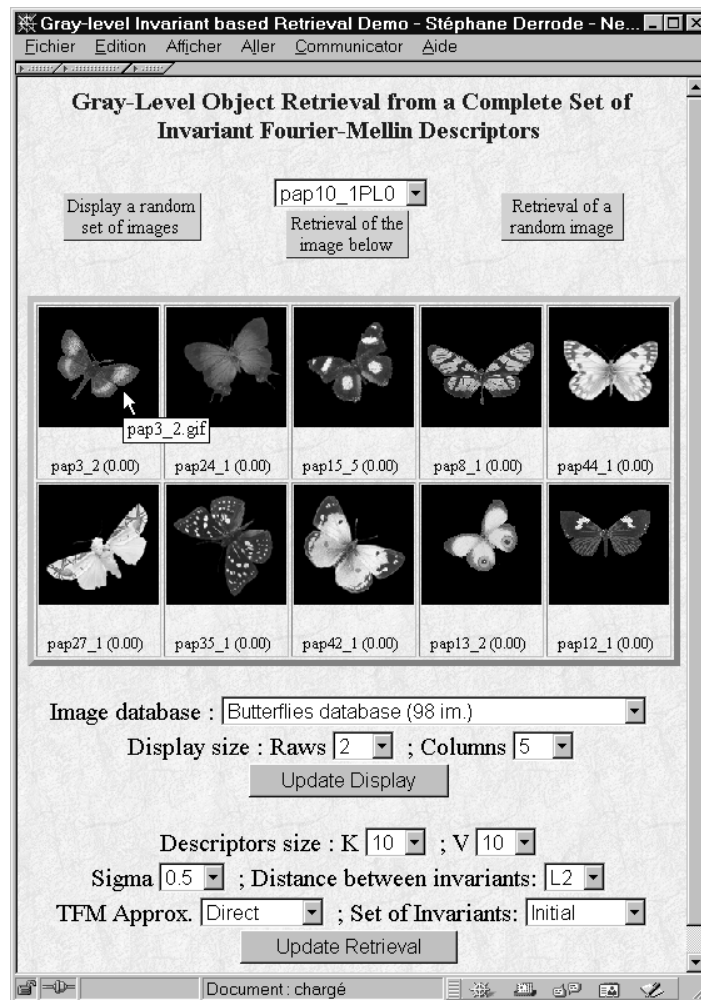


Figure 5.2 — Interface utilisateur pour l'interrogation des bases d'objets.

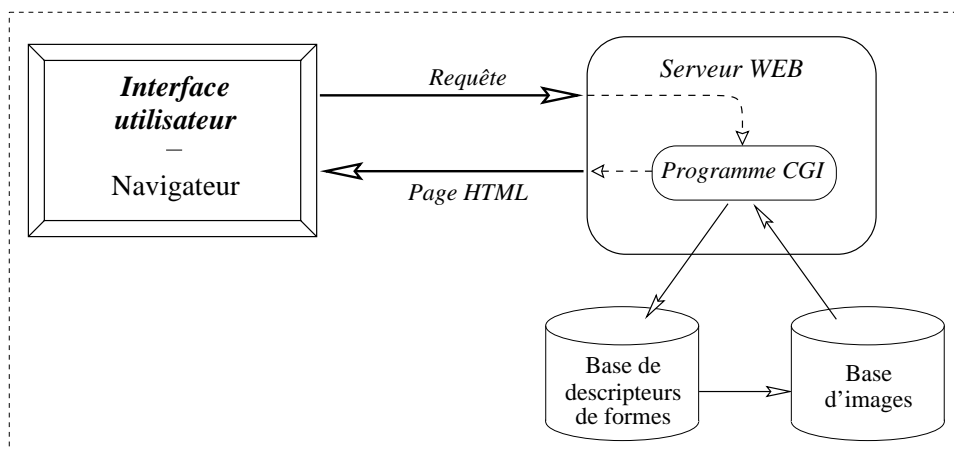


Figure 5.3 — Principe de fonctionnement de la plateforme de test.

base de modèles selon l'algorithme présenté dans le paragraphe précédent, et retourne au serveur les images les plus similaires sous la forme d'une page HTML.

- Enfin, la page HTML est retransmise au navigateur par le biais du serveur et l'utilisateur voit apparaître le résultat de sa requête. Le système met donc à jour de manière dynamique l'interface utilisateur au fur et à mesure des réponses aux différentes requêtes.

Pour les expériences des paragraphes suivants, l'image requête fait partie de la base de modèles. Cela ne pose pas de problème particulier car la méthode de recherche est exhaustive et déterministe. Ainsi, la fenêtre de résultats présente toujours l'image requête en première position avec une distance exactement nulle. Les huit images les plus similaires à la requête sont alors affichées dans l'ordre croissant de la distance (de gauche à droite et de bas en haut). Pour chacune d'elles, nous précisons la valeur de la distance entre parenthèses.

Dans les sections suivantes, nous utilisons cette interface pour présenter différents résultats de recherche obtenus à partir des deux bases d'objets.

5.2.3 Quelques résultats pour les trois méthodes d'approximation de la TFMA

Dans ce paragraphe, nous présentons les résultats de recherche d'une image de la base de papillons et d'une image de la base de Columbia. Ces résultats sont organisés sous la forme d'un tableau :

- Les lignes correspondent au nombre de descripteurs invariants utilisés pour mesurer la similarité entre les images (5, 41 et 181).
- Les colonnes correspondent aux trois méthodes d'approximation de la TFMA utilisées pour estimer les descripteurs invariants (algorithmes direct TFMA-D, rapide TFMA-F et cartésien TFMA-C).

La figure 5.4 présente les résultats de recherche pour l'image requête `Pap28_2`. Dans le cas des approximations directe et rapide, nous retrouvons bien en priorité les quatre images similaires à la requête et l'image du papillon de forme légèrement différente (`Pap28_4`), quel que soit le nombre de descripteurs invariants utilisés pour mesurer la similarité. Par contre, pour l'approximation cartésienne, les résultats de recherche se détériorent au fur et à mesure que le nombre de descripteurs invariants augmente. L'image `Pap28_3` qui présente une différence de taille importante avec la requête, n'apparaît jamais dans les huit premières images sélectionnées. Ce constat confirme les remarques des chapitres précédents concernant les problèmes d'isotropie de la grille cartésienne par rapport aux facteurs d'échelle en particulier.

La figure 5.5 montre les résultats de recherche obtenus pour l'image requête `Obj19_48` de la base de Columbia. Un aperçu général du tableau révèle que les premières images sélectionnées représentent bien le même objet que la requête, même si des images d'un second objet viennent s'insérer lorsque le nombre de descripteurs invariants augmente. Il est intéressant de noter que les descripteurs invariants estimés par les trois méthodes d'approximation de la TFMA retrouvent en premier lieu les mêmes images, même si leur ordre de priorité est légèrement différent pour les trois méthodes d'approximation de la TFMA.

Dans le paragraphe suivant, nous proposons de comparer les résultats des descripteurs issus de l'approximation rapide de la TFMA (algorithme TFMA-F) avec des méthodes de représentation globale des objets issues des moments géométriques de l'image.





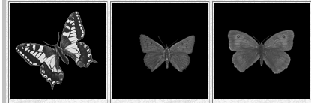


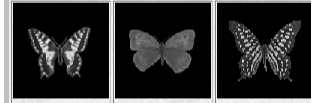




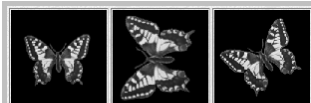











D	Approx. directe	Approx. rapide	Approx. cartésienne	
5	 <p>pap28_2 (0.00) pap28_5 (0.03) pap28_3 (0.03)</p>  <p>pap28_1 (0.04) pap28_4 (0.06) pap1_2 (0.14)</p>  <p>pap43_1 (0.15) pap1_1 (0.17) pap60_1 (0.18)</p>	 <p>pap28_2 (0.00) pap28_5 (0.04) pap28_4 (0.05)</p>  <p>pap28_1 (0.06) pap47_1 (0.08) pap43_1 (0.12)</p>  <p>pap28_3 (0.13) pap1_2 (0.14) pap59_1 (0.15)</p>	 <p>pap28_2 (0.00) pap28_1 (0.04) pap28_5 (0.06)</p>  <p>pap28_4 (0.09) pap43_1 (0.15) pap1_1 (0.16)</p>  <p>pap1_2 (0.16) pap15_4 (0.17) pap60_1 (0.19)</p>	
	41	 <p>pap28_2 (0.00) pap28_3 (0.15) pap28_5 (0.22)</p>  <p>pap28_1 (0.24) pap28_4 (0.27) pap1_2 (0.63)</p>  <p>pap1_1 (0.65) pap47_1 (0.73) pap59_1 (0.73)</p>	 <p>pap28_2 (0.00) pap28_5 (0.18) pap28_1 (0.19)</p>  <p>pap28_4 (0.27) pap28_3 (0.35) pap1_1 (0.63)</p>  <p>pap1_2 (0.72) pap59_1 (0.78) pap47_1 (0.82)</p>	 <p>pap28_2 (0.00) pap28_1 (0.14) pap28_4 (0.28)</p>  <p>pap1_1 (0.60) pap28_5 (0.62) pap43_1 (0.66)</p>  <p>pap47_1 (0.69) pap1_2 (0.70) pap15_2 (0.75)</p>
		181	 <p>pap28_2 (0.00) pap28_3 (0.23) pap28_5 (0.37)</p>  <p>pap28_4 (0.40) pap28_1 (0.40) pap1_2 (0.86)</p>  <p>pap1_1 (0.92) pap47_1 (0.93) pap43_1 (0.96)</p>	 <p>pap28_2 (0.00) pap28_5 (0.30) pap28_1 (0.31)</p>  <p>pap28_4 (0.40) pap28_3 (0.52) pap1_1 (0.93)</p>  <p>pap1_2 (0.96) pap47_1 (1.01) pap43_1 (1.06)</p>

Figure 5.4 — Résultats de recherche de l'image du papillon Pap28_2 avec 5 ($K = V = 1$), 41 ($K = V = 4$) et 181 ($K = V = 9$) descripteurs invariants pour les trois méthodes d'approximation de la TFMA.

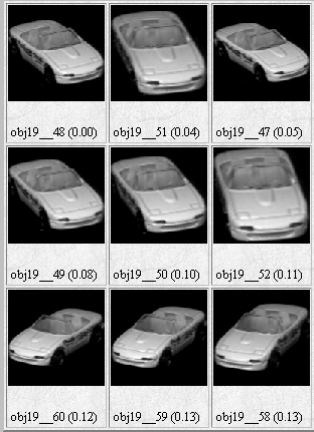
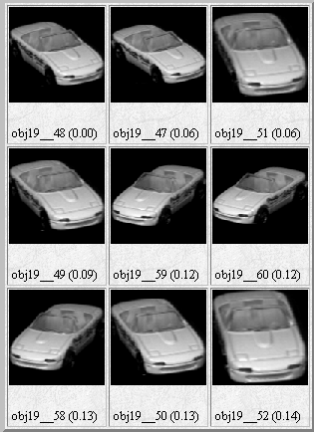
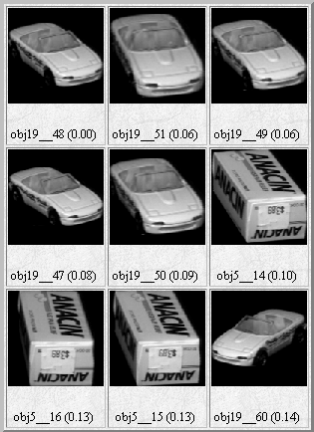
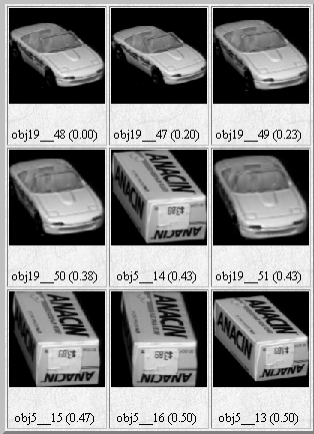
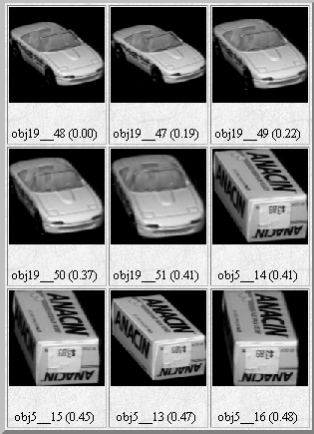
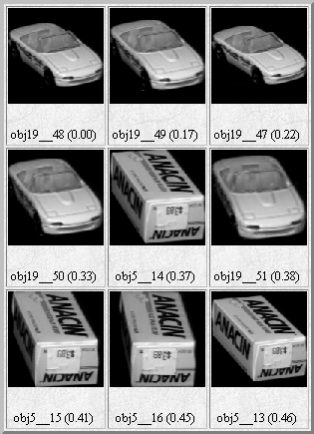
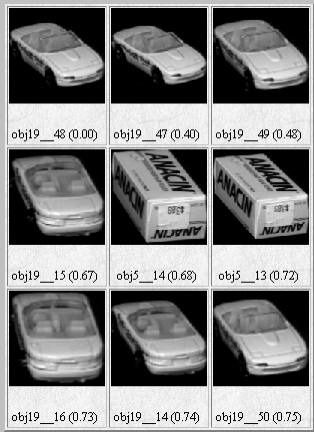
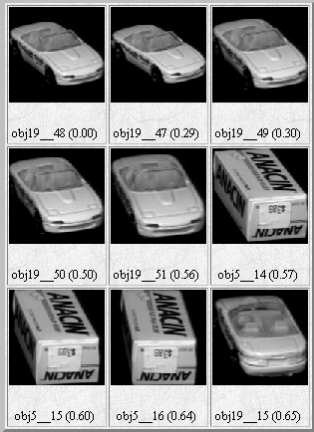
D	Approx. directe	Approx. rapide	Approx. cartésienne	
5	 <p>obj19_48 (0.00) obj19_51 (0.04) obj19_47 (0.05)</p> <p>obj19_49 (0.08) obj19_50 (0.10) obj19_52 (0.11)</p> <p>obj19_60 (0.12) obj19_59 (0.13) obj19_58 (0.13)</p>	 <p>obj19_48 (0.00) obj19_47 (0.06) obj19_51 (0.06)</p> <p>obj19_49 (0.09) obj19_59 (0.12) obj19_60 (0.12)</p> <p>obj19_58 (0.13) obj19_50 (0.13) obj19_52 (0.14)</p>	 <p>obj19_48 (0.00) obj19_51 (0.06) obj19_49 (0.06)</p> <p>obj19_47 (0.08) obj19_50 (0.09) obj5_14 (0.10)</p> <p>obj5_16 (0.13) obj5_15 (0.13) obj19_60 (0.14)</p>	
	41	 <p>obj19_48 (0.00) obj19_47 (0.20) obj19_49 (0.23)</p> <p>obj19_50 (0.38) obj5_14 (0.43) obj19_51 (0.43)</p> <p>obj5_15 (0.47) obj5_16 (0.50) obj5_13 (0.50)</p>	 <p>obj19_48 (0.00) obj19_47 (0.19) obj19_49 (0.22)</p> <p>obj19_50 (0.37) obj19_51 (0.41) obj5_14 (0.41)</p> <p>obj5_15 (0.45) obj5_13 (0.47) obj5_16 (0.48)</p>	 <p>obj19_48 (0.00) obj19_49 (0.17) obj19_47 (0.22)</p> <p>obj19_50 (0.33) obj5_14 (0.37) obj19_51 (0.38)</p> <p>obj5_15 (0.41) obj5_16 (0.45) obj5_13 (0.46)</p>
		181	 <p>obj19_48 (0.00) obj19_47 (0.40) obj19_49 (0.48)</p> <p>obj19_15 (0.67) obj5_14 (0.68) obj5_13 (0.72)</p> <p>obj19_16 (0.73) obj19_14 (0.74) obj19_50 (0.75)</p>	 <p>obj19_48 (0.00) obj19_47 (0.29) obj19_49 (0.30)</p> <p>obj19_50 (0.50) obj19_51 (0.56) obj5_14 (0.57)</p> <p>obj5_15 (0.60) obj5_16 (0.64) obj19_15 (0.65)</p>

Figure 5.5 — Résultats de recherche de l'objet Co119_48 avec 5 ($K = V = 1$), 41 ($K = V = 4$) et 181 ($K = V = 9$) descripteurs invariants pour les trois méthodes d'approximation de la TFMA.

5.2.4 Comparaison avec des descripteurs des moments

Afin de mesurer l'apport des descripteurs de forme issus de la TFMA pour l'indexation de bases d'images, nous proposons de comparer les résultats de recherche avec deux ensembles de descripteurs invariants par similitudes planes issus des moments de l'image. Le premier ensemble est constitué des sept invariants de Hu. Leur expression est donnée dans l'annexe B (équation B.4, page 120). Ces descripteurs sont utilisés par exemple dans le moteur de recherche QBIC. Le second ensemble de descripteurs provient d'une nouvelle famille complète d'invariants issue des moments complexes de l'image (cf. paragraphe 1.2.2.1 page 16). Nous la présentons maintenant.

En utilisant la relation liant les moments complexes de deux images de positions, d'orientations et de tailles différentes, et en appliquant le procédé de normalisation du chapitre 3, nous avons construit une nouvelle famille de descripteurs invariants par similitude issus des moments complexes [82] :

$$\mathcal{C}_f(p, q) = (\Lambda_f)^{p+q+2} e^{i(p-q)\Psi_f} c_f(p, q) \quad (5.1)$$

où les coefficients c_f désignent les moments complexes de l'image f . Λ_f et Ψ_f sont les paramètres de normalisation de la famille. Pour obtenir l'invariance, ces paramètres peuvent être choisis de la manière suivante :

$$\begin{cases} \Lambda_f = |c_f(p_0, q_0)|^{\frac{-1}{p_0+q_0+2}} \\ \Psi_f = -\arg(c_f(p_1 + 1, p_1)), \end{cases} \quad (5.2)$$

où p_0, q_0 et p_1 désignent trois entiers naturels quelconques.

On vérifie aisément que cette famille est complète et représente les formes de manière unique. Pour mesurer la similarité entre deux objets f et g indépendamment des similitudes, nous utilisons une erreur quadratique entre les descripteurs invariants :

$$\xi_2(f, g) = \left(\sum_{p \in \mathbb{N}} \sum_{q \in \mathbb{N}} |c_f(p, q) - c_g(p, q)|^2 \right)^{\frac{1}{2}} \quad (5.3)$$

Les deux paragraphes suivants comparent quelques résultats de recherche pour ces trois ensembles d'invariants.

5.2.4.1 Base de Papillons

Exemple 1 : La figure 5.6 présente les résultats de recherche obtenus pour l'image `Pap13_1` pour les trois ensembles de descripteurs invariants. Toutes les méthodes retrouvent en priorité les deux images similaires à la requête, et cela quelque soit le nombre de descripteurs invariants utilisés. Cela permet de vérifier l'invariance aux similitudes des deux ensembles issus des moments. On remarque que les images classées sont essentiellement les mêmes pour les trois ensembles.

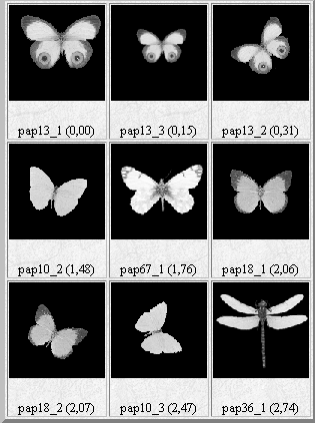
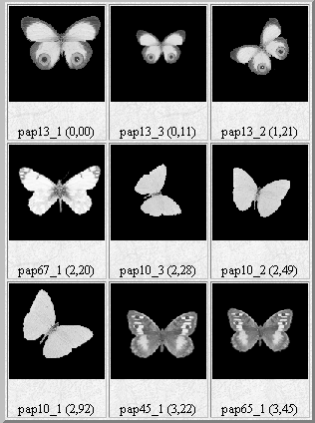
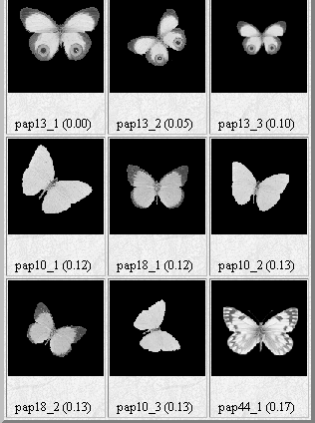
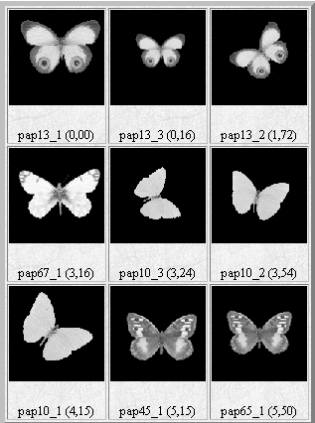
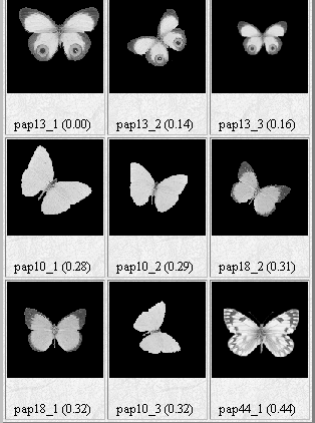
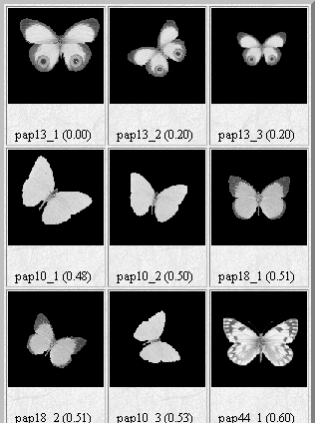
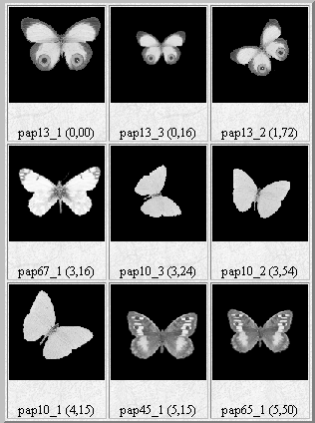
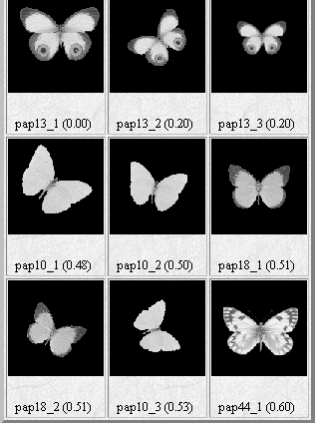
D	Inv. de Hu (10^{-12})	Inv. des mom. cplx. (10^{-6})	Inv. de la TFMA	
5	 <p>pap13_1 (0,00) pap13_3 (0,15) pap13_2 (0,31)</p> <p>pap10_2 (1,48) pap67_1 (1,76) pap18_1 (2,06)</p> <p>pap18_2 (2,07) pap10_3 (2,47) pap36_1 (2,74)</p>	 <p>pap13_1 (0,00) pap13_3 (0,11) pap13_2 (1,21)</p> <p>pap67_1 (2,20) pap10_3 (2,28) pap10_2 (2,49)</p> <p>pap10_1 (2,92) pap45_1 (3,22) pap65_1 (3,45)</p>	 <p>pap13_1 (0,00) pap13_2 (0,05) pap13_3 (0,10)</p> <p>pap10_1 (0,12) pap18_1 (0,12) pap10_2 (0,13)</p> <p>pap18_2 (0,13) pap10_3 (0,13) pap44_1 (0,17)</p>	
	40	 <p>pap13_1 (0,00) pap13_3 (0,16) pap13_2 (1,72)</p> <p>pap67_1 (3,16) pap10_3 (3,24) pap10_2 (3,54)</p> <p>pap10_1 (4,15) pap45_1 (5,15) pap65_1 (5,50)</p>	 <p>pap13_1 (0,00) pap13_2 (0,14) pap13_3 (0,16)</p> <p>pap10_1 (0,28) pap10_2 (0,29) pap18_2 (0,31)</p> <p>pap18_1 (0,32) pap10_3 (0,32) pap44_1 (0,44)</p>	 <p>pap13_1 (0,00) pap13_2 (0,20) pap13_3 (0,20)</p> <p>pap10_1 (0,48) pap10_2 (0,50) pap18_1 (0,51)</p> <p>pap18_2 (0,51) pap10_3 (0,53) pap44_1 (0,60)</p>
		180	 <p>pap13_1 (0,00) pap13_3 (0,16) pap13_2 (1,72)</p> <p>pap67_1 (3,16) pap10_3 (3,24) pap10_2 (3,54)</p> <p>pap10_1 (4,15) pap45_1 (5,15) pap65_1 (5,50)</p>	 <p>pap13_1 (0,00) pap13_2 (0,20) pap13_3 (0,20)</p> <p>pap10_1 (0,48) pap10_2 (0,50) pap18_1 (0,51)</p> <p>pap18_2 (0,51) pap10_3 (0,53) pap44_1 (0,60)</p>

Figure 5.6 — Résultats de recherche de l'objet Pap13_1 avec environ 5, 40 et 180 descripteurs invariants. Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA.

Exemple 2 : La figure 5.7 présente les résultats de recherche obtenus pour l'image `Pap15_1`. Ce papillon est représenté par quatre images similaires dans la base. Dans le cas des invariants de Hu, les quatre images sont bien présentes mais ne sont pas placées exactement aux quatre premières places. Certains papillons de formes légèrement différentes sont mieux classés, ce qui peut être directement imputé aux problèmes bien connus d'approximation des moments géométriques de l'image. Par contre, les invariants issus des moments complexes permettent de retrouver toutes les images similaires, quel que soit le nombre d'invariants utilisés. Enfin, pour les invariants issus de la TFMA, nous retrouvons trois images sur quatre pour $D = 5$ et les quatre images pour des ensembles de descripteurs plus importants ($D = 40$ et $D = 180$).

5.2.4.2 Base de Columbia

Dans ce paragraphe, nous présentons deux résultats de recherche dans la base de Columbia parmi les plus défavorables que nous ayons rencontrés lors de nos expériences. Ceci nous permet de comparer la robustesse de chacun des ensembles de descripteurs.

Les figures 5.8 et 5.9 illustrent les résultats de recherche obtenus pour les images `obj5_64` et `obj8_1`. Pour ces deux exemples, les invariants de Hu ne permettent de retrouver aucun objet similaire à l'image requête, ce qui confirme qu'ils ne sont pas très stables par rapport aux faibles variations de formes. Les invariants issus des moments complexes permettent de retrouver entre 2 et 4 images similaires dans le premier exemple, et entre 4 et 5 dans le second exemple. Les descripteurs issus de la TFMA permettent, quant à eux, de retrouver entre 7 et 9 images similaires dans le premier cas, et entre 7 et 8 dans le second cas. Les résultats de recherche sont donc sensiblement meilleurs dans le cas de la TFMA. Ceci a pu être confirmé par de nombreuses expériences complémentaires.

Les expériences montrent que les descripteurs développés dans le cadre du 3^e chapitre de ce mémoire peuvent sensiblement améliorer les résultats de recherche.

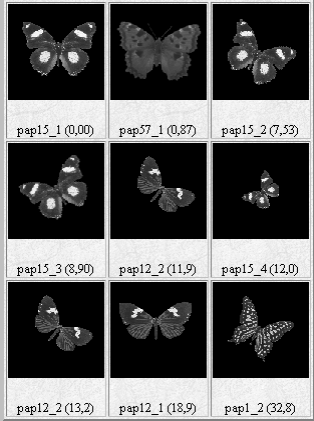
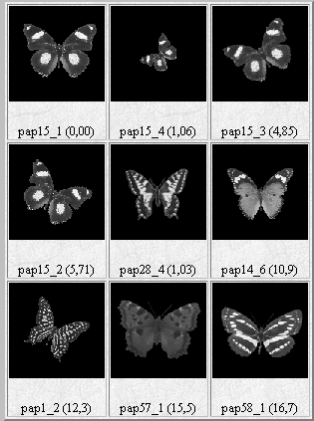
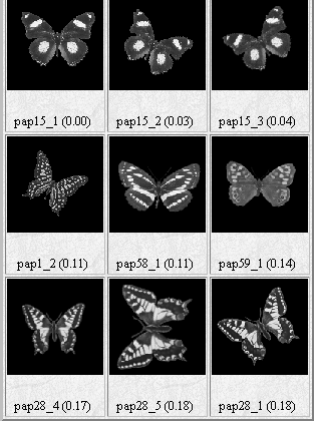

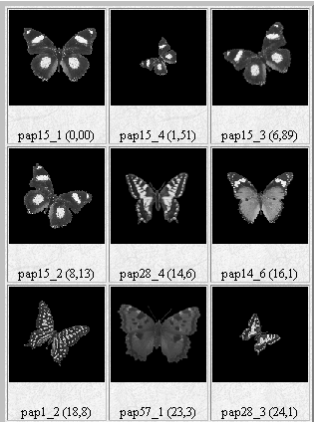
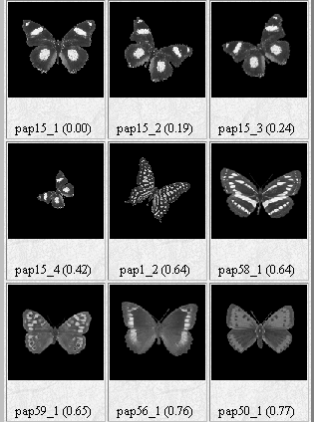

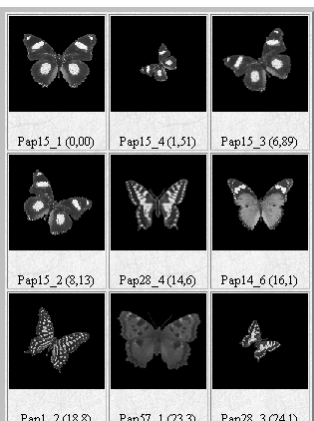
D	Inv. de Hu (10^{-12})	Inv. des mom. cplx. (10^{-6})	Inv. de la TFMA	
5	 <p>pap15_1 (0,00) pap57_1 (0,87) pap15_2 (7,53)</p> <p>pap15_3 (8,90) pap12_2 (11,9) pap15_4 (12,0)</p> <p>pap12_2 (13,2) pap12_1 (18,9) pap1_2 (32,8)</p>	 <p>pap15_1 (0,00) pap15_4 (1,06) pap15_3 (4,85)</p> <p>pap15_2 (5,71) pap28_4 (1,03) pap14_6 (10,9)</p> <p>pap1_2 (12,3) pap57_1 (15,5) pap58_1 (16,7)</p>	 <p>pap15_1 (0,00) pap15_2 (0,03) pap15_3 (0,04)</p> <p>pap1_2 (0,11) pap58_1 (0,11) pap59_1 (0,14)</p> <p>pap28_4 (0,17) pap28_5 (0,18) pap28_1 (0,18)</p>	
	40	 <p>pap15_1 (0,00) pap15_4 (1,51) pap15_3 (6,89)</p> <p>pap15_2 (8,13) pap28_4 (14,6) pap14_6 (16,1)</p> <p>pap1_2 (18,8) pap57_1 (23,3) pap28_3 (24,1)</p>	 <p>pap15_1 (0,00) pap15_2 (0,19) pap15_3 (0,24)</p> <p>pap15_4 (0,42) pap1_2 (0,64) pap58_1 (0,64)</p> <p>pap59_1 (0,65) pap56_1 (0,76) pap50_1 (0,77)</p>	 <p>pap15_1 (0,00) pap15_2 (0,34) pap15_3 (0,45)</p> <p>pap15_4 (0,58) pap59_1 (0,96) pap1_2 (0,98)</p> <p>pap50_1 (1,02) pap56_1 (1,07) pap58_1 (1,07)</p>
		180	 <p>Pap15_1 (0,00) Pap15_4 (1,51) Pap15_3 (6,89)</p> <p>Pap15_2 (8,13) Pap28_4 (14,6) Pap14_6 (16,1)</p> <p>Pap1_2 (18,8) Pap57_1 (23,3) Pap28_3 (24,1)</p>	 <p>Pap15_1 (0,00) Pap15_2 (0,34) Pap15_3 (0,45)</p> <p>Pap15_4 (0,58) Pap59_1 (0,96) Pap1_2 (0,98)</p> <p>Pap50_1 (1,02) Pap56_1 (1,07) Pap58_1 (1,07)</p>

Figure 5.7 — Résultats de recherche de l'objet Pap15_1 avec environ 5, 40 et 180 descripteurs invariants. Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA.



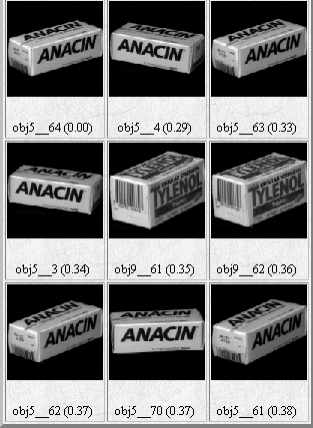





D	Inv. de Hu (10^{-12})	Inv. des mom. cplx. (10^{-6})	Inv. de la TFMA	
5	 <p>Obj5_64 (0,00) Obj9_22 (1,48) Obj6_6 (1,85)</p> <p>Obj6_3 (2,28) Obj10_4 (2,28) Obj10_40 (2,28)</p> <p>Obj10_41 (2,28) Obj10_42 (2,28) Obj10_43 (2,28)</p>	 <p>Obj5_64 (0,00) Obj5_27 (1,66) Obj9_57 (1,90)</p> <p>Obj9_63 (2,21) Obj9_22 (2,37) Obj9_21 (2,40)</p> <p>Obj9_58 (2,64) Obj6_21 (2,80) Obj9_27 (2,84)</p>	 <p>obj5_64 (0,00) obj5_4 (0,29) obj5_63 (0,33)</p> <p>obj5_3 (0,34) obj9_61 (0,53) obj9_62 (0,36)</p> <p>obj5_62 (0,37) obj5_70 (0,37) obj5_61 (0,38)</p>	
	40	 <p>Obj5_64 (0,00) Obj5_27 (3,06) Obj9_63 (4,49)</p> <p>Obj5_65 (5,09) Obj19_9 (5,29) Obj9_64 (5,51)</p> <p>Obj5_28 (5,70) Obj6_21 (5,71) Obj9_57 (5,80)</p>	 <p>Obj5_64 (0,00) Obj5_27 (3,06) Obj9_63 (4,49)</p> <p>Obj5_65 (5,09) Obj19_9 (5,29) Obj9_64 (5,51)</p> <p>Obj5_28 (5,70) Obj6_21 (5,71) Obj9_57 (5,80)</p>	 <p>obj5_64 (0,00) obj5_63 (0,72) obj5_3 (0,77)</p> <p>obj5_61 (0,78) obj5_34 (0,92) obj5_35 (0,93)</p> <p>obj5_33 (0,95) obj5_36 (0,95) obj5_60 (0,95)</p>
		180	 <p>Obj5_64 (0,00) Obj5_27 (3,06) Obj9_63 (4,49)</p> <p>Obj5_65 (5,09) Obj19_9 (5,29) Obj9_64 (5,51)</p> <p>Obj5_28 (5,70) Obj6_21 (5,71) Obj9_57 (5,80)</p>	 <p>Obj5_64 (0,00) Obj5_27 (3,06) Obj9_63 (4,49)</p> <p>Obj5_65 (5,09) Obj19_9 (5,29) Obj9_64 (5,51)</p> <p>Obj5_28 (5,70) Obj6_21 (5,71) Obj9_57 (5,80)</p>

Figure 5.8 — Résultats de recherche de l'objet Obj5_64 avec environ 5, 40 et 180 descripteurs invariants. Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA.




















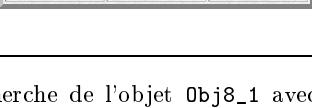
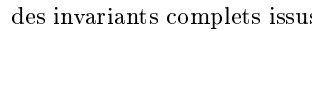


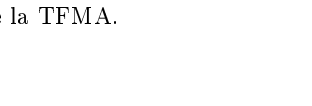

D	Inv. de Hu (10^{-15})	Inv. des mom. cplx. (10^{-6})	Inv. de la TFMA
5	 Obj8_1 (0,00) Obj16_3 (7,53) Obj16_30 (7,53)	 Obj8_1 (0,00) Obj8_2 (0,27) Obj9_38 (0,67)	 obj8_1 (0,00) obj8_0 (0,04) obj8_2 (0,04)
	 Obj16_31 (7,53) Obj16_32 (7,53) Obj16_33 (7,53)	 Obj9_34 (0,80) Obj8_70 (0,89) Obj4_38 (1,40)	 obj8_71 (0,07) obj8_70 (0,09) obj2_69 (0,10)
	 Obj16_34 (7,53) Obj16_35 (7,53) Obj16_36 (7,53)	 Obj8_3 (1,50) Obj5_35 (1,55) Obj8_69 (1,57)	 obj2_70 (0,10) obj8_3 (0,11) obj2_68 (0,13)
40	 Obj8_1 (0,00) Obj8_2 (0,60) Obj9_34 (2,31)	 Obj8_3 (2,60) Obj8_0 (2,70) Obj4_31 (3,16)	 obj8_1 (0,00) obj8_0 (0,13) obj8_2 (0,23)
	 Obj9_34 (3,27) Obj4_63 (3,37) Obj4_39 (3,45)	 obj8_71 (0,28) obj8_70 (0,34) obj8_69 (0,42)	
	 Obj8_1 (0,00) Obj8_2 (0,60) Obj9_34 (2,31)	 Obj8_3 (2,60) Obj8_0 (2,68) Obj4_31 (3,16)	 obj8_1 (0,00) obj8_0 (0,23) obj8_2 (0,35)
180	 Obj9_34 (3,27) Obj4_63 (3,37) Obj4_39 (3,45)	 obj8_71 (0,43) obj8_70 (0,49) obj8_69 (0,60)	
	 Obj8_1 (0,00) Obj8_2 (0,60) Obj9_34 (2,31)	 Obj8_3 (2,60) Obj8_0 (2,68) Obj4_31 (3,16)	 obj8_1 (0,00) obj8_0 (0,23) obj8_2 (0,35)
	 Obj9_34 (3,27) Obj4_63 (3,37) Obj4_39 (3,45)	 obj8_66 (0,52) obj8_68 (0,55) obj4_16 (0,57)	 obj8_66 (0,71) obj7_42 (0,75) obj8_68 (0,76)

Figure 5.9 — Résultats de recherche de l'objet Obj8_1 avec environ 5, 40 et 180 descripteurs invariants. Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA.

5.3 Conclusion du chapitre

Ce cinquième et dernier chapitre a présenté l'application des familles complètes de descripteurs invariants issus de la TFMA pour la recherche d'images en vue de l'indexation de bases d'objets par le contenu. Pour cela, nous avons développé une plateforme de test permettant de comparer les résultats de recherche en fonction de certains paramètres tels que le nombre de descripteurs invariants et la méthode d'approximation de la TFMA utilisée pour estimer ces descripteurs.

Les expériences conduites sur deux bases d'objets ont permis de tester et vérifier la pertinence de cet outil. Nous avons notamment comparé les résultats obtenus avec, d'une part, les invariants de Hu que l'on retrouve dans plusieurs moteurs de recherche, et d'autre part, avec une nouvelle famille complète de descripteurs invariants par similitudes planes issus des moments complexes de l'image. Cette famille a été construite en utilisant le procédé de normalisation introduit dans le 3^e chapitre.

Les deux familles complètes présentent de bien meilleurs résultats de recherche que les descripteurs invariants de Hu qui s'avèrent très peu robustes aux faibles distorsions de formes. La complétude est une propriété intéressante qui offre une représentation riche et flexible des formes pour la recherche d'images. Nous avons également constaté que la famille de descripteurs issue de la TFMA se montre plus robuste que les invariants des moments dans le cas de la base d'objets de Columbia.

Le système de recherche proposé ouvre sur plusieurs perspectives de développement permettant de compléter et enrichir la plateforme de test. Notamment, un point important concerne l'accélération de la recherche en évitant le parcourt exhaustif de tous les modèles de la base. Ces perspectives sont développées dans la conclusion générale de ce mémoire que nous abordons maintenant.

Conclusion générale

CETTE thèse a été consacrée à l'étude de la représentation globale des images planes à niveaux de gris dans le contexte de l'indexation de bases d'objets par le contenu. L'approche proposée dans ce mémoire est basée sur la description des objets par la transformée de Fourier-Mellin. Après avoir rappelé, au cours du premier chapitre, les principales méthodes de description globale des objets plan, nous avons introduit la transformée de Fourier-Mellin comme la transformée de Fourier sur les groupe des similitudes planes. Cela nous a permis de mettre en évidence que la TFM, dans sa définition standard, n'est pas adaptée pour les fonctions représentant des objets à niveaux de gris, ce qui a largement dévalorisé son utilisation en reconnaissance de formes par exemple.

Ainsi, le premier point que nous avons abordé concerne l'approximation de la TFM. Pour pallier au problème de divergence de l'intégrale de Mellin, le prolongement analytique de la transformée de Fourier-Mellin (TFMA) nous a permis de proposer trois algorithmes originaux pour l'estimation de la TFMA : l'approximation directe TFMA-D, l'approximation rapide TFMA-F et l'approximation cartésienne TFMA-C. Ces trois algorithmes diffèrent par la manière dont les images sont ré-échantillonnées sur de nouvelles grilles et par les méthodes d'intégration adéquates pour ces grilles. Par approximation de la transformée de Fourier-Mellin inverse, nous avons reconstruit des images à niveaux de gris complexes avec une qualité visuelle satisfaisante, ce qui n'avait pas été fait jusqu'à présent. Lorsque les spectres numériques sont de faibles dimensions, les résultats de reconstruction montrent cependant la présence de halos concentriques qui proviennent de la troncature abrupte des spectres numériques par une fenêtre rectangulaire. Il serait donc intéressant d'étendre les fenêtres développées pour la transformée de Fourier en utilisant le lien que cette dernière entretient avec la transformée de Fourier-Mellin par un changement de variable log-polaire.

Le second point que nous avons étudié concerne la description invariante des formes par similitudes. À l'aide de certaines propriétés héritées de la transformée de Fourier et adaptées au groupe des similitudes planes, nous avons caractérisé une forme comme l'ensemble des objets à niveaux de gris équivalents aux sens des similitudes. Ainsi, un procédé original de normalisation de la TFMA a été proposé pour construire des familles complètes et intégrables de descripteurs invariants par similitudes. La propriété de complétude garantie que chaque forme est représentée de manière unique et la convergence nous permet d'identifier l'espace des invariants à un espace métrique. Ainsi, une vraie distance est utilisée pour mesurer la similarité des objets indépendamment de leur pose, leur orientation et leur taille. Ces familles de descripteurs invariants reposent sur deux paramètres de normalisation que l'on peut choisir parmi un large éventail de possibilités. Une famille robuste a alors pu être proposée sur la base de critères numériques portant sur l'invariance effective aux similitudes, la discrimination des formes et la robustesse des descripteurs aux approximations et aux faibles distorsions de

formes.

Nous avons également développé une méthode d'appariement d'images permettant d'estimer le mouvement global entre deux objets de même forme à partir de leur TFMA. L'algorithme est construit à partir de la distance euclidienne exprimée sur l'espace transformé de Fourier-Mellin et l'estimation de mouvement est formulée comme un problème d'optimisation. Le minimum de la fonction d'appariement fournit également une mesure de non similarité entre formes permettant de déterminer si les paramètres estimés entre deux objets quelconques sont significatifs. Cette mesure de similarité est un écart entre formes pour le groupe des similitudes et la distance de Hausdorff dans le cas où le mouvement perçu des objets est restreint au groupe compact des rotations du plan. Une autre mesure qu'il serait intéressant d'étudier est la convolution sur le groupe des similitudes. Pour cela, nous pourrions utiliser l'expression de la convolution sur l'espace transformé de Fourier-Mellin. L'algorithme que nous avons développé prend en compte toute l'information disponible sur l'objets, contrairement aux techniques d'appariement basées sur des primitives préalablement extraites des images telles que les points d'intérêt, les segments, les coins... Les expériences montrent que la méthode est particulièrement robuste tant pour l'estimation des paramètres de mouvement que pour la reconnaissance et la discrimination des formes.

Nous avons finalement présenté l'application des descripteurs de formes issus de la TFMA en vue de l'indexation de bases d'objets par le contenu. Après avoir décrit le principe d'un système d'indexation et les méthodes de représentation des images des moteurs de recherche actuels, nous avons montré comment les familles d'invariants introduites dans le chapitre 3 pouvaient être utilisées pour rechercher les objets les plus similaires à une requête indépendamment de leur pose, leur orientation et leur taille. L'accent est mis sur les propriétés particulières de nos descripteurs, notamment l'unicité des invariants et la distance entre formes comme mesure de similarité. Nous avons réalisé des expériences à partir de deux bases d'objets et comparés les résultats obtenus avec deux ensembles de descripteurs invariants issus des moments géométriques de l'image. Nous avons d'abord mis en évidence que les résultats de recherche pouvaient être améliorés en augmentant le nombre de descripteurs invariants utilisés pour estimer la similarité entre les objets. Il est également apparu que l'approximation rapide de la TFMA (algorithme TFMA-F) conduisait à des résultats particulièrement robustes même lorsque le mouvement des objets 3D est tel que ses images laissent entrevoir de légères occlusions. Pour toutes ces raisons, nous pensons que les descripteurs sont une alternative intéressante aux descripteurs de formes généralement utilisés dans les moteurs de recherche actuels et s'inscrivent bien dans le cadre de travail de la future norme MPEG7.

Différentes perspectives sont envisageables pour compléter et enrichir le système de recherche proposé. L'indexation directe d'un objet à partir de ses invariants permet de compléter une base de modèles sans qu'il y ai lieu de modifier les descripteurs déjà stockés. Cependant, cette technique impose une confrontation exhaustive de l'objet requête avec tous les éléments de la base, ce qui peut engendrer des temps de recherche relativement importants. Une des voies les plus intéressantes à explorer concerne donc la phase d'indexation des descripteurs et l'accélération de la recherche. Pour éviter de scruter l'ensemble des modèles, nous avons vu qu'une solution consistait à construire une structure arborescente par quantification vectorielle des descripteurs invariants. L'identification précise de l'espace des formes nous permet d'en-

visager un apprentissage adapté au support des descripteurs invariants. Dans ce cas, il serait également intéressant de tester la robustesse des approximations de la TFMA aux erreurs de quantification, et notamment son influence sur la reconstruction des images. Une autre perspective importante à ce travail consiste à faire collaborer ces familles avec des descripteurs de formes contours ayant des propriétés similaires, ainsi qu'à intégrer des descripteurs locaux d'objets à niveaux de gris, mieux à même de décrire les déformations locales des formes ainsi que les occlusions.

A

Petite biographie sur J. Fourier et H. Mellin

A.1 Joseph Fourier

JEAN Batiste Joseph Fourier est né le 21 mars 1768 à Auxerre. Issu d'un milieu pauvre, Fourier demeure longtemps inconnu des milieux scientifiques parisiens. Lorsque, en 1789, il envoie à l'Académie un mémoire sur l'approximation des racines des équations polynômiales, la tourmente révolutionnaire voue à l'échec cette première tentative de faire connaître ses travaux. Fourier s'éloigne de sa ville natale pour suivre les cours de l'École Normale de l'an III. Il s'y fait remarquer de Monge, Laplace et Lagrange et, à la fondation de l'École Polytechnique, il est nommé assistant de Lagrange ; il lui succédera en 1797. Ses recherches suivent la variété de ses enseignements : des équations numériques à la mécanique rationnelle. En 1798, il accompagne le corps expéditionnaire français en Égypte et manifeste un vif intérêt pour les monuments antiques et leur interprétation, pour le calcul des erreurs d'observation et pour la statistique. En août 1799, il devient administrateur civil de l'Égypte.

De retour en France en 1802, il est nommé par Napoléon préfet à Grenoble. Bon diplomate, Fourier réussit à harmoniser les différentes tendances politiques en présence. Mais le préfet consciencieux reste géomètre. En 1804, il rédige une publication de ses travaux sur les équations numériques. Peu après, il édifie sa théorie de la propagation de la chaleur et en 1807, présente sur ce sujet un mémoire à l'Académie, mémoire que celle-ci couronnera en 1812. À la suite de son attitude ambiguë lors des événements de 1814-1815, il revient à Paris totalement démuné. Heureusement, au Bureau de la statistique du département de la Seine, il peut s'adonner plus librement à ses activités scientifiques. Ce n'est qu'alors qu'il commence à publier régulièrement : « théorie de la chaleur, équations numériques, statistique ». En avril 1816, le roi refuse d'entériner son élection comme académicien. Il acceptera pourtant, le 21 mai 1817, d'entériner une nouvelle élection, nommant ainsi Fourier membre de la commission de physique. Celui-ci fait alors partie de la communauté scientifique et, quelques mois après avoir publié son oeuvre majeure, *Théorie analytique de la chaleur* devient en 1822 secrétaire perpétuel de l'Académie des sciences. Il meurt, le 16 mai 1830, laissant près de soixante publications dont une, inachevée, l'*Analyse des équations déterminées*.



[extrait de l'*Encyclopédia Universalis*, Paris, 1989]

A.2 Hjalmar Mellin

ROBERT Hjalmar Mellin, fils d'ecclésiastique, est né à Liminka en Finlande en 1854. Il grandit à Hämeenlinna (environ 100 km au nord d'Helsinki) et suivit ses cours universitaires à Helsinki, où son professeur était le mathématicien suédois G. Mittag-Leffler. Pendant l'automne 1881, Mellin soutient sa thèse de doctorat sur les fonctions algébriques d'une variable complexe. Il séjourna par deux fois à Berlin, en 1881 et en 1882 pour étudier sous la direction de K. Weierstrass. En 1883-1884, il retourne à Stockholm pour continuer ses études avec Mittag-Leffler.

H. Mellin fût ensuite nommé enseignant à l'université de Stockholm de 1884 à 1891 mais n'y a jamais donné aucun cours magistral. En 1884, il fût nommé professeur en mathématiques au récent Institut Polytechnique qui devint, en 1908, l'Université Technique de Finlande. En 1901, Mellin abandonna son poste pour la chaire de mathématiques à l'Université d'Helsinki en faveur de son illustre ami et compatriote E. Lindelöf. Pendant la période 1904-1907, Mellin fût directeur de l'Institut Polytechnique et, en 1908, il devint le premier professeur de mathématiques dans la nouvelle université. Il y resta pendant 42 ans, se retirant en 1926 à l'âge de 72 ans. H. Mellin fût l'un des fondateurs de l'Académie des Sciences de Finlande en 1908. De 1908 jusqu'à sa mort en 1933, il représenta son pays dans le conseil éditorial de *Acta Mathematica*.



Les travaux de Mellin portent essentiellement dans le domaine de la théorie des fonctions. Il a étudié systématiquement la transformée qui porte actuellement son nom, qu'il a appliqué pour l'étude des fonctions Gamma, des fonctions hyper-géométriques, des séries de Dirichlet, ... Par extension à plusieurs dimensions, celle-ci lui a également permis de résoudre des équations aux dérivées partielles. La transformée inverse a été développée comme un outil puissant pour l'étude asymptotique de sommes numériques et harmoniques (cf. [46] pour une application récente de ce travail en informatique théorique). Il est intéressant de noter que, durant la dernière décennie de sa vie, Mellin, assez curieusement, fût préoccupé par la théorie de la relativité d'Einstein et écrivit une dizaine d'articles à ce sujet.

[Extrait traduit de l'article de R. Paris,
Université d'Abertay, Écosse, Décembre 1998]

Les représentations de Fourier et Mellin sont issues du site *Web The MathTutor History of Mathematics archive* (<http://www-history.mcs.st-and.ac.uk/~history/>).

B

Les moments géométriques et les invariants de Hu

La théorie des moments est une théorie classique en physique et en statistique. Les moments d'ordre faible, tels que le centre de masse et les moments d'inertie sont utilisés dans bien des disciplines. Historiquement, H. Ming-Kuel [86] fût le premier à les utiliser pour la reconnaissance de formes. Il décrivit comment les transformées telles que les translations, rotations, réflexions et facteurs d'échelle peuvent être évaluées avec un ensemble de moments géométriques et introduisit le concept des *invariants des moments*.

Le *moment géométrique* $m_f(p, q)$, d'ordre $(p + q)$, d'une fonction f continue par morceaux, est défini par :

$$m_f(p, q) = \iint_{\mathbb{R}^2} x^p y^q f(x, y) dx dy, \quad (\text{B.1})$$

où p et q désignent deux entiers naturels.

L'ensemble des moments $m_f(p, q)$ est déterminé de manière unique si l'intégrale de $f(x, y)$ sur \mathbb{R}^2 existe. Réciproquement, l'image $f(x, y)$ est uniquement déterminée par l'ensemble de ses moments $m_f(p, q)$. En principe, une image peut donc être reconstruite à partir de l'ensemble infini de ses moments. Cependant, il est bien connu que la reconstruction numérique d'une image depuis ses moments est particulièrement coûteuse en temps de calcul et est très instable [97]. De plus, les moments d'ordres élevés s'avèrent particulièrement sensibles aux bruits.

Le moment d'ordre 0 représente l'énergie totale de l'image et les moments d'ordre 1 permettent de calculer le barycentre d'une image :

$$(\bar{x}, \bar{y}) = \left(\frac{m_f(1, 0)}{m_f(0, 0)}, \frac{m_f(0, 1)}{m_f(0, 0)} \right).$$

On définit ainsi les *moments centrés* de l'image $\mu_f(p, q)$, invariants par translations :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad \mu_f(p, q) = \iint_{\mathbb{R}^2} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy. \quad (\text{B.2})$$

Si on considère maintenant deux images f et g représentant des objets identiques à un facteur d'échelle α près, *i.e.* $g(x, y) = f(\frac{x}{\alpha}, \frac{y}{\alpha})$, alors la relation entre les moments centrés des images s'écrit :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad \mu_g(p, q) = \alpha^\gamma \mu_f(p, q), \quad \text{avec } \gamma = p + q + 2.$$

Ce résultat est à l'origine des *moments réduits* $\eta_f(p, q)$, dont l'expression est donnée par :

$$\forall (p, q) \in \mathbb{N} \times \mathbb{N}, \quad \eta_f(p, q) = \mu_f(0, 0)^{-\frac{7}{2}} \mu_f(p, q). \quad (\text{B.3})$$

Les moments réduits sont invariants par translations et homothéties du plan. Pour obtenir l'invariance par rotation, H. Ming-Kuel [86] a proposé en 1963 un ensemble de 7 descripteurs appelés *invariants de Hu*. Ceux-ci sont construits comme des combinaisons non-linéaires des moments jusqu'à l'ordre 3 (la notation $\eta_f(p, q)$ est remplacée par η_{pq} pour des commodités d'écriture) :

$$\begin{aligned} \phi_f^1 &= \eta_{2,0} + \eta_{0,2} \\ \phi_f^2 &= (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2 \\ \phi_f^3 &= (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2 \\ \phi_f^4 &= (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2 \\ \phi_f^5 &= (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2 \right] \\ &\quad + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \left[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right] \\ \phi_f^6 &= (\eta_{2,0} - \eta_{0,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right] \\ &\quad + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{0,3} + \eta_{2,1}) \\ \phi_f^7 &= (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2 \right] \\ &\quad - (\eta_{3,0} - 3\eta_{2,1})(\eta_{1,2} + \eta_{0,3}) \left[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right] \end{aligned} \quad (\text{B.4})$$

Cette famille a été étendue à des invariants d'ordres supérieurs [14, 131] sans parvenir à une description invariante complète par rotation du plan. La principale raison est qu'il n'existe pas de relation simple liant les moments géométriques de deux objets de même forme mais orientés différemment. Pour cette raison, les moments géométriques sont généralement considérés comme inadéquats pour l'étude des rotations. De plus, ces descripteurs se sont souvent avérés décevants en reconnaissance de formes. Les principales sources d'erreurs par rapport aux propriétés théoriques d'invariance sont liées aux problèmes de discrétisation et d'intolérance aux bruits.

Les recherches se sont alors concentrées sur des représentations des images en coordonnées polaires plutôt qu'en coordonnées cartésiennes, puisqu'elles s'avèrent mieux adaptées à l'étude de l'invariance par similitude. L'alternative de Reddy [102] consiste à utiliser les *moments radiaux et angulaires* d'une image décrite en coordonnées polaires par rapport à son centre d'inertie (k, p, q désignent des entiers naturels) :

$$\psi_f(k, p, q) = \int_0^\infty \int_0^{2\pi} f(r, \theta) r^k \cos^p \theta \sin^q \theta d\theta dr.$$

Cette méthode a permis d'étendre l'ensemble des invariants de Hu à des ordres supérieurs. Cependant, elle reste essentiellement identique à la méthode précédente dans la mesure où il n'existe pas de relation simple liant les moments radiaux et angulaires de deux objets tournés l'un par rapport à l'autre.

ANNEXE

C

Bases d'objets utilisées pour les tests

C.1 Base de papillons

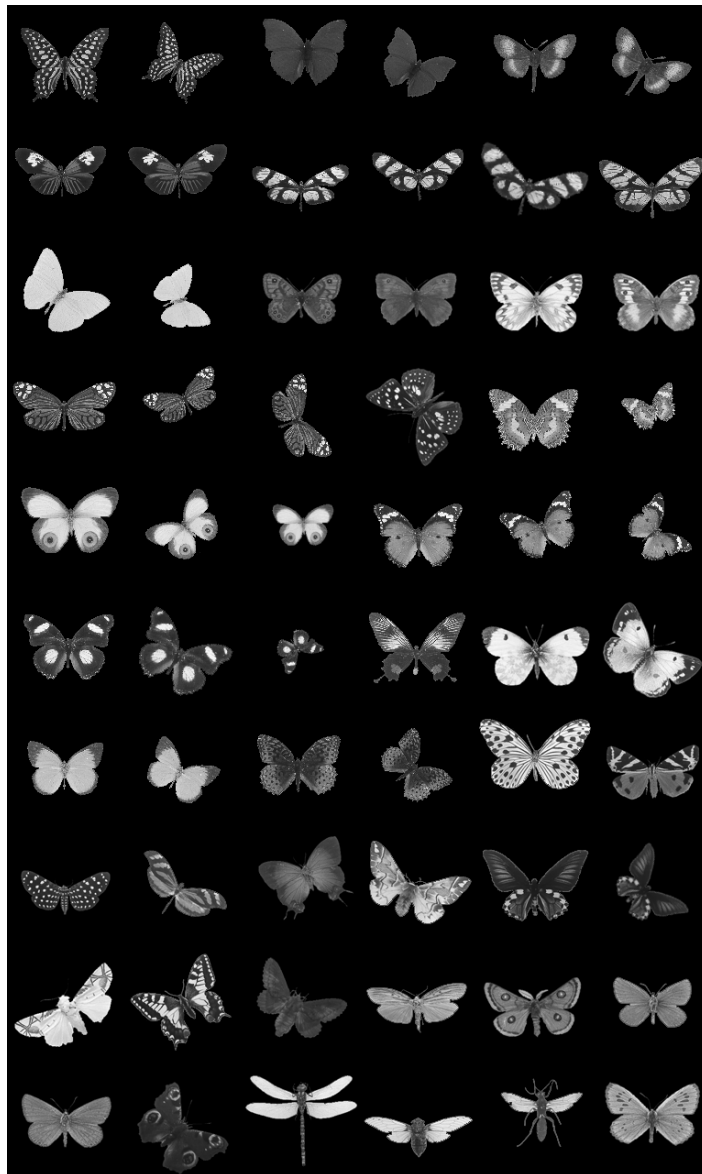


Figure C.1 — Échantillon de 60 images de papillons. La base complète contient 98 images au total, dont 69 représentent des papillons de formes différentes. La taille des images est de 128×128 pixels.

C.2 Bases d'objets de l'université de Columbia

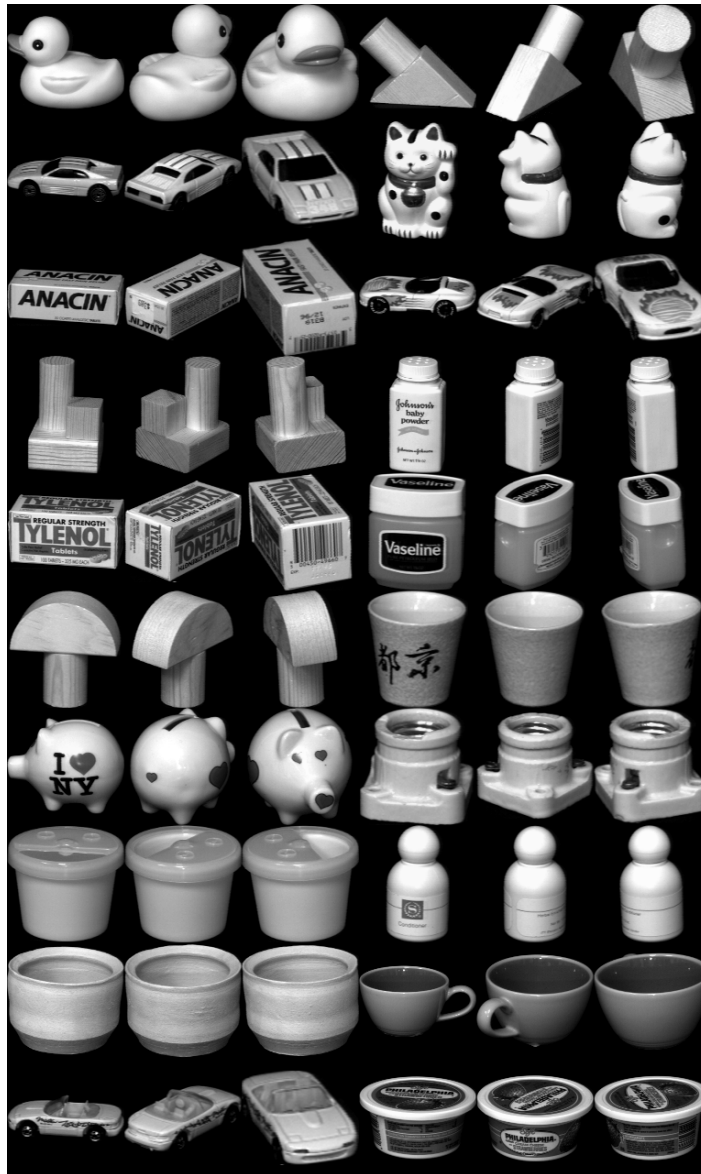


Figure C.2 — Échantillon de 60 images composé de 3 images de chacun des 20 objets de la base de Columbia. La base complète contient 1440 images des 20 objets pris dans 72 positions différentes. La taille des images est de 128×128 pixels.

<http://www.cs.columbia.edu/CAVE/coil-20.html>

Liste des figures

1.1	La transformée de Fourier des fonctions à support sur \mathbb{R} et illustration du théorème du retard.	7
1.2	La corrélation des fonctions à support sur \mathbb{R}	8
1.3	Action d'une similitude et d'une transformation affine sur la représentation d'un papillon et de son profil extérieur.	11
1.4	La représentation originale des signaux discrets, périodiques et causaux, et action des groupes de transformations sur leur transformée de Fourier respective.	12
1.5	Représentation des images par les fonctions circulaires harmoniques.	16
2.1	Représentation 2,5D de la surface d'un demi-disque en coordonnées cartésiennes et polaires.	26
2.2	Amplitude et phase de la TFM analytique d'un demi-disque	26
2.3	Représentation de la grille cartésienne et image du papillon servant à illustrer les trois méthodes d'approximation.	29
2.4	Illustration de l'algorithme directe d'approximation des TFMA directe et inverse.	30
2.5	Illustration de l'algorithme rapide d'approximation des TFMA directe et inverse.	33
2.6	Illustration de l'algorithme cartésien d'approximation des TFMA directe et inverse.	35
2.7	Images de test utilisées pour estimer l'erreur de reconstruction	36
2.8	Reconstruction de l'image de la lettre F en fonction de K et V (algorithme TFMA-C).	37
2.9	Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-D).	39
2.10	Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-F)	40

2.11	Reconstruction de l'image d'un papillon en fonction de K et V (algorithme TFMA-C).	41
2.12	Représentation de l'erreur de reconstruction pour l'image du papillon de la figure 2.7 en fonction des trois méthodes d'approximation.	42
3.1	Illustration de l'importance de la phase dans la représentation des images par la TFMA.	47
3.2	Caractérisation des formes dans le domaine de représentation original et dans le domaine transformé de Fourier-Mellin.	51
3.3	Module et phase de quelques descripteurs invariants de faibles ordres d'un demi-disque	57
3.4	Interprétation du rôle des paramètres de normalisation et caractérisation de l'objet prototype d'une forme.	58
3.5	Images de papillons utilisées pour tester la fiabilité des paramètres de normalisation.	60
3.6	Influence de l'erreur d'estimation des paramètres de normalisation sur le calcul de la distance invariante.	61
3.7	Reconstruction numérique des objets prototypes d'une image pour les familles initiales et étendues de descripteurs invariants.	63
3.8	Huit images de test utilisées pour tester la robustesse et la discrimination des familles de descripteurs invariants.	65
3.9	Évolution de la distance invariante entre des objets de même forme pour les trois méthodes d'approximation de la TFMA en fonction du nombre d'invariants	66
3.10	Évolution de la distance invariante entre des objets de formes différentes pour les trois méthodes d'approximation de la TFMA en fonction du nombre d'invariants	67
3.11	Évaluation de la discrimination numérique des familles initiales et étendues de descripteurs invariants pour les trois méthodes d'approximation de la TFMA.	68
3.12	Reconstruction des objets prototypes de deux papillons de formes très différentes pour des ensembles discrets et finis de descripteurs invariants issus de la famille étendue.	69
3.13	Évaluation de la discrimination numérique des familles initiales et étendues entre deux papillons de formes proches mais différentes (étirement horizontal).	70
3.14	Reconstruction des objets prototypes de deux papillons de formes proches mais différentes (étirement horizontal) pour des ensembles discrets et finis de descripteurs invariants issus de la famille étendue.	70
3.15	Évaluation de la discrimination numérique des familles initiales et étendues entre deux papillons de formes proches mais différentes (faibles variations locales de texture).	72
3.16	Reconstruction des objets prototypes de deux papillons de formes proches mais différentes (variations naturelles de la texture des ailes) pour des ensembles discrets et finis de descripteurs invariants issus de la famille étendue.	72

4.1	Fonction d'appariement entre deux images de demi-disques et estimation des paramètres de mouvement	79
4.2	Illustration de la fonction d'appariement entre des papillons similaires et des papillons de formes différentes.	81
4.3	Chemin parcouru par l'algorithme d'optimisation à la surface de la fonction d'appariement calculée pour deux objets de même forme. Zoom sur le puits de la figure 4.2.	83
4.4	Transformée de Fourier-Mellin analytique et fonction d'appariement d'objets possédant des symétries par rotation.	87
4.5	Dérivée de la fonction d'appariement calculée sur une image présentant des axes de symétrie par rotation.	89
4.6	Huit images de test utilisées pour tester la robustesse et la discrimination de l'écart entre formes.	90
4.7	Évolution de l'écart et des paramètres de mouvement estimés entre des objets de même forme pour les trois méthodes d'approximation de la TFMA.	91
4.8	Évaluation de la discrimination de l'écart pour les trois méthodes d'approximation de la TFMA.	92
4.9	Évolution de l'écart et du ratio de discrimination entre des objets de formes presque identiques pour les trois méthodes d'approximation de la TFMA.	93
5.1	Architecture générique d'un système d'indexation de bases d'images.	97
5.2	Interface utilisateur pour l'interrogation des bases d'objets.	102
5.3	Principe de fonctionnement de la plateforme de test.	102
5.4	Résultats de recherche sur la base de papillons — Comparaison des trois méthodes d'approximation de la TFMA.	104
5.5	Résultats de recherche sur la base d'objets Columbia — Comparaison des trois méthodes d'approximation de la TFMA.	105
5.6	Résultats de recherche sur la base d'objets Columbia (exemple 1) — Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA	107
5.7	Résultats de recherche sur la base d'objets Columbia (exemple 2) — Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA	109
5.8	Résultats de recherche sur la base d'objets Columbia — Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA	110
5.9	Résultats de recherche sur la base d'objets Columbia — Comparaison des invariants de Hu, des invariants complets issus des moments complexes et des invariants complets issus de la TFMA	111
C.1	Échantillon de 60 images provenant de la base de papillons.	121

C.2 Échantillon de 60 images provenant de la base de Columbia.	122
--	-----

Liste des tableaux

4.1	Itérations de l'algorithme d'optimisation pour l'estimation des paramètres de mouvement entre des objets de même forme	82
4.2	Quelques itérations de l'algorithme d'optimisation entre des objets de formes différentes	83

Bibliographie

- [1] *Special issue on content-based image retrieval systems*, IEEE Computers Magazine, vol. 28, no. 9, 1995. Editeurs : V.N. Guidivada et J.V. Raghavan.
- [2] MPEG7 : *context and objectives (version 5.0)*, Rapport technique ISO/IEC JTC1/SC29/WG11 N1920, MPEG, octobre 1997.
- [3] MPEG7 : *proposal package description (PPD - version 1.0)*, Rapport technique ISO/IEC JTC1/SC29/WG11 N1923, MPEG, octobre 1997.
- [4] *Special issue on visual information management*, Communication of the ACM, décembre 1997. Editeur : R. Jain.
- [5] Y. ABU-MOSTAFA et D. PSALTIS, *Recognitive aspects of moment invariants*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 6, no. 6, pp. 698–706, novembre 1984.
- [6] Y. ABU-MOSTAFA et D. PSALTIS, *Image normalisation by complex moments*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 7, no. 1, pp. 46–55, janvier 1985.
- [7] P. AIGRAIN et AL, *Content-based representation and retrieval of visual media : a state-of-the-art review*, Multimedia Tools and Applications, vol. 3, no. 3, novembre 1996.
- [8] R. ALTES, *The Fourier-Mellin transform and mammalian hearing*, J. of the Acoustical Society of America, vol. 63, pp. 174–183, janvier 1978.
- [9] K. ARBTER, *From pixels to features*, chapitre Affine-invariant Fourier descriptors, (pp. 153–164), Elsevier Science, Amsterdam (NL), J.C. Simons édition, 1989.
- [10] K. ARBTER, W. SNYDER, H. BURKHARDT et G. HIRZINGER, *Application of affine-invariant Fourier descriptors to recognition of 3-D objects*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 640–647, juillet 1990.
- [11] H. ARSENAULT et Y. SHENG, *Properties of the circular harmonic expansion for rotation-invariant pattern recognition*, Applied optics, vol. 25, no. 18, pp. 3225–3229, septembre 1986.
- [12] C. ASMUTH, *An application of group representation theory to picture recognition*, Comput. Math. Applic., vol. 13, no. 4, pp. 363–365, 1987.

-
- [13] J. BACH, *The Virage image search engine : An open framework for image management*, dans *Storage and retrieval for still image and video databases IV, IS&T/SPIE*, vol. 2670, (pp. 76–87), 1996.
- [14] I. BALSLEV, *Noise tolerance of moments invariant in pattern recognition*, *Pattern Recognition Letters*, vol. 19, pp. 1183–1189, 1998.
- [15] T. BANNOUR et F. GHORBEL, *Isotropie de la représentation des surfaces ; Application à la description et la visualisation d'objets 3D*, dans *A paraître dans 12^e colloque francophone en reconnaissance des formes et intelligence artificielle*, Paris (Fr.), 1-3 février 2000.
- [16] J. BEIS et D. LOWE, *Shape indexing using approximate nearest-neighbor search in high-dimensional spaces*, dans *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, (pp. 1000–1006), Puerto Rico, juin 1997.
- [17] S. BELKASIM, M. SHRIDHAR et M. AHMADI, *Pattern recognition with moment invariants : A comparative study and new results*, *Pattern Recognition*, vol. 12, pp. 1117–1138, 1991.
- [18] L. BEN YOUSSEF et F. GHORBEL, *Matching planar and gray-level shapes using the generalized correlation*, dans *soumis à la 15^e Conf. on Pattern Recognition*, Barcelone (Esp.), 3-8 septembre 2000.
- [19] A. BLAKE, M. TAYLOR et A. COX, *Grasping visual symmetry*, dans *Proc. of the Int. Conf. on Computer Vision*, (pp. 724–733), Berlin (All.), 11-14 mai 1993.
- [20] R. BRANDT et F. LIN, *Representations that uniquely characterize images modulo translation, rotation, and scaling*, *Pattern Recognition Letters*, vol. 17, pp. 1001–1015, 1996.
- [21] R. BRENT, *Algorithms for minimization without derivatives*, Englewood cliffs, NJ : Prentice-Hall, 1973.
- [22] C. CANN, S. DERRODE, F. GHORBEL et V. BURDIN, *A model of open surface based on Fourier-Mellin invariants obtained from a depth image*, dans *Third Int. Conf. on Curves and Surfaces*, Chamonix (Fr.), 27 juin - 3 juillet 1996.
- [23] D. CASASENT et D. PSALTIS, *Scale invariant optical transform*, *Optical Engineering*, vol. 15, no. 3, pp. 258–261, mai/juin 1976.
- [24] S. CHANG et J. SMITH, *Extracting multi-dimentionnal signal features for content-based visual query*, dans *Proc. SPIE Symposium on Visual Communications and Signal Processing*, 1995.
- [25] J. CHASSERY et A. MONTANVER, *Géométrie discrète en analyse d'images*, Édition Hermès (Fr.), 1991.
- [26] Q. CHEN, M. DEFRISE et F. DECONINCK, *Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, 1994.
- [27] D. CLEMENS et D. JACOBS, *Model group indexing for recognition*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 1007–1017, octobre 1991.
- [28] I. COX, M. MILLER, S. HOMOHUNDRO et P. YIANILOS, *Pichunter : Bayesian relevance feedback for image retrieval*, dans *Proc. of the IEEE Int. Conf. on Pattern Recognition*, (pp. 361–369), 1998.

- [29] T. CRIMMINS, *A complete set of Fourier descriptors for two-dimensional shapes*, IEEE trans. on Systems, Man, and Cybernetics, vol. 12, no. 6, pp. 848–855, novembre/décembre 1982.
- [30] M. DAOUDI, F. GHORBEL, A. MOKADEM, O. AVARO et H. SANSON, *Shape distances for contour tracking and motion estimation*, Pattern Recognition, vol. 32, pp. 1297–1306, juin 1999.
- [31] P. DAVIS, *Plane regions determined by complex moments*, Journal of Approximation Theory, vol. 19, pp. 148–153, 1977.
- [32] J. DE BOUGRENET DE LA TOCNAYE et F. GHORBEL, *Scale-rotation invariant pattern recognition applied to image data compression*, Pattern Recognition Letters, vol. 8, pp. 55–58, juillet 1988.
- [33] L. DENIAU et J. BLANC-TALON, *A new nonlinear operator for multiresolution signal and image analysis*, dans *4th International Symposium on Signal Processing and its Applications*, Gold Coast (Au.), 25-28 août 1996.
- [34] S. DERRODE, M. DAOUDI et F. GHORBEL, *Invariant content-based image retrieval using a complete set of Fourier-Mellin descriptors*, dans *First Int. Conf. on Multimedia Computing and Systems*, (pp. 877–881), Florence (It.), 7-11 juin 1999.
- [35] S. DERRODE et F. GHORBEL, *Estimation de mouvement et reconstruction d'objets à niveaux de gris décrits par leur transformée de Fourier-Mellin*, dans *3^e journées d'études et d'échanges CORESA*, Issy-Les-Moulineaux (Fr.), 26-27 mars 1997.
- [36] S. DERRODE et F. GHORBEL, *Rapport final de la convention CNET/ENIC*, Rapport technique Contrat 94-7407, École Nouvelle d'Ingénieurs en Communication, septembre 1997.
- [37] S. DERRODE et F. GHORBEL, *Extraction de familles complètes de descripteurs invariants par similitudes pour la reconnaissance d'objets à niveaux de gris*, dans *1^{er} atelier de Traitement et Analyse d'Images : Méthodes et Applications (TAIMA'99)*, Hammamet (Tn.), 20-26 mars 1999.
- [38] S. DERRODE et F. GHORBEL, *Comparaison de deux familles complètes de descripteurs de formes pour l'indexation de bases d'objets 2D à niveaux de gris*, à paraître dans *Annales des Télécom*, 2000.
- [39] S. DERRODE, F. GHORBEL, O. AVARO et H. SANSON, *Transformée de Fourier-Mellin numérique - Reconstruction et estimation de mouvement d'objets à niveaux de gris*, dans *16^e Colloque GRETSI*, (pp. 655–658), Grenoble (Fr.), 15-19 septembre 1997.
- [40] S. DERRODE, R. MEZHOUD et F. GHORBEL, *Reconnaissance de formes par invariants complets et convergents - Application à l'indexation de bases d'objets à niveaux de gris*, dans *17^e Colloque GRETSI*, (pp. 119–122), Vannes (Fr.), 13-17 septembre 1999.
- [41] J. DIEUDONNÉ, *Éléments d'analyse*, Cahiers scientifiques, Vol. 2, 3^e édition, Gauthier-Villars, Paris (Fr.), 1983.
- [42] J. DOW, *Content-based retrieval in multimedia imaging*, dans *Proc. of the SPIE Conf. on Visual Communication and Image Processing*, San Jose (Ca.), février 1993.
- [43] P. FAUDEMAY et Y. RUI, *Proposal to the MPEG7 proposal package description (ppd) document*, Rapport technique ISO/IEC JTC1/SC29/WG11 M3039, MPEG, 1998.

-
- [44] O. FAUGERAS et M. HEBERT, *The representation, recognition and positioning of 3D shapes from range data*, Techniques for 3D machine perception, Elsevier Science Publishers (B.V.), North Holland, 1986.
- [45] M. FERRARO, *Invariant pattern representations and Lie groups theory*, Advances in Electronics and Electron Physics, vol. Chapter 3, pp. 131–196, 1992.
- [46] P. FLAJOLET, X. GOURDON et P. DUMAS, *Mellin transforms and asymptotics : Harmonic sums*, Theoretical Computer Science, vol. 144, no. 1-2, pp. 3–58, juin 1994.
- [47] D. FLEET et A. JEPSON, *Stability of phase information*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 12, pp. 1253–1268, décembre 1993.
- [48] D. FLEET et A. JEPSON, *Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 16, no. 12, pp. 1156–1168, décembre 1994.
- [49] M. FLICKNER et AL, *Query by image and video content : the QBIC system*, IEEE trans. on Computers, vol. 28, no. 9, pp. 23–32, 1995.
- [50] H. FONGA, *Pattern recognition in gray-level images by Fourier analysis*, Pattern Recognition Letters, vol. 17, pp. 1477–1489, 1996.
- [51] H. FREEMAN, *Shape description via the use of critical points*, Pattern Recognition, vol. 10, pp. 159–166, 1978.
- [52] J. GAUTHIER, G. BORNARD et M. SILBERMANN, *Motion and pattern analysis : Harmonic analysis on motion groups and their homogeneous spaces*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 21, no. 1, pp. 159–171, janvier/février 1991.
- [53] F. GHORBEL, *Vers une approche mathématique unifiée des aspects géométriques et statistiques de la reconnaissance de formes planes*, Thèse de doctorat, Université de Rennes I, 1990.
- [54] F. GHORBEL, *Stability of invariant Fourier descriptors and its inference in the shape classification*, dans *Proc. of the IEEE Int. Conf. on Pattern Recognition*, (pp. 130–134), The Hague (NL), 2-9 août 1992.
- [55] F. GHORBEL, *A complete invariant description for gray-level images by the harmonic analysis approach*, Pattern Recognition Letters, vol. 15, pp. 1043–1051, octobre 1994.
- [56] F. GHORBEL, *Towards a unitary formulation for invariant image description ; Application to image coding*, Annales des Télécommunications, vol. 53, no. 3, pp. 143–163, mai/juin 1998.
- [57] F. GHORBEL et J. DE BOUGRENET DE LA TOCNAYE, *Automatic control of lamelli-branch larva growth using contour invariant feature extraction*, Pattern Recognition, vol. 23, no. 3/4, pp. 319–323, mars 1990.
- [58] S. GOH, *The Mellin transformation : Theory and digital filter implementation*, Ph.D. dissertation, Purdue University, West Lafayette (In.), 1985.
- [59] A. GRACE et M. SPANN, *A comparison between Fourier-Mellin descriptors and moment based features for invariant object recognition using neural networks*, Pattern Recognition Letters, vol. 12, pp. 635–643, octobre 1991.
- [60] G. GRANLUND, *Fourier preprocessing for hand printed character recognition*, IEEE trans. on Computers, vol. 21, pp. 195–201, février 1972.

-
- [61] P. GROS, O. BOURNEZ et E. BOYER, *Using geometric quasi-invariants to match and model images of line segments*, Rapport de recherche 2606, INRIA, juillet 1995.
- [62] A. GUPTA et R. JAIN, *Visual information retrieval*, Communications of the ACM, vol. 40, no. 5, pp. 70–79, 1997.
- [63] R. HARALICK, K. SHANMUGAM et I. DINSTEN, *Texture features for image classification*, IEEE trans. on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610–621, 1973.
- [64] D. HUTTENLOCHER et K. KEDEM, *Computing the minimum Hausdorff distance for point sets under translation*, Proc of the ACM Symp. on Computational Geometry, (pp. 340–349), 1990.
- [65] D. HUTTENLOCHER, G. KLANDERMAN et W. RUCKLIDGE, *Comparing images using the Hausdorff distance*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 9, pp. 850–863, septembre 1993.
- [66] D. HUTTENLOCHER et W. RUCKLIDGE, *A multi-resolution technique for comparing images using the Hausdorff distance*, Internal report TR 92-1321, Cornell University, décembre 1992.
- [67] D. HUTTENLOCHER et S. ULLMAN, *Recognizing solid objects by alignment with an image*, Int. J. of Computer Vision, vol. 5, no. 2, pp. 195–212, 1990.
- [68] M. INGGIS et A. ROBINSON, *Neural approaches to ship target recognition*, dans *Proc. of the IEEE Int. Radar Conference*, (pp. 386–391), Alexandria, 8-11 mai 1995.
- [69] T. IRINO, *A Gammachirp function as an optimal auditory filter with the Mellin transform*, dans *Proc. of the IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, (pp. 981–984), Atlanta, 7-10 mai 1996.
- [70] K. KANATANI, *Group theoretical methods in image understanding*, Lecture notes in computer science, Springer-Verlag, Berlin, Germany, 1990.
- [71] A. KHOTANZAD et Y. HONG, *Invariant image recognition by Zernike moments*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 489–497, mai 1990.
- [72] A. KHOTANZAD et J. LU, *Classification of invariant image representations using a neural network*, IEEE trans. on Acoustics, Speech and Signal Processing, vol. 38, no. 6, pp. 1028–1190, juin 1990.
- [73] Y. LAMDAN et H. WOLFSON, *Geometric hashing : a general and efficient model-based recognition scheme*, dans *Proc. of the Int. Conf. on Computer Vision*, (pp. 238–249), Tampa (Fl.), décembre 1988.
- [74] R. LENZ, *Describing and recognizing group-invariant pattern classes with group-sampling*, Pattern Recognition Letters, vol. 9, pp. 169–173, avril 1989.
- [75] R. LENZ, *Group theoretical methods in image processing*, vol. 413 de *Lecture notes in computer science*, Springer-Verlag, Berlin, Germany, G. Goos and J. Hartmanis édition, 1990.
- [76] Y. LI, *Reforming the theory of invariant moments for pattern recognition*, Pattern Recognition, vol. 25, no. 7, pp. 723–730, 1992.
- [77] S. LIAO et M. PAWLAK, *On image analysis by moments*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 18, no. 3, pp. 254–266, mars 1996.

- [78] S. LIAO et M. PAWLAK, *On the accuracy of Zernike moments for image analysis*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 20, no. 12, pp. 1358–1364, décembre 1998.
- [79] F. LIN et R. BRANDT, *Towards absolute invariants of images under translation, rotation, and dilatation*, Pattern Recognition Letters, vol. 14, pp. 369–379, mai 1993.
- [80] A. MATHENY et D. GOLDOF, *The use of three- and four-dimensional surface harmonics for rigid and non-rigid surface recovery and representation*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 10, pp. 967–981, octobre 1995.
- [81] S. MATUSIAK, *Description invariante et locale des formes planes – Application à l’indexation d’une base d’images*, Thèse de doctorat, Université de Valenciennes et du Hainaut-Cambresis, 1999.
- [82] R. MEZHOUD, *Description invariante et complète des images à niveaux de gris obtenue par les moments géométriques*, Rapport de dea, École Nationale des Sciences de l’Informatique, Rue des entrepreneurs, La Charguia II, Tunis-Carthage, Tunisie, Juin 1998.
- [83] R. MEZHOUD, T. BANNOUR, S. DERRODE et F. GHORBEL, *Invariants des moments et applications*, dans *1^{er} atelier de Traitement et Analyse d’Images : Méthodes et Applications (TAIMA’99)*, Hammamet (Tn.), 20-26 mars 1999.
- [84] R. MEZHOUD, S. DERRODE et F. GHORBEL, *Comparaison de deux familles de descripteurs de formes pour l’indexation de bases d’images 2D*, dans *5^e journées d’études et d’échanges CORESA*, Sophia-Antipolis (Fr.), 14-15 juin 1999.
- [85] R. MILANESE et M. CHERBULIEZ, *A rotation-, translation-, scale-invariant approach to content-based image retrieval*, J. of Visual Communication and Image Representation, vol. 10, pp. 186–196, 1999.
- [86] H. MING-KUEL, *Visual pattern recognition by moment invariants*, IRE trans. on Information Theory, vol. 8, pp. 179–187, 1962.
- [87] T. MINKA et R. PICARD, *Interactive learning using a society of models*, Pattern Recognition, vol. 30, no. 4, pp. 565–581, 1997.
- [88] A. MOKADEM, *Distances entre formes : Application au codage orienté objet*, Thèse de doctorat, Université de Compiègne, 1996.
- [89] J. MUNDY, A. ZISSERMAN et D. FORSYTH. (ED.), *Applications of invariance in computer vision*, Lecture notes in computer science, Springer-Verlag, Berlin, Germany, 1993.
- [90] H. MURASE et S. NAYAR, *Visual learning and recognition of 3d objects from appearance*, Int. J. of Computer Vision, vol. 14, pp. 5–24, 1995.
- [91] C. NASTAR, N. BOUJEMAA, M. MITSCHKE et C. MEILHAC, *Surfimage : un système flexible d’indexation et de recherche d’images*, dans *4^e journées d’études et d’échanges CORESA*, Lannion (Fr.), 9-10 juin 1998.
- [92] W. NIBLACK et AL, *The QBIC project : Querying images by content using color, texture and shape*, dans *Proc. of the SPIE Conf. on Visual Communication and Image Processing*, (pp. 173–181), San Jose (Ca.), février 1993.
- [93] P. OHANIAN et R. DUBES, *Performance evaluation for four classes of texture features*, Pattern Recognition, vol. 25, no. 8, pp. 819–833, 1992.

-
- [94] C. OLSON, *A probabilistic formulation for Hausdorff matching*, dans *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, (pp. 150–156), Santa Barbara (Ca.), juin 1998.
- [95] A. OPENHEIM et J. LIM, *The importance of phase in signals*, *Proceedings of the IEEE*, vol. 69, pp. 529–541, 1981.
- [96] M. PAVEL, *Fundamentals of pattern recognition*, Pure and applied mathematics, Marcel Dekker, Inc., New York, 1989.
- [97] M. PAWLAK, *On the reconstruction aspects of moment descriptors*, *IRE trans. on Information Theory*, vol. 38, no. 6, pp. 1698–1708, novembre 1992.
- [98] A. PENTLAND, R. PICARD et S. SCLAROFF, *Photobook : Tools for content-based manipulation of image databases*, *Int. J. of Computer Vision*, vol. 18, no. 3, pp. 233–254, 1996.
- [99] E. PERSOON et K. FU, *Shape description using Fourier descriptors*, *IEEE trans. on Systems, Man, and Cybernetics*, vol. 7, no. 3, pp. 170–179, mars 1977.
- [100] W. PRESS, S. TEUKOLSKY, A. VETTERLING et B. FLANNERY, *Numerical recipes in C, The art of scientific computing*, Cambridge university press, 2^e édition, 1983.
- [101] R. PROKOP et A. REEVES, *A survey of moment-based techniques for unoccluded object representation and recognition*, *Graphical Models and Image Processing*, vol. 54, no. 5, pp. 438–460, septembre 1992.
- [102] S. REDDI, *Radial and angular moment invariants for image identification*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 3, no. 2, pp. 240–242, mars 1981.
- [103] T. REISS, *Recognising planar objects using invariant image features*, vol. 676 de *Lecture notes in computer science*, Springer-Verlag, Berlin, Germany, 1993.
- [104] I. ROTHE, K. VOSS et H. SUESSE, *The method of normalization to determine invariants*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 4, pp. 366–375, avril 1996.
- [105] W. RUCKLIDGE, *Efficient visual recognition using the Hausdorff distance*, vol. 1173 de *Lecture notes in computer science*, Springer-Verlag, Berlin, Germany, 1996.
- [106] W. RUDIN, *Fourier analysis on groups*, Wiley Classics, New York, 1990.
- [107] Y. RUI, T. HUANG et S. CHANG, *Image retrieval : current techniques, promising directions, and open issues*, *J. of Visual Communication and Image Representation*, vol. 10, pp. 39–62, 1999.
- [108] Y. RUI, T. HUANG, M. ORTEGA et S. MEHROTRA, *Relevance feedback : a power tool in interactive content-based image retrieval*, *IEEE trans. on Circuits and Systems for Video Technology*, special issue on Interactive Multimedia Systems for the Internet, vol. 8, no. 5, pp. 644–655, septembre 1998.
- [109] F. SADJADI et E. HALL, *Three dimensional moments invariants*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 2, no. 2, pp. 127–136, mars 1980.
- [110] C. SCHMID, *Appariement d'images par invariants locaux à niveaux de gris*, Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [111] C. SCHMID et R. MOHR, *Local greyvalue invariants for image retrieval*, *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–535, mai 1997.

-
- [112] H. SCHULZ-MIRBACH, *Constructing invariant features by averaging techniques*, dans *Proc. of the IEEE Int. Conf. on Pattern Recognition*, (pp. 387–390), Jerusalem, Israel, 25-28 octobre 1994.
- [113] D. SHEN et H. IP, *Symmetry detection by generalized complex moments : a closed form solution*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 21, no. 5, pp. 466–476, mai 1999.
- [114] Y. SHENG et H. ARSENAULT, *Experiments on pattern recognition using invariant Fourier-Mellin descriptors*, J. of the Optical Society of America, vol. 3, no. 6, pp. 771–776, juin 1986.
- [115] Y. SHENG et J. DUVERNOY, *Circular-Fourier radial-Mellin transform descriptors for pattern recognition*, J. of the Optical Society of America, vol. 3, no. 6, pp. 885–888, juin 1986.
- [116] Y. SHENG et C. LEJEUNE, *Invariant pattern recognition using Fourier-Mellin transforms and neural networks*, J. of Optics, vol. 22, no. 5, pp. 223–228, 1991.
- [117] Y. SHENG, C. LEJEUNE et H. ARSENAULT, *Frequency-domain Fourier-Mellin descriptors for invariant pattern recognition*, Optical Engineering, vol. 27, no. 8, pp. 354–357, mai 1988.
- [118] Y. SHENG et L. SHEN, *Orthogonal Fourier-Mellin moments for invariant pattern recognition*, J. of the Optical Society of America, vol. 11, no. 6, pp. 1748–1757, juin 1994.
- [119] M. STRICKER et M. ORENGO, *Similarity of color images*, dans *Proc. of the SPIE Conf. on Visual Communication and Image Processing*, vol. 2420, (pp. 381–392), 1995.
- [120] M. SWAIN et D. BALLARD, *Color indexing*, Int. J. of Computer Vision, vol. 7, no. 1, pp. 11–32, 1991.
- [121] D. SWEENEY, E. OCHOA et G. SCHILS, *Experimental use of iteratively designed rotation-invariant correlation filters*, Applied optics, vol. 26, no. 16, pp. 3458–3465, août 1987.
- [122] H. TAMURA, S. MORI et T. YAMAWAKI, *Texture features corresponding to visual perception*, IEEE trans. on Systems, Man, and Cybernetics, vol. 8, no. 6, pp. 460–473, 1978.
- [123] M. TEAGUE, *Image analysis via the general theory of moments*, J. of the Optical Society of America, vol. 70, no. 8, pp. 920–930, août 1980.
- [124] C. TEH et R. CHIN, *On image analysis by the method of moments*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 10, no. 4, pp. 496–513, juillet 1988.
- [125] T. TZANETAKIS, M. TRAKA et G. TZIRITAS, *Motion estimation based on affine moments invariants*, dans *IX European Signal Processing Conf., Eupisco'98*, (pp. 925–929), Island of Rhodes, (Gr.), 8-11 septembre 1998.
- [126] A. WALLIN et O. KÜBLER, *Complete sets of complex Zernike invariants and the role of the pseudo-invariants*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 11, pp. 1106–1110, novembre 1995.
- [127] H. WECHSLER et G. ZIMMERMAN, *2D invariant object recognition using distributive associative memory*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 10, no. 6, pp. 811–821, novembre 1988.
- [128] I. WEISS, *Projective invariants of shapes*, dans *Proc. DARPA, Image understanding workshop*, (pp. 1125–1134), 1989.

-
- [129] I. WEISS, *Geometric invariants and object recognition*, Int. J. of Computer Vision, vol. 10, pp. 207–231, 1993.
- [130] R. WU et H. STARK, *Rotation-invariant pattern using a vector reference*, Applied optics, vol. 23, no. 6, pp. 838–840, mars 1984.
- [131] R. WU et H. STARK, *Rotation-invariant pattern recognition using optimum feature extraction*, Applied optics, vol. 24, no. 2, pp. 179–184, janvier 1985.
- [132] L. YANG, F. ALBREGTSEN et T. TAXT, *Fast computation of three dimensional geometric moments using a discrete divergence theorem and a generalization to higher dimensions*, Graphical Models and Image Processing, vol. 59, no. 2, pp. 97–108, 1997.
- [133] X. YI et O. CAMPS, *Line feature-based recognition using Hausdorff distance*, dans *Proc. of the Int. Symposium on Computer Vision*, (pp. 79–84), Coral Gables (Fl.), novembre 1995.
- [134] E. ZALUBAS et W. WILLIAMS, *Discrete scale transform for signal analysis*, dans *Proc. of the IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, (pp. 1557–1560), Detroit, 9-12 mai 1995.
- [135] C. ZHAN et R. ROSKIES, *Fourier descriptors for plane closed curves*, IEEE trans. on Computers, vol. 21, no. 3, pp. 269–281, 1972.
- [136] M. ZRIBI, *Les fonctions spéciales et les représentations des groupes pour la reconnaissance de formes – Application à l’imagerie médicale*, Thèse de doctorat, Université de Rennes I, 1997.
- [137] P. ZWICKE et Z. KISS, *A new implementation of the Mellin transform and its application to radar classification*, IEEE trans. on Pattern Analysis and Machine Intelligence, vol. 5, no. 2, pp. 191–199, mars 1983.

