

# THESE

*Préparée*

à l'École Supérieure des Communications de Tunis  
Université de Carthage

Unité de Recherche en Imagerie Satellitaire et ses Applications (URISA)

*En vue d'obtenir le diplôme de*

## Docteur

*En*

Technologies de l'Information et de la Communication

*Par*

### Sarra SAKJI-NSIBI

Ingénieur en Télécommunications de SUP'COM

*Titre*

**Indexation d'images multispectrales dans le domaine  
transformé en ondelettes**

**Directrice de thèse : Pr Amel BENZAZZA**

**Soutenue le 25 Février 2011 devant le jury d'examen composé de :**

**Présidente : Mme Mariem Jaidane, Professeure à l'ENIT**

**Rapporteurs : M. Marc Antonini, Directeur de Recherches CNRS, I3S, Sophia-Antipolis  
M. Chokri Benamar, Maître de Conférences à l'ENIS**

**Examineur : M. Moncef Gabbouj, Professeur à l'Université de Technologie de Tampere**

**Directrice de thèse : Mme Amel Benazza, Professeure à SUP'COM**

# Remerciements

Les travaux présentés dans ce mémoire ont été réalisés au sein de l'Unité de Recherche en Imagerie Satellitaire et ses Applications (URISA) à l'Ecole Supérieure de Communications de Tunis (SUP'COM).

Je tiens d'abord à exprimer toute ma gratitude à ma directrice de thèse Mme Amel Benazza, Professeure à Sup'Com pour m'avoir offert l'opportunité de travailler au sein de son équipe et d'avoir encadré ce travail de recherche. Je la remercie pour ses directives solides, ses conseils précieux, ainsi que sa bienveillance au bon déroulement de cette thèse.

Je suis, également, particulièrement reconnaissante aux membres du jury pour l'intérêt qu'ils ont porté à ce travail. Mes sincères remerciements s'adressent à M. Marc Antonini, Directeur de Recherches CNRS au laboratoire I3S à Nice-Sophia Antipolis (France) et à M. Chokri Benamar, Maître de Conférence à l'ENIS (Sfax, Tunisie) pour avoir accepté la lourde charge d'être les rapporteurs de ce mémoire. Je voudrais aussi adresser mes plus vifs remerciements à Mme Meriem Jaïdane, Professeure à l'ENIT, pour avoir accepté de présider mon jury de thèse ainsi qu'à M. Moncef Gabbouj, Professeur à l'Université de Technologie de Tampere (Finlande) de me faire l'honneur d'examiner ce travail.

Enfin, j'adresse mes remerciements à mes collègues de l'URISA pour leur collaboration, leurs conseils et leur soutien moral.

# Résumé

Cette thèse est consacrée à l'indexation d'images multispectrales dans le domaine transformé par des décompositions multirésolutions.

Dans un premier temps, nous avons dressé un état d'art sur les systèmes de recherche ainsi que sur les transformations en ondelettes et plus généralement des décompositions multirésolution.

Nous avons présenté ensuite nos contributions concernant l'étape d'extraction des signatures. La première consiste à établir grâce à *la théorie des copules* un modèle multivarié de la distribution conjointe des coefficients de toutes les composantes spectrales.

Concernant l'étape de la procédure de recherche, un nouvel algorithme de recherche arborescente exploitant la propriété d'échelonnabilité des décompositions multirésolution a été conçu afin d'assurer une recherche rapide et efficace selon une stratégie de parcours des résolutions les plus grossières vers les plus fines.

L'apport de nos contributions en terme de performances du système de recherche et de sa complexité opératoire a été mis en évidence par des simulations menées sur des bases d'apprentissage et de test contenant des images satellitaires multispectrales SPOT4 de plusieurs régions de la Tunisie.

*Mots-clés* : recherche d'images par le contenu, images multispectrales, transformations en ondelettes, décompositions multirésolutions, modèles de lois multivariées, dépendance interspectrales, dépendances inter-échelles, théorie des copules, recherche arborescente.

# Table des matières

<b>1</b>	<b>Systèmes de recherche d'images par le contenu</b>	<b>16</b>
1.1	Introduction . . . . .	16
1.2	Nécessité de systèmes de recherche d'images par le contenu . . . . .	16
1.3	Principe d'un système de recherche d'images par le contenu . . . . .	18
1.4	Descripteurs utilisés . . . . .	20
1.4.1	Les descripteurs de forme . . . . .	20
1.4.2	Les descripteurs de texture . . . . .	23
1.5	Exemples de systèmes de recherche d'images par le contenu . . . . .	24
1.6	Conclusion . . . . .	25
<b>2</b>	<b>Décompositions multirésolutions</b>	<b>27</b>
2.1	Rappels sur les transformations en ondelettes . . . . .	28
2.1.1	Principe . . . . .	28
2.1.2	Localisation temps-fréquence . . . . .	28
2.1.3	Ondelettes discrètes et analyse multirésolution . . . . .	29
2.1.4	Ondelettes à supports compacts . . . . .	32
2.1.5	Ondelettes biorthogonales . . . . .	33
2.1.6	Analyse et synthèse par algorithme rapide . . . . .	33
2.1.7	Cas de signaux numériques . . . . .	34
2.1.8	Extension multidimensionnelle . . . . .	34
2.1.9	Transformations en ondelettes $M$ bandes . . . . .	34
2.1.10	Ondelettes de deuxième génération . . . . .	35
2.2	Extensions de la transformation en ondelettes . . . . .	37
2.2.1	Les pyramides orientables . . . . .	37
2.2.2	La transformée en ondelettes complexe en arbre dual . . . . .	38
2.2.3	La transformée en ondelette bidimensionnelle $M$ bandes en arbre dual . . . . .	39
2.3	Intérêt de la transformation en ondelettes en compression . . . . .	41
2.3.1	Principe de la compression . . . . .	41
2.3.2	La norme JPEG2000 . . . . .	42

2.4	Conclusion . . . . .	42
<b>3</b>	<b>Contributions à l'extraction de descripteurs d'images multispectrales</b>	<b>43</b>
3.1	Signatures dans le domaine transformé en ondelettes . . . . .	44
3.1.1	Motivations . . . . .	44
3.1.2	Cas d'images monochromes . . . . .	44
3.1.3	Cas d'images multispectrales . . . . .	47
3.2	Contribution 1 : Modèle multivarié inter-spectral . . . . .	51
3.2.1	Motivations . . . . .	51
3.2.2	Bref rappel sur la théorie des copules . . . . .	52
3.2.3	Principes des modèles proposés . . . . .	54
3.2.4	Choix des distributions marginales . . . . .	55
3.2.5	Familles de copules proposées . . . . .	58
3.2.6	Sélection de la famille de copules . . . . .	63
3.3	Modèles prenant en compte les dépendances inter-échelles et inter-spectrales .	65
3.3.1	Motivation . . . . .	65
3.3.2	Principe . . . . .	66
3.3.3	Lois marginales . . . . .	67
3.3.4	Copules inter-spectrale et copules inter-échelle . . . . .	67
3.3.5	Signatures proposées . . . . .	67
3.3.6	Récapitulons ! . . . . .	68
3.4	Contribution 3 : agir sur la décomposition multi-échelle . . . . .	68
3.4.1	Motivation . . . . .	68
3.4.2	Recours au schéma de lifting vectoriel . . . . .	69
3.4.3	Recours à la transformation bidimensionnelle $M$ bande en arbre dual .	70
3.5	Conclusion . . . . .	71
<b>4</b>	<b>Accélération de la procédure de recherche</b>	<b>73</b>
4.1	Choix de la mesure de similarité . . . . .	73
4.1.1	Intérêt d'une fonction de similarité . . . . .	73
4.1.2	Mesures de similarité de type distances de Minkowski . . . . .	74
4.1.3	Distance symétrique de Kullback-Leibler . . . . .	75
4.2	Conception d'un algorithme de recherche arborescente . . . . .	78
4.2.1	Etape d'organisation des signatures . . . . .	79
4.2.2	Etape de recherche hiérarchique . . . . .	83
4.2.3	Avantages de notre approche . . . . .	83
4.3	Conclusion . . . . .	84

<b>5 Résultats expérimentaux</b>	<b>85</b>
5.1 Cadre expérimental . . . . .	85
5.1.1 Présentation des bases d'apprentissage et de test . . . . .	85
5.1.2 Critères d'évaluation de performances . . . . .	87
5.1.3 Travaux de références . . . . .	87
5.2 Performances du modèle multivarié reposant sur la théorie des copules . . . . .	88
5.2.1 Choix des densités marginales . . . . .	88
5.2.2 Prise en compte des dépendances inter-spectrales . . . . .	89
5.2.3 Prise en compte simultanée des dépendances inter-spectrales et inter-échelles . . . . .	91
5.2.4 Influence du choix de la copule sur les performances . . . . .	92
5.3 Influence du choix de la décomposition . . . . .	95
5.3.1 Décomposition en schéma de lifting vectoriel . . . . .	95
5.3.2 Apport de la décomposition 2D TOM-AD et invariance par rotation . . . . .	96
5.4 Influence du choix de la distance . . . . .	97
5.5 Apport de l'algorithme de recherche arborescente . . . . .	99
5.5.1 Performances et complexité opératoire de la recherche arborescente . . . . .	99
5.5.2 Influence du nombre d'échelles considérées . . . . .	101
5.6 Conclusion . . . . .	101
<b>A Estimation bayésienne de la famille de copule</b>	<b>106</b>
A.1 Position du problème . . . . .	106
A.2 Stratégie d'estimation bayésienne . . . . .	106
A.3 Sélection de la meilleure copule . . . . .	109
<b>B Généralisation du théorème de Sklar</b>	<b>110</b>
<b>C Schéma de lifting vectoriel</b>	<b>111</b>
<b>D Approximation du centroïde d'une classe au sens de la DSKL</b>	<b>113</b>

# Table des figures

1.1	Architecture générale d'un système de recherche d'images par le contenu . . .	19
2.1	Cellule élémentaire d'analyse. . . . .	31
2.2	Cellule élémentaire de synthèse. . . . .	31
2.3	Analyse par décomposition en ondelettes sur $J = 3$ niveaux de résolution. . .	32
2.4	Synthèse de $a_0(n)$ à partir des coefficients d'ondelettes pour $J = 3$ niveaux de résolution. . . . .	32
2.5	Structure d'analyse du schéma de lifting. . . . .	36
2.6	Structure de synthèse du schéma de lifting . . . . .	37
2.7	Transformée en ondelettes complexe en arbre dual à deux étages de décomposition. . . . .	38
2.8	Transformée en ondelettes $M$ bandes en arbre dual à deux étages de décomposition. . . . .	40
3.1	Approche univariée pour l'indexation d'images multispectrales codées dans le domaine des ondelettes. . . . .	48
3.2	Différentes parties à considérer dans l'étape d'extraction des signatures. . . .	50
3.3	Histogrammes conjoints empiriques des coefficients d'ondelettes des deux premières composantes de deux images SPOT4 différentes, pour $j = 1$ . . . . .	51
3.4	Densités gaussiennes généralisées bivariées ( $B = 2$ ) de paramètres respectifs $\rho_1 = 0.77$ , $s_1 = 0.54$ et $\rho_2 = 0.8$ , $s_2 = 0.73$ modélisant les distributions conjointes des coefficients d'ondelettes des deux premières composantes des deux images SPOT4 précédentes, pour $j = 1$ , où $\rho_1$ , (resp. $\rho_2$ ) est le coefficient de corrélation linéaire entre les deux composantes de l'image 1 (resp. l'image 2). . . . .	52
3.5	Histogrammes empiriques des coefficients d'ondelettes des 2 premières composantes d'une image SPOT4 pour $j = 1$ . . . . .	56
3.6	Histogrammes empiriques des valeurs absolues des coefficients d'ondelettes des 2 premières composantes d'une image SPOT4 pour $j = 1$ . . . . .	57

3.7	Densités gaussiennes généralisées de paramètres respectifs $\beta_1 = 0.19$ , $\alpha_1 = 0.47$ et $\beta_2 = 0.15$ , $\alpha_2 = 0.45$ modélisant les coefficients d'ondelettes horizontales des 2 premières composantes de l'image SPOT4 précédente pour $j = 1$ . . . . .	58
3.8	Densités gamma de paramètres respectifs $\gamma_1 = 0.28$ , $\delta_1 = 4.29$ et $\gamma_2 = 0.4$ , $\delta_2 = 3.32$ modélisant les valeurs absolues des coefficients d'ondelettes horizontales des 2 premières composantes de l'image SPOT4 précédente pour $j = 1$ . . . . .	59
3.9	Densités de probabilité $f_{w_j}$ résultants de la modélisation des distributions conjointes des coefficients d'ondelettes des deux premières composantes ( $B = 2$ ) de deux images SPOT4 par des DGG et des copules gaussiennes de paramètres respectifs $\rho_1 = 0.74$ et $\rho_2 = 0.78$ , $j = 1$ , où $\rho_1$ , respectivement $\rho_2$ est le coefficient de corrélation linéaire entre les deux composantes. . . . .	60
3.10	Dépendances inter-échelle entre coefficients d'ondelettes. . . . .	65
3.11	Modélisation des coefficients d'ondelettes prenant en compte les distributions marginales, les dépendances inter-spectrales et inter-échelles . . . . .	66
4.1	Organisation de la base des signatures. . . . .	80
4.2	Procédure de recherche arborescente. . . . .	84
5.1	Différentes catégories d'images dans la base d'apprentissage. . . . .	86
5.2	Courbes précision-rappel obtenues avec l'approche univariée <b>a</b> : densités marginales modélisées par des DGG <b>b</b> : densités marginales modélisées par des DG. . . . .	89
5.3	Gain en performances dû à l'utilisation d'une approche multivariée <b>a</b> : approche univariée avec des densités marginales modélisées par des DGG, <b>b</b> : densité conjointe modélisée par une DGGM. . . . .	89
5.4	Apport en terme de rappel-précision de notre première contribution : <b>a</b> : approche univariée avec des densités marginales modélisées par des DG, <b>b</b> : approche basée sur l'utilisation d'une ACI, <b>c</b> : approche multivariée utilisant une DGGM, <b>d</b> : dépendances inter-spectrales modélisées par une copule gaussienne. . . . .	90
5.5	Apport en rappel-précision de la prise en compte simultanée des dépendances inter-canaux et inter-échelles : <b>a</b> : dépendances inter-spectrales et inter-échelles ignorées, <b>b</b> : dépendances inter-spectrales modélisées par une copule gaussienne et dépendances inter-échelles ignorées, <b>c</b> : dépendances inter-échelles modélisées par une copule gaussienne et dépendances inter-spectrales ignorées, <b>d</b> : prise en compte simultanée des dépendances inter-spectrales et inter-échelles. . . . .	92
5.6	Apport en précision-rappel de l'approche bayésienne : <b>a</b> : dépendances inter-spectrales ignorées, <b>b</b> : dépendances inter-spectrales modélisées par une copule gaussienne, <b>c</b> : dépendance inter-spectrales modélisées par une copule de Gumbel, <b>d</b> : approche bayésienne. . . . .	94



5.7	Comparaison des performances en terme de précision-rappel obtenues avec la décomposition 5/3 et la décompositions DEC3 : <b>a</b> : décomposition 5/3 suivie de l'approche univariée, <b>b</b> : décomposition 5/3 et dépendances inter-spectrales modélisées par une copule gaussienne, <b>c</b> : décomposition DEC3 suivie de l'approche univariée, <b>d</b> : décomposition DEC3 et dépendances inter-spectrales modélisées par une copule gaussienne. . . . .	96
5.8	Comparaison des performances en terme de précision-rappel obtenues par différentes décompositions multirésolutions : <b>a</b> : S-TOD, <b>b</b> : TOC-AD , <b>c</b> : TOM-AD, $M = 2$ , <b>d</b> : TOM-AD, $M = 3$ , <b>e</b> : Décomposition en pyramides orientables, $O = 6$ . . . . .	97
5.9	Courbes précision-rappel obtenues avec $B = 3$ , $J = 3$ , $j_u = 1$ pour différents modèles paramétriques et différentes distances, (à gauche) modèles DG et DGG de lois marginales et (à droite) modèle DG pour les lois marginales. . . . .	99
5.10	Courbes précision-rappel en pourcentage obtenues avec $J = 3$ , $j_u = 1$ et la distance DSKL dans les deux cas de la recherche exhaustive ou arborescente. . . . .	100

# Liste des tableaux

1.1	Moteurs de recherche commerciaux et expérimentaux les plus connus. . . . .	26
2.1	Propriétés des différentes décompositions multirésolutions, $O$ : nombre de filtres orientés, $M$ : paramètre de la décomposition $M$ bandes. . . . .	41
3.1	Signatures issues des modèles multivariés proposés. . . . .	64
5.1	Précision moyenne en pourcentage obtenue avec différents couples de familles de copules traduisant les dépendances inter-canaux (en colonnes) et inter-échelles (en lignes). . . . .	93
5.2	Rappel moyen en pourcentage obtenu avec différents couples de familles de copules traduisant les dépendances inter-canaux (en colonnes) et inter-échelles (en lignes). . . . .	93
5.3	Rapport moyen entre le nombre des DSKL calculées lors de la recherche arborescente et celui des DSKL calculées lors de la recherche exhaustive pour différentes valeurs du nombre $R$ d'images retournées. . . . .	100
5.4	Nombre final de sous-classes, précision moyenne et rappel moyen obtenus avec la recherche arborescente pour les 2 algorithmes de classification K-moyennes et ACC pour $R = 20$ et $j_u = 1, 2, 3$ . . . . .	101

# Liste des abréviations

- AAC : Algorithme d'Agglomération Compétitive.
- ACI : Analyse en Composantes Indépendantes.
- AHM : Ali-Mikhail-Haq.
- CBIR : *Content-Based Image Retrieval System*.
- CNES : Centre National d'Etudes Spatiales.
- DEN : Distance Euclidienne Normalisée.
- DG : Distribution Gamma.
- DGG : Distribution Gaussienne Généralisée.
- DGGM : Distribution Gaussienne Généralisée Multivariée.
- DSKL : Distance Symétrique de Kullback-Leibler.
- EBCOT : Embedded Block Coding with Optimal Truncation Points.
- EROS : Earth Resources Observation and Science Center.
- EZBC : *Embedded Zero Block Coder*.
- EZW : *Embedded Zerotree Wavelet coding*.
- FBI : *Federal Bureau of Investigation*.
- FCM : *Fuzzy C-means*.
- IRM : Imagerie par Résonance Magnétique.
- JPEG : *Joint Photographic Experts Group*.
- PR : Précision.
- RC : Rappel.
- RRMNM : Rang de Recherche Modifié Normalisé Moyen.
- RIF : Réponse Impulsionnelle Finie.
- SPACL : *Signal Processing and Coding Laboratory*.
- SPIHT : *Set Partitioning In Hierarchical Trees*.
- S-TOD : Transformée en Ondelettes Discrète Dyadique et Séparable.
- TCD : Transformée en Cosinus Discrète.
- TFD : Transformée de Fourier Discrète.
- TFMA : Transformée de Fourier-Mellin Analytique.
- TOC-AD : Transformée en Ondelettes Complexe en Arbre Dual.

- TOM-AD : Transformée en Ondelettes  $M$ -bandes en Arbre Dual.
- USGS : *United States Geological Survey*.
- ZTE : Zero Tree Entropy.

# Introduction générale

Ces dernières années ont été marquées par un progrès considérable des techniques d'acquisition et de stockage d'images numériques qui a été à l'origine de la production permanente de bases d'images numériques intervenant dans plusieurs domaines tels que la santé, l'audiovisuel, l'architecture, la télésurveillance et la télédétection, domaine auquel nous nous intéressons dans cette thèse. En effet, plusieurs satellites de télédétection parcourent notre ciel munis de capteurs optiques ou radar à haute résolution tels que IKONOS, SPOT5, ALOS, GeoEye-1, WorldView-2 et bientôt TerraSAR-X. Ces capteurs de plus en plus sophistiqués en terme de résolution spatiale et spectrale ont permis la constitution de bases d'images satellitaires panchromatiques, multispectrales et hyperspectrales très souvent stockées par les agences spatiales telles que les bases d'images des satellites SPOT, les bases d'images aériennes et satellitaires du Earth Resources Observation and Science Center (EROS), les bases d'images Microsoft TerraServer. Ces collections d'images sont régulièrement consultées par les scientifiques et les experts en imagerie pour servir dans plusieurs applications telles que l'agriculture, la géologie, la cartographie et l'urbanisme. A moyen terme, toutes ces applications de traitement d'images nécessitent la mise en place de systèmes automatiques permettant un accès rapide et sélectif aux images de ces bases. En particulier, il est indispensable de concevoir des systèmes permettant à l'utilisateur de retrouver rapidement des images répondant à une requête donnée. Les limitations des systèmes de recherche d'images par mots-clé ont été rapidement mises en évidence comme la subjectivité de la description et ont motivé le développement de procédures de recherche d'images par le contenu. Leur principe réside dans l'indexation *automatique* des bases d'images en exploitant uniquement le *contenu visuel* de ces images.

Généralement, un système de recherche d'images par le contenu s'articule autour des deux étapes principales que sont l'extraction des descripteurs du contenu visuel et la procédure de recherche dans la base des descripteurs servant d'identifiants ou de signatures aux images de la base.

- L'extraction des signatures :

L'étape d'extraction des signatures a pour rôle de décrire le contenu visuel de l'image par un ensemble d'attributs. Elle peut se faire, au moment où l'image est insérée dans

la base et *hors- ligne*. Par conséquent, elle ne nécessite pas forcément une complexité opératoire réduite.

– La procédure de recherche :

La procédure de recherche consiste à trouver les images de la base dont les descripteurs sont les plus proches des descripteurs de la requête relativement à une mesure de similarité prédéfinie. C'est une étape *en ligne* qui s'effectue chaque fois qu'un utilisateur a besoin de consulter la base d'images pour répondre à sa requête.

Ainsi, les performances d'un système de recherche d'images par le contenu sont fortement liées aux choix des descripteurs, de la distance et de l'algorithme de recherche. Par ailleurs, les bases d'images satellitaires sont souvent volumineuses. En effet, les satellites de nos jours fournissent, quotidiennement, jusqu'à 450 images multispectrales par satellite qui sont elles-même très volumineuse. Une image SPOT5 à 5 m de résolution spatiale contient par exemple,  $12000 \times 12000$  pixels typiquement codés sur 8 bpp. Ces tailles croissantes des bases posent des problèmes au niveau de leur stockage et de leur transmission sur des canaux de bande passante réduite. Pour remédier à ces contraintes, la solution qui s'impose est le recours à des techniques de compression afin de réduire les redondances présentes au niveau des données d'une image et de générer des représentations compactes nécessitant moins de ressources mémoire ou de temps de transmission que les données d'origine. A cet égard, plusieurs méthodes de compression ont été développées et rapportées dans la littérature : elles peuvent être avec ou sans pertes. Parmi celles-ci, beaucoup d'intérêt a été porté aux méthodes de codage dans le domaine transformé dont les plus connus sont le domaine de la transformée en cosinus discrète (TCD) utilisé dans la norme de compression JPEG et le domaine des transformations en ondelettes de seconde génération utilisé dans la norme de compression JPEG2000. Ces transformations en ondelettes ont la particularité d'offrir une représentation parcimonieuse de l'image à différentes résolutions. Pour indexer des bases d'images satellitaires monochromes ou multispectrales, stockées sous une forme codée, une approche intuitive consiste à décoder les images de la base avant d'extraire les signatures. Cette étape de décodage requiert des ressources supplémentaires en terme de calculs et de mémoire de stockage. Pour l'éviter, une solution est d'indexer ces images *directement* dans le domaine codé en calculant les signatures non plus sur les pixels, mais plutôt sur les coefficients transformés de l'image. Dans le cas d'images codées dans le domaine transformé en ondelettes, ceci permet, en plus d'éviter le décodage, d'exploiter la nature de la représentation échelle-espace en vue d'une meilleure caractérisation visuelle du contenu de l'image.

Pour toutes ces raisons, la motivation dans cette thèse consiste à concevoir un système de recherche d'images multispectrales opérant dans le domaine transformé en ondelettes. En effet, si un grand nombre de travaux a porté sur l'indexation d'images monochromes dans le domaine transformé, à notre connaissance, peu d'études ont été consacrées au cas des images multispectrales au moment du démarrage de cette thèse début 2006.

Nos contributions ont touché à la fois, à l'extraction des signatures et à la procédure de recherche.

- Concernant l'étape d'extraction des signatures, une première contribution a consisté à proposer un modèle de loi multivariée paramétrique pour décrire la distribution statistique conjointe des coefficients d'ondelettes de toutes les composantes spectrales de l'image. En plus de sa flexibilité, ce modèle présente l'avantage de prendre en compte tous les types de dépendance statistique entre les coefficients d'ondelettes relatifs à toutes les échelles et à toutes les composantes spectrales. Les paramètres de ce modèle ont alors servi de signatures des images.

La seconde contribution a été d'appliquer des transformations multirésolutions sophistiquées pour permettre d'extraire des signatures pertinentes et éventuellement invariantes par rapport à certaines "perturbations" de l'image.

- Pour l'étape de procédure de recherche, notre apport a été d'abord, d'étudier l'influence du choix de la mesure de similarité sur les performances de la recherche. En exploitant la propriété d'échelonnabilité des transformations multirésolutions, nous avons également conçu un algorithme de recherche et arborescente dans le but de diminuer la complexité opératoire.

## Plan de la thèse

Ce manuscrit est organisé de la manière suivante :

- **Le chapitre 1** est une introduction des systèmes de recherche d'images par le contenu. Nous soulignons la nécessité du recours à un système de recherche d'images par le contenu et nous présentons en détail son principe et ses principales étapes. Nous établissons par la suite un bref état de l'art des différents descripteurs d'images monochromes rapportés dans la littérature ainsi que des moteurs de recherche d'images par le contenu les plus connus.
- **Dans le chapitre 2**, nous présentons les transformations en ondelettes et plus généralement les transformations multirésolutions en insistant sur leur principe ainsi que leurs propriétés attrayantes pour la compression et la recherche d'images par le contenu.
- **Dans le chapitre 3**, nous présentons nos deux principales contributions pour la description par le contenu d'images multispectrales à partir de représentations multirésolutions. La première consiste à établir des modèles statistiques aptes à refléter les propriétés statistiques des coefficients d'ondelettes d'images multicomposantes. La seconde contribution consiste à utiliser des décompositions multirésolutions sophistiquées permettant de calculer des signatures pertinentes et robustes par rapport à certaines "perturbations" de l'image.

- **Le chapitre 4** est consacré à nos apports au niveau de l'étape de la procédure de recherche. Dans un premier temps, nous décrivons les distances utilisées dans notre travail pour comparer les vecteurs des signatures. Nous proposons par la suite un algorithme de recherche arborescente permettant un accès échelonné en résolution des images de la base.
- **Dans le chapitre 5**, nous étayons l'ensemble de nos contributions par des résultats expérimentaux en veillant à nous comparer aux travaux de l'état de l'art.
- Enfin, une conclusion générale et les perspectives laissées ouvertes font l'objet du dernier chapitre.



# Chapitre 1

## Systemes de recherche d'images par le contenu

### 1.1 Introduction

Ces dernières années, les systèmes de recherche d'images par le contenu sont de plus en plus utilisés dans différents domaines d'application tels que la télédétection. Ils ont pour but de permettre, entre autres, aux scientifiques et experts en imagerie de consulter régulièrement des bases d'images de grande taille (certaines archives d'images satellites SPOT comptent par exemple plus d'une dizaine de millions d'images). Dans ce chapitre, nous commençons, dans la première partie, par mettre en évidence la nécessité du recours à des systèmes de recherche d'images par le contenu. Ensuite, dans la deuxième partie, nous présentons le principe sur lequel repose un système de recherche d'images par le contenu ainsi que ses principales étapes. Dans la troisième partie, nous dressons un panorama des principaux descripteurs décrits dans la littérature pour caractériser le contenu des images tandis qu'un bref état de l'art des moteurs de recherche d'images par le contenu les plus connus est établi dans la quatrième partie.

### 1.2 Nécessité de systèmes de recherche d'images par le contenu

Depuis une dizaine d'années, on assiste à une véritable révolution en imagerie numérique du point de vue technologique et celui des applications. Cette révolution est due au progrès considérable des techniques d'acquisition et de stockage (capteurs numériques, médias de stockage de données, microprocesseurs, périphériques d'affichage). Il en résulte une production permanente et considérable d'images numériques de grande taille et riches en informations visuelles très utiles dans beaucoup de domaines tels que la télédétection, la santé, l'audiovisuel, l'architecture et la télésurveillance. On distingue deux types de bases d'images [1, 2] :

les bases généralistes contenant des millions d'images [2] de contenu très varié comme les bases d'images des agences de presse ou de publicité ou les bases d'images sur Internet (dans ce dernier cas, on peut parler de milliards d'images circulant sur le réseau [2]) et les bases spécifiques contenant des images relatives à un domaine d'application bien précis tel que la médecine, la sécurité ou l'imagerie satellitaire. Ce dernier domaine, auquel nous nous intéressons dans cette thèse, connaît un perfectionnement continu et considérable des capteurs optiques ou radars embarqués sur des plateformes aériennes ou à bord de satellites à haute résolution, tels que IKONOS, SPOT5, ALOS, GeoEye-1, WorldView-2 [3] et TerraSAR-X [4] qui fournissent quotidiennement des images monochromatiques et multispectrales, très volumineuses (une image SPOT5 monochromatique contient  $12000 \times 12000$  pixels [5]). D'autres satellites seront également lancés prochainement comme Pléiades [6] pour lequel on prévoit jusqu'à 450 images quotidiennes à 70 cm de résolution et GeoEye-2 capable de discerner des objets jusqu'à une taille de 25 cm. Il en résulte d'énormes bases d'images satellitaires multi/hyperspectrales stockées par les agences spatiales telles que les archives d'images de satellites SPOT qui comptent plus de 12 millions d'images selon le site du Centre National d'Etudes Spatiales (CNES), la base d'images aériennes et satellitaires du Earth Resources Observation and Science Center (EROS) [7] contenant, en partie, des images fournies par les satellites Landsat 7, ASTER et MODIS, et la base d'images Microsoft TerraServer [8], considérée comme étant la plus grande base d'images aériennes et satellitaires, contenant des images fournies par les agences spatiales United States Geological Survey (USGS) et Sovinformspuznik (l'agence spatiale fédérale russe).

Les bases d'images aériennes et satellitaires sont régulièrement consultées par les scientifiques et les experts en imagerie pour servir dans plusieurs applications. En effet, la télédétection trouve entre autres ses applications en :

- agriculture (pour la classification des types de cultures, l'évaluation de l'état des cultures ou la cartographie des caractéristiques du sol) ;
- foresterie (pour la mesure des propriétés biophysiques des forêts ou la détection des zones brûlées) ;
- géologie (pour l'exploration minière ou pétrolière, la détection des déformations de terrain, la surveillance des phénomènes naturels comme les tornades, inondations, volcans) ;
- hydrologie (pour l'évaluation de l'humidité du sol et la prévention des inondations) ;
- urbanisme (pour la cartographie de base pour des données GIS et l'étude de l'expansion et du développement urbain) ;
- océanographie (pour l'évaluation des réserves de poissons et des mammifères marins et la gestion du transport maritime).

Néanmoins, ces diverses applications nécessitent souvent un accès rapide et sélectif aux bases volumineuses d'images aériennes et satellitaires. Une technique ancienne de recherche

d'images est la recherche par texte associé [9, 10]. Elle repose sur l'annotation des images d'une base par des attributs textuels tels que les mot-clé décrivant le contenu de l'image, le titre, l'auteur. Ainsi, la recherche d'images s'apparente à une recherche de mots-clés. Néanmoins, cette approche présente des inconvénients majeurs. D'une part, elle nécessite l'intervention humaine pour annoter les images de la base. Ceci est non seulement pénible, voire irréalisable sur de grandes bases, mais présente aussi l'inconvénient de décrire les images d'une manière subjective et souvent approximative ou inexacte. D'autre part, la recherche textuelle est souvent confrontée aux contraintes linguistiques telles que le multilinguisme et les synonymes [11].

Pour remédier à ces problèmes, des systèmes de recherche d'images par le contenu (ou CBIR en anglais pour *Content-Based Image Retrieval System*) ont permis d'indexer *automatiquement* des bases d'images en exploitant uniquement leur *contenu visuel*. Dans la section suivante, nous présentons le principe général d'un système de recherche d'images par le contenu et nous y détaillons les différentes étapes.

### 1.3 Principe d'un système de recherche d'images par le contenu

La recherche d'images par le contenu [1] définit l'ensemble des techniques utilisées pour chercher automatiquement, à partir d'une base d'images, les images répondant à une requête donnée en s'appuyant seulement sur des descripteurs de leur contenu visuel tels que la couleur, la texture, la forme ou n'importe quelle information extraite directement de l'image. La recherche d'images par le contenu peut être classée selon les trois catégories qui sont les suivantes.

- La recherche ciblée ou recherche d'une image spécifique : l'utilisateur dispose d'une image unique bien déterminée qu'il cherche dans la base d'images considérée par exemple l'image satellitaire d'une ville ou d'une zone bien déterminée. L'utilisateur a besoin uniquement de cette image.
- La recherche par catégorie : Contrairement à la recherche ciblée, l'utilisateur ne cherche pas une image spécifique bien connue mais l'ensemble des images appartenant à une catégorie bien déterminée comme par exemple les images d'une zone urbaine en imagerie satellitaire.
- La recherche navigation libre ou butinage : le but de la recherche n'est pas précis et peut évoluer au cours de la navigation.

En ce qui concerne le type de la requête, la recherche par l'exemple est la plus connue et la plus élémentaire. Dans ce cas, l'utilisateur choisit une image exemple et le système détermine les images de la base les plus semblables visuellement à l'image requête. Néanmoins, dans certains cas, l'utilisateur ne dispose pas d'image requête similaire à l'image ou aux images recherchées. On parle dans ce cas de recherche par image mentale [12, 13] : l'utilisateur a

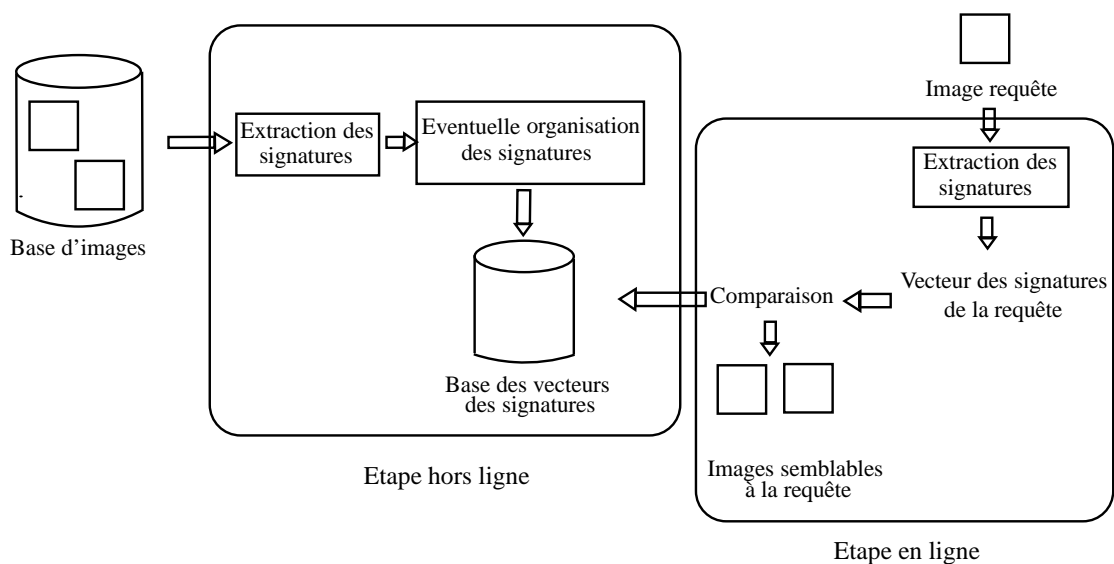


FIG. 1.1 – Architecture générale d'un système de recherche d'images par le contenu

une idée plus au moins précise de ce qu'il cherche et il utilise les fonctionnalités du système de recherche d'images par le contenu pour exprimer et satisfaire sa requête. Dans ce cadre, il a été proposé de conduire la recherche à partir d'un croquis dessiné par l'utilisateur ou à partir d'un dictionnaire regroupant des parties d'image, caractérisant la base, proposées à l'utilisateur pour l'aider à composer et à exprimer sa requête [14, 15].

D'autre part, la requête peut être globale ou locale. Dans le premier cas, la recherche de l'utilisateur porte sur des images dont l'apparence visuelle approche celle de l'image requête (si elle existe) dans toute sa globalité. Dans le second cas, l'intérêt de l'utilisateur porte sur une ou plusieurs entités associées à l'image. On parle alors de recherche par requête partielle qui peut porter sur un objet, un bloc, une région d'intérêt ou des points d'intérêt de l'image considérée. La recherche vise à extraire les images de la base contenant la même région d'intérêt définie par la requête ou une région d'intérêt qui lui est semblable [16].

Pour assurer une recherche fiable et efficace, un système de recherche d'images par le contenu repose, comme le montre la figure 1.1, sur deux étapes principales [1].

– L'extraction des signatures :

Cette étape consiste à représenter les images de la base ainsi que l'image requête par un ensemble d'attributs, appelés aussi descripteurs ou signatures, qui reflètent le contenu visuel de l'image. Pour chaque image, on définit alors un vecteur de descripteurs. Un descripteur peut caractériser une image dans son ensemble. On parle alors de descripteur global ou bien uniquement une partie de l'image (objet, région d'intérêt), il est ainsi appelé descripteur local. Qu'elle soit globale ou locale une signature appartient généralement à une des trois catégories suivantes : les descripteurs de couleur, les des-

cripteurs de forme et les descripteurs de texture.

Vu le caractère subjectif de la recherche d'images par le contenu, le choix d'un descripteur dépend de son habilité à décrire le plus fidèlement possible le contenu de l'image et ceci, selon le domaine d'application et les besoins de l'utilisateur. Les descripteurs des images de la base sont alors extraits et stockés dans la base des signatures.

– La procédure de recherche :

La procédure de recherche a pour rôle d'extraire les images de la base dont les vecteurs des signatures sont les plus proches du vecteur des signatures de la requête relativement à une distance ou critère de similarité prédéfini [1, 17]. Le choix de la distance dépend de la nature des vecteurs des signatures [17]. Cette étape est alors une étape dite *en ligne* qui s'effectue chaque fois qu'un utilisateur a besoin de consulter la base d'images pour répondre à sa requête.

Dans certains travaux, la procédure de recherche contient un système de bouclage de pertinence [15, 18] : la recherche se fait alors sur plusieurs passes. A chaque passe, l'utilisateur indique au système la pertinence des images retournées par rapport à ce qu'il cherche par exemple en sélectionnant les images retournées par le système répondant le mieux à son besoin. Grâce à un algorithme d'apprentissage, le système utilise ces indications pour raffiner le recherche à la passe suivante jusqu'à la satisfaction de l'utilisateur.

Dans la section suivante, nous nous focalisons davantage sur l'étape d'extraction des descripteurs et nous passons en revue les principales signatures utilisées pour décrire les images monochromatiques (à niveaux de gris).

## 1.4 Descripteurs utilisés

En général, les descripteurs sont des descripteurs de forme et/ou de texture et/ou de couleur. Dans cette partie, nous nous intéressons uniquement aux descripteurs de forme et de texture utilisés dans le cas d'images monochromatiques. Le cas d'images couleur et (en général multispectrales) sera présenté plus loin.

### 1.4.1 Les descripteurs de forme

Une image est généralement caractérisée par les différents objets qui la composent ainsi que les matériaux qui composent les objets. Cette distinction permet de définir deux propriétés complémentaires de l'image : les formes et les textures [19].

La forme est une propriété globale d'un objet permettant de caractériser son apparence globale comme son contour [19]. La forme d'un objet demeure inchangée quelle que soit sa position dans l'image [20]. En général, un prétraitement consistant à segmenter l'image pour

en extraire les différentes formes existantes est souvent nécessaire. Une étape d'extraction de contour peut aussi être envisagée [21, 22].

Afin d'assurer une description fiable d'objets, en plus d'être significatifs, les descripteurs de forme doivent satisfaire certaines propriétés telles que l'invariance par rapport à certaines transformations géométriques (translation, rotation, facteur d'échelle), la robustesse vis-à-vis du bruit ou de faibles variations de la forme et éventuellement la complétude c'est-à-dire le fait qu'un objet est identifié de manière unique à partir de ses descripteurs à une transformation près [23, 24]. Dans la littérature, les descripteurs de forme peuvent être classés en deux catégories principales : les descripteurs de contours et les descripteurs de régions.

– Descripteurs de contours

Ces descripteurs sont particulièrement utilisés quand l'information sur le contour est disponible et facile à extraire par exemple dans le cas d'objets simples. Les descripteurs de contours les plus utilisés sont les suivants.

– *Signatures de contours dans le domaine spatial :*

Ce sont des descripteurs calculés directement à partir de la représentation d'un contour dans le domaine spatial [21, 25] telles l'aire, la circularité, le périmètre, l'excentricité, les coordonnées complexes et la courbure. Néanmoins, bien que simples à calculer, ces descripteurs n'ont souvent pas l'habileté de discriminer des formes légèrement différentes [21]. De plus, la plupart de ces descripteurs ne sont pas robustes au bruit entachant les intensités des pixels [21]. Enfin, extraire des signatures invariantes par rotation directement dans le domaine spatial s'avère complexe du point de vue opératoire [21]. Pour toutes ces raisons, plusieurs travaux visent à définir des signatures à partir d'une représentation du contour dans le domaine fréquentiel tel que le domaine de la Transformée de Fourier Discrète (TFD).

– *Signatures de contour dans le domaine de la TFD :*

Ces signatures sont largement utilisées en analyse et reconnaissance de forme ainsi qu'en recherche d'images par le contenu [21, 26, 27, 28]. La concentration de l'énergie d'un contour sur un faible nombre de coefficients basses fréquences de sa TFD permet une représentation parcimonieuse de l'objet. En effet, les coefficients de Fourier aux basses fréquences fournissent une information générale et approximative sur la forme tandis que les coefficients aux hautes fréquences contiennent une information concernant les détails fins de la forme. En plus, ces descripteurs ont la particularité d'être robustes au bruit. Enfin, comme les coefficients de la TFD contiennent une information relative à la taille, à l'orientation et à la position de l'objet, il est possible de les rendre invariants par rapport au facteur d'échelle, à l'orientation et à la translation [21, 27, 28].

– Descripteurs de régions

Quand l'information sur le contour est insuffisante ou indisponible, il est nécessaire

d'extraire des descripteurs directement à partir de tous les pixels de l'objet. Dans ce contexte, on peut citer deux familles de descripteurs souvent citées dans la littérature.

– *Les moments de l'image :*

Les moments sont des quantités très utilisées en statistique pour la caractérisation des distributions de variables aléatoires et en physique pour décrire la distribution spatiale de masse d'un objet. Ainsi, en traitement d'images, les moments ont l'habilité de traduire les propriétés statistiques d'une forme.

Les moments les plus connus sont les moments géométriques [29, 30] qui offrent une représentation des propriétés géométriques de la forme telles que la surface traduite par le moment d'ordre zéro, le barycentre exprimé à l'aide de moments de premier ordre et les principaux axes d'un objet exprimés en fonction des moments de second ordre. Les moments de troisième ordre expriment le degré de déviation de l'objet par rapport à une symétrie centrale (coefficient de dissymétrie) tandis que les moments d'ordre quatre nous renseignent sur l'excentricité de la forme (coefficient d'aplatissement). Des descripteurs invariants par transformations (translation, rotation ou facteur d'échelles) appelés moments invariants, sont généralement calculés à partir de ces moments géométriques [29]. Il est également montré que l'ensemble infini des moments géométriques assure la propriété de complétude c'est-à-dire qu'un objet peut être décrit et reconstruit d'une manière unique à partir de l'ensemble infini de ses moments géométriques. Néanmoins, comme cette reconstruction est impraticable, d'autres moments ont été introduits tels que les moments orthogonaux (moments de Zernike et moments de Legendre) [29, 31] permettant une reconstruction quasi-parfaite d'une image à partir d'un petit nombre de moments ainsi que les moments complexes [32] qui sont une généralisation des moments géométriques. Ils ont été introduits afin de traiter simultanément les changements d'orientation et d'échelle des objets.

– *Les descripteurs de la transformée de Fourier-Mellin analytique :*

La transformée de Fourier-Mellin analytique (TFMA) introduite par Ghorbel [33] est une transformation réversible qui offre une description complète de l'objet transformé. En plus, la TFMA peut être facilement approximée en utilisant différents algorithmes d'estimation [34]. Dans [24, 35], Derrode et Ghorbel ont proposé une famille complète de descripteurs d'un objet à niveaux de gris issus de sa TFMA. Ces descripteurs ont la particularité d'être invariants par translation, rotation et facteur d'échelle. Ils sont aussi stables vis-à-vis de faibles variations de forme de l'objet. D'autre part, les auteurs ont montré que l'objet original peut être facilement reconstruit à partir d'un nombre fini de ses descripteurs en utilisant une approximation numérique de la TFMA inverse.

## 1.4.2 Les descripteurs de texture

Contrairement à la forme qui définit l'apparence globale d'un objet, la texture caractérise les matériaux composant un objet comme l'ensemble des motifs relativement fins, nombreux, se répétant généralement d'une manière périodique qui tapissent la partie de l'image correspondant à un objet [19]. En d'autres termes, la texture reflète les propriétés locales des pixels de l'objet ainsi que les relations avec leur voisinage. Dans la littérature, les approches de description de texture peuvent être classées en 3 catégories.

- Les approches statistiques :

Elles ont pour but de décrire les propriétés statistiques spatiales de la texture. Dans ce cadre, la matrice de co-occurrence spatiale est l'outil le plus utilisé [36, 37]. Elle est définie de la manière suivante. Pour une image ayant des niveaux de gris entre 0 et  $I - 1$ , une matrice de taille  $I \times I$  est définie pour une distance  $d$  donnée telle que son élément générique  $(i, j)$  correspond au nombre de fois qu'un pixel prenant une valeur de niveau de gris  $i$  est situé à la distance  $d$  d'un pixel prenant une valeur de niveau de gris  $j$ . Le succès de la matrice de co-occurrence tient à la fidélité avec laquelle elle décrit la texture en traduisant la relation spatiale entre les différents niveaux de gris dans une image. En pratique, plusieurs descripteurs peuvent être calculés à partir de ces matrices comme par exemple l'entropie [37, 38].

- Les approches structurelles :

Les approches structurelles exploitent le fait qu'une région texturée présente un ensemble de propriétés spatiales locales qui se répètent d'une manière quasi-périodique. Ainsi, ces approches caractérisent une texture par un ensemble de primitives qui apparaissent dans l'image suivant un arrangement spatial répétitif (ou périodique) [39].

Une primitive de texture est définie comme étant un ensemble de deux ou plus de pixels adjacents ayant des intensités de niveau de gris bien déterminées et dont la relation entre eux est définie par un attribut donné, par exemple : même niveau de gris ou forme polygonale bien déterminée caractérisée par sa longueur ou son élongation [40].

Une texture est ainsi décrite par ses primitives ainsi que par la manière dont elles sont organisées spatialement dans l'image traduite par une loi de placement.

- Les approches spectrales :

Contrairement aux travaux cités précédemment qui extraient des descripteurs de texture directement à partir de la représentation spatiale de l'image, d'autres travaux utilisent une représentation de la texture dans le domaine fréquentiel. Les transformées suivantes sont parmi les plus utilisées pour l'analyse de la texture.

- *La TFD :*

Comme la TFD est généralement l'outil adéquat pour l'analyse de fonctions périodiques, la description de la texture dans le domaine de la TFD est motivée d'une part,



par la nature périodique des textures. D'autre part, la représentation de la texture dans le domaine de la TFD facilite l'extraction de signatures invariantes par translation, rotation ou facteur d'échelle, tâche plus compliquée dans le domaine spatial [41]. Néanmoins, comme la TFD est une fonction localisée en fréquence, elle permet uniquement une analyse globale de la texture. Ainsi la TFD n'est pas adéquate pour la description, par exemple, de textures hétérogènes qui varient d'une position à l'autre de l'image. Ce problème peut être résolu par une représentation espace-fréquence de l'image.

– *Les décompositions multirésolutions :*

Un inconvénient commun à tous les travaux de description de texture cités précédemment consiste à considérer l'image à une seule échelle d'analyse. Or, des études sur le système visuel humain [42, 43] ont prouvé que le cortex visuel peut être modélisé par un ensemble de canaux indépendants caractérisés par des orientations et des fréquences différentes. Ainsi une représentation multirésolution offrirait une meilleure caractérisation visuelle du contenu de l'image surtout dans le domaine de la reconnaissance de la texture. En conséquence, beaucoup de travaux se sont intéressés à l'extraction de descripteurs de texture après l'application d'une transformation espace-fréquence ou multi-échelle de l'image. Une fois, la transformation multirésolution appliquée, les signatures peuvent s'apparenter aux descripteurs de texture avec une approche statistique dans le domaine spatial mais cette fois calculées sur les coefficients multirésolutions de l'image comme les signatures d'énergie [44, 45], la matrice de co-occurrence [149] et les paramètres d'un modèle statistique [46, 47, 48]. Ainsi, la possibilité de calculer ces signatures à différentes orientations et différentes échelles permet une description de la texture plus proche de celle offerte par les mécanismes de la vision humaine.

La forme et la texture jouent des rôles complémentaires pour la caractérisation des images monochromes [19, 49]. Pour cette raison, certains travaux et systèmes de recherche d'images combinent les signatures de forme et de texture pour assurer une meilleure description de l'image [49]. Dans la section suivante, nous allons décrire quelques systèmes de recherche d'image les plus utilisés en précisant les descripteurs choisis par chaque système.

## 1.5 Exemples de systèmes de recherche d'images par le contenu

Plusieurs moteurs de recherche d'images par le contenu ont été développés. Actuellement, la recherche d'images par le contenu commence à peine à avoir un succès commercial avéré. C'est pourquoi la plupart des moteurs de recherche sont des moteurs expérimentaux. Ceci est dû au fait que la plupart des utilisateurs de bases d'images préfèrent spécifier sémantiquement

le contenu de l'image (ou des images) qu'ils cherchent. Dans le tableau 1.1, nous présentons certains moteurs de recherche par le contenu les plus connus en citant, pour chaque système, la nature des images concernées, les différents types de descripteurs utilisés ainsi que le type de la requête. Notons que seuls les moteurs QBIC et VIRAGE sont des moteurs commerciaux, tandis que les autres moteurs sont des moteurs expérimentaux. De manière générale, les signatures sont très variées comprenant aussi bien des attributs de haut que de bas niveau qui se renforcent et coopèrent de manière complémentaire sous forme de combinaison de caractéristiques.

## 1.6 Conclusion

Dans ce chapitre, nous avons introduit les systèmes de recherche d'images par le contenu en présentant leur nécessité, en particulier en imagerie satellitaire, leur principe, ainsi que leurs différentes étapes. Nous avons également passé en revue les différents descripteurs de texture et de forme calculés sur les images monospectrales en niveaux de gris. Nous avons par la suite, présenté les différents systèmes de recherche d'images commerciales et académiques les plus connus.

Par ailleurs, nous avons souligné que l'extraction de signatures à partir de représentations multirésolutions de l'image est une des méthodes de description d'images les plus utilisées. Ainsi, dans le chapitre suivant, nous nous focalisons sur la présentation des décompositions multirésolutions ainsi que leur application en recherche et en compression d'images.

Moteur	Laboratoire/Entreprise	Type d'images	Type de la requête	Type des descripteurs	Autres propriétés
QBIC [50]	IBM (USA)	Couleur	Par image mentale ou par l'exemple Requête globale ou partielle	Couleur, texture et forme	Bouclage de pertinence
VIRAGE [51]	Virage Manages and Monitors Multimedia San Francisco, USA	Couleur : base de Sybase, Object design, Objectivity	Par l'exemple ou par image mentale	Couleur et texture	Inclusion de mots clés, Bouclage de pertinence
Photobook [52]	Massachussets Institute of Technology (MIT), USA	Naturelles (Vistex) Visage (FaceID)	Par image mentale puis par l'exemple	Texture et forme	Reconnaissance de visage, Bouclage de pertinence
MARS [53]	Université de l'Illinois et de Californie, USA	Couleur	Par l'exemple ou par image mentale	Couleur, texture et forme	Bouclage de pertinence
VisualSEEK [54]	Image and Advanced Television Lab Université de Colombie, USA	Couleur	Par image mentale puis par l'exemple Requête partielle	Couleur et forme	
NETRA [55]	Département de Génie Electrique et Informatique de l'Université de Californie, USA	Couleur	Par l'exemple Requête globale ou partielle	Couleur, texture et forme	
IKONA [56]	INRIA, France	Satellitaires Photos professionnelles Objets 3D Base de visages	Par l'exemple Requête globale ou partielle	Descripteurs génériques : couleur, texture, forme spécifiques : signatures de visage	Architecture client-serveur Descripteurs textuels Bouclage de pertinence
SIMPLICITY [57]	Université de Standford Californie, USA	Bases génériques	Par l'exemple	Signatures adaptatives : couleur forme ou texture	Etape de classification sémantique
CIRES [58]	Computer and Vision Research Center Université de Texas, USA	Naturelles contenant des objets structurés	Par l'exemple	Bas niveau : texture, forme Hauts niveau : Coins, polygones	Bouclage de pertinence
MUVIS [59]	Université de Technologie de Tampere, Finlande	Génériques	Par l'exemple, requête globale ou partielle	Couleur, texture et forme	Base multimédia vidéo et audio

TAB. 1.1 – Moteurs de recherche commerciaux et expérimentaux les plus connus.

## Chapitre 2

# Décompositions multirésolutions

Comme on l'a vu au chapitre 1, en indexation d'images, à l'instar des autres applications de traitement de signal, on a souvent besoin de générer une représentation qui soit plus exploitable. La transformation la plus populaire est la transformée de Fourier. Elle permet de passer d'une représentation dans le domaine spatial à une représentation dans le domaine fréquentiel. Néanmoins, la transformée de Fourier présente l'inconvénient de fournir une information fréquentielle globale de l'image sans prendre en considération la localisation spatiale de la distribution des fréquences. La transformée de Fourier à court terme a été conçue pour remédier à ce problème. Toutefois, cette transformation ne peut représenter l'image qu'à une seule échelle ce qui n'est pas forcément adéquat pour discriminer les détails fins des détails grossiers contenus dans l'image. L'une des alternatives pour pallier ces limitations est la transformée en ondelettes et de manière générale les décompositions multirésolutions. En effet, leur principe est d'analyser l'image par des atomes modulés en échelle, centrés en différentes positions afin de fournir une représentation échelle-espace. Les décompositions multirésolutions s'avèrent aussi intéressantes en compression d'image grâce à la compacité des représentations qu'elles génèrent.

Ce chapitre est organisé de la manière suivante. Dans la section 2.1, nous commençons par rappeler le principe d'une transformation en ondelettes dans le cas 1D, nous enchaînons avec la présentation de la décomposition en ondelettes séparable dans le cas 2D et nous finirons par détailler le principe des ondelettes de seconde génération basées sur le schéma de lifting. Dans la section 2.2, nous décrivons d'autres transformations multirésolutions considérées comme une généralisation de la transformation en ondelettes. Dans la section 2.3, nous expliquons l'intérêt de la transformation en ondelettes en compression d'images notamment pour la norme de compression JPEG2000. La conclusion de ce chapitre est donnée dans la section 2.4.

## 2.1 Rappels sur les transformations en ondelettes

### 2.1.1 Principe

Les ondelettes sont des fonctions de  $L^2(\mathbb{R})$ . Elles sont générées par une translation du paramètre de localisation  $u \in \mathbb{R}$  et une dilatation de facteur d'échelle  $s \in \mathbb{R}_+^*$  à partir d'une seule fonction appelée ondelette-mère  $\psi \in L^2(\mathbb{R})$  intégrable, de moyenne nulle et qui oscille localement et ceci de la manière suivante :

$$\forall t \in \mathbb{R} \quad \psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right). \quad (2.1)$$

Quand les paramètres  $u$  et  $s$  varient, l'ondelette  $\psi_{u,s}(t)$  change respectivement de localisation temporelle et d'échelle, dans ce dernier cas, elle peut couvrir différents points d'intérêt du signal à analyser à différentes échelles. La transformée en ondelettes continue d'une fonction  $f$  de  $L^2(\mathbb{R})$  correspond au calcul des coefficients  $c(u, s)$  tels que :

$$c(s, u) = \int_{\mathbb{R}} f(t) \psi_{u,s}(t) dt. \quad (2.2)$$

La transformée en ondelettes inverse s'obtient aisément :

$$f(t) = \int_{\mathbb{R}_+^*} \int_{\mathbb{R}} c(s, u) \psi_{u,s}(t) ds du. \quad (2.3)$$

Notons que cette transformée inverse existe si et seulement si la fonction ondelette-mère satisfait la condition d'admissibilité suivante :

$$\int \frac{|\widehat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty \quad (2.4)$$

où  $\widehat{\psi}$  est la transformée de Fourier de  $\psi$ . Cette condition implique la nullité de la moyenne temporelle de  $\widehat{\psi}$  et donc que  $\widehat{\psi}$  soit oscillante (une onde).

### 2.1.2 Localisation temps-fréquence

Supposons que :

$$\int_{\mathbb{R}} \psi(t)^2 dt = 1 \quad \text{et} \quad \int_{\mathbb{R}} t \psi(t)^2 dt = 0 \quad (2.5)$$

On définit respectivement les supports temporel  $\Delta t > 0$  et fréquentiel  $\Delta \omega > 0$  de la fonction  $\psi$  de la manière suivante :

$$\begin{aligned} (\Delta t)^2 &= \int_{\mathbb{R}} t^2 \psi(t)^2 dt \\ (\Delta \omega)^2 &= \int_{\mathbb{R}} (\omega - \bar{\omega})^2 |\widehat{\psi}(\omega)|^2 d\omega \end{aligned} \quad (2.6)$$

où  $\bar{\omega} \triangleq 2 \int_0^{+\infty} \psi(\omega) d\omega$ . Les équations montrent que les valeurs de  $\psi(t)$  (respectivement  $\psi(\omega)$ ) sont essentiellement réparties dans l'intervalle  $[-\frac{\Delta t}{2}, \frac{\Delta t}{2}]$  (respectivement  $[\bar{\omega} - \frac{\Delta \omega}{2}, \bar{\omega} + \frac{\Delta \omega}{2}]$ ). Par ailleurs, on montre aisément que  $\psi_{u,s}$  a un support temporel égal à  $\frac{\Delta t}{s^2}$  centré

en  $u$  et un support fréquentiel de taille  $s^2\Delta\omega$  centré en  $s\bar{\omega}$ . Par conséquent, le produit des supports temporel et fréquentiel de  $\psi_{u,s}$  est le même que celui de  $\psi(t)$  et ce, quel que soit le couple  $(u, s)$ . Rappelons que ce produit est borné en vertu du principe d'incertitude de Heisenberg-Gabor puisqu'on peut montrer que :

$$\left(\frac{\Delta t}{s^2}\right)(s^2\Delta\omega) \geq \frac{1}{2}. \quad (2.7)$$

La différence fondamentale avec l'analyse de Fourier à court terme est que les résolutions temporelles et fréquentielles varient dans le plan temps-fréquence. Ainsi l'analyse du signal  $f(t)$  à  $\Delta\omega$  constant permet d'améliorer la localisation temporelle quand la fréquence augmente.

La flexibilité de cette analyse présente divers avantages :

- Il s'agit d'une analyse proche de celle réalisée par le système visuel (et auditif) humain [42, 43].
- Elle permet une analyse des phénomènes transitoires.

### 2.1.3 Ondelettes discrètes et analyse multirésolution

Dans la cas où les paramètres  $(s, u)$  sont discrétisés  $s = 2^j$  et  $u = 2^j n$ , l'ondelette  $\psi(t)$  est construite à partir d'une fonction dite d'échelle  $\phi(t)$  qui satisfait les équations suivantes :

$$\begin{aligned} \forall t \in \mathbb{R} \quad \phi(t) &= \sqrt{2} \sum_{\ell=-\infty}^{+\infty} h(\ell)\phi(2t - \ell) \\ \psi(t) &= \sqrt{2} \sum_{\ell=-\infty}^{+\infty} g(\ell)\phi(2t - \ell) \end{aligned} \quad (2.8)$$

où  $h(\ell)$  et  $g(\ell)$  sont des séquences de réels qui définissent respectivement des filtres passe-bas et passe-haut. Par ailleurs, la donnée de la fonction  $\phi(t)$  et donc de  $\psi(t)$  aboutit à ce qu'on appelle une analyse multirésolution.

Les principes de base de l'analyse multirésolution ont été posés par Mallat [60]. Une analyse multirésolution de  $L^2(\mathbb{R})$  permet d'organiser l'information contenue dans un signal 1D en un ensemble de signaux approximations et de détails à différents niveaux successifs de résolution  $j = 1, \dots, J$ . Le signal détail à un étage  $j > 1$  est par définition égal à la différence entre l'approximation du signal original à l'étage  $j - 1$  et son approximation à l'étage  $j$ . Ainsi, en allant des échelles les plus grossières vers les échelles les plus fines, on accède à des représentations de plus en plus précises du signal initial. Les signaux approximation s'obtiennent par une succession de projections orthogonales du signal original sur une suite emboîtée de sous-espaces vectoriels fermés  $V_j$  de  $L^2(\mathbb{R})$ , qui présentent les propriétés suivantes.

1.  $\forall j \in \mathbb{Z}, V_{j+1} \subset V_j$
2.  $\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f(t/2) \in V_{j+1}$
3.  $\lim_{j \rightarrow +\infty} \bigcap_{j \in \mathbb{Z}} V_j = 0$  et  $\lim_{j \rightarrow -\infty} \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$

4. Il existe une fonction  $\phi(t) \in L^2(\mathbb{R})$  appelée fonction d'échelle telle que  $\phi(t - n), n \in \mathbb{Z}$  est une base orthonormale de  $V_0$ .

Le signal détail à chaque niveau  $j$  représente une projection orthogonale du signal original sur le sous-espace vectoriel  $W_j$  orthogonal à  $V_j$  et vérifiant  $V_{j-1} = V_j \oplus W_j$ .

L'analyse multirésolution est généralement associée à une décomposition multirésolution par les ondelettes. En effet, Mallat [60] a défini la décomposition on ondelettes orthogonales en  $J$  étages d'un signal monodimensionnel continu à valeurs réelles  $f(t)$  de  $L^2(\mathbb{R})$  comme les projections successives  $f_{V_j}$  et  $f_{W_j}$  sur les sous-espaces  $V_j$  et  $W_j$ , de bases orthonormales respectives  $\{2^{-j/2}\phi(t/2^j - n)\}_{n \in \mathbb{Z}}$  et  $\{2^{-j/2}\psi(t/2^j - n)\}_{n \in \mathbb{Z}}$  :

$$f_{V_j}(t) = \sum_{n=-\infty}^{+\infty} a_j(n)2^{-j/2}\phi(t/2^j - n), \quad f_{W_j}(t) = \sum_{n=-\infty}^{+\infty} w_j(n)2^{-j/2}\psi(t/2^j - n) \quad (2.9)$$

où les coefficients  $a_j(n)$  et  $w_j(n)$ ,  $n \in \mathbb{Z}$  sont appelés respectivement coefficients d'échelle (ou d'approximation) et coefficients d'ondelettes (ou détail), ils sont définis par :

$$a_j(n) = 2^{-j/2} \int_{-\infty}^{+\infty} f(t)\phi(t/2^j - n)dt, \quad w_j(n) = 2^{-j/2} \int_{-\infty}^{+\infty} f(t)\psi(t/2^j - n)dt. \quad (2.10)$$

Mallat [60] a démontré que les coefficients approximation  $a_j(n)$  et détail  $w_j(n)$  à un étage  $j > 0$  peuvent se déduire à partir des coefficients d'échelle à l'étage  $j - 1$  par filtrage avec les filtres  $\tilde{h}$  et  $\tilde{g}$  suivi d'une décimation d'un facteur 2 de la manière suivante :

$$a_j(n) = \sum_{\ell=-\infty}^{\ell=\infty} a_{j-1}(\ell)\tilde{h}(\ell - 2n), \quad w_j(n) = \sum_{\ell=-\infty}^{\ell=\infty} a_{j-1}(\ell)\tilde{g}(\ell - 2n) \quad (2.11)$$

où  $\tilde{h}(\ell) = h(-\ell)$  et  $\tilde{g}(\ell) = g(-\ell)$ . L'application récursive de ce banc de filtres, appelé banc de filtres d'analyse, de l'étage 0 à l'étage le plus grossier  $J$  correspond à l'algorithme récursif de Mallat. Cet algorithme suppose qu'on dispose de coefficients d'approximation  $a_0(n)$  de  $f(t)$  à l'étage  $j = 0$ .

Mallat [60] a également montré que le signal approximation  $a_{j-1}(n)$  à un étage  $j - 1$  peut être reconstruit à partir des signaux approximation et détail  $a_j(n)$  et  $w_j(n)$  à l'étage  $j$  après application d'un banc de filtres de synthèse dual correspondant à un filtre passe-bas de réponse impulsionnelle  $h$  et un filtre passe-haut de réponse impulsionnelle  $g$  :

$$a_{j-1}(n) = \sum_{\ell=-\infty}^{\ell=\infty} h(n - 2\ell)a_j(\ell) + \sum_{\ell=-\infty}^{\ell=\infty} g(n - 2\ell)w_j(\ell). \quad (2.12)$$

Ainsi, les coefficients approximation  $a_0(n)$  correspondant à l'espace vectoriel  $V_0$  de base  $\phi(t - n), n \in \mathbb{Z}$  peuvent être déduits récursivement des coefficients approximation  $a_j$  à l'étage le plus grossier  $J$  et les coefficients détail  $w_1, \dots, w_J$  aux étages  $1, \dots, J$ .

La figure 2.1 représente la cellule basique de l'analyse qui permet de générer  $a_j$  et  $w_j$  à partir de  $a_{j-1}$ . Elle met en évidence :

- un décimateur qui divise par 2 le nombre de coefficients en éliminant les échantillons impairs.
- une paire de filtres numériques en quadrature miroir (*Quadrature Mirror Filter* ou *QMF*)

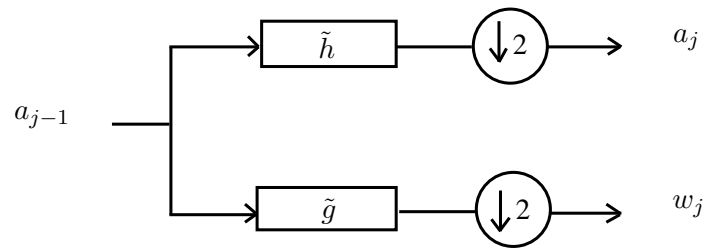


FIG. 2.1 – Cellule élémentaire d’analyse.

La figure 2.2 fournit le schéma de la cellule élémentaire de synthèse qui est duale de celle de l’analyse. Un interpolateur d’ordre 2 permet de doubler la longueur des coefficients d’entrée en insérant un zéro entre deux échantillons successifs de l’entrée.

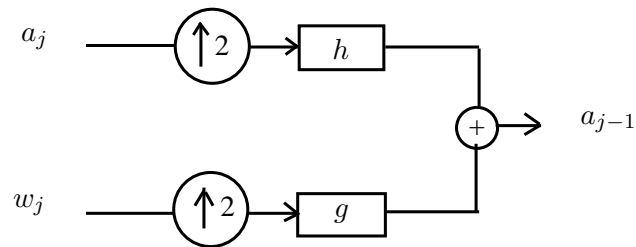


FIG. 2.2 – Cellule élémentaire de synthèse.



En mettant en cascade les cellules élémentaires d'analyse tel qu'illustré dans la figure 2.3, des approximations successives  $a_1(n), \dots, a_J(n)$  de  $f(t)$  sont obtenues de manière récursive ainsi que les détails perdus  $w_1(n), \dots, w_J(n)$ .

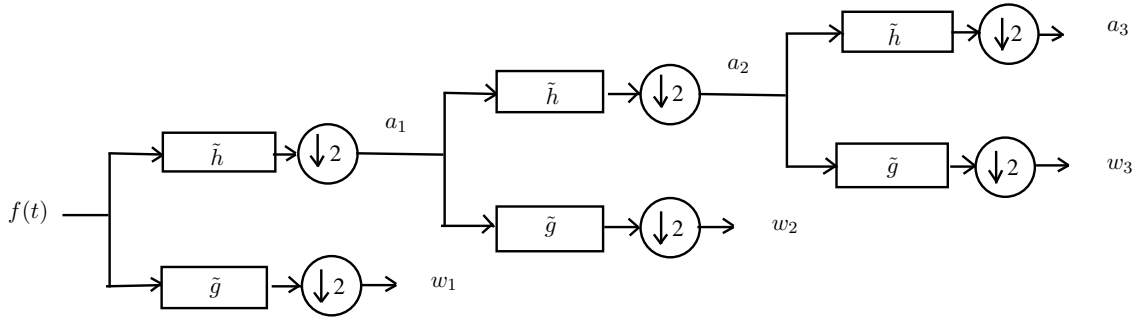


FIG. 2.3 – Analyse par décomposition en ondelettes sur  $J = 3$  niveaux de résolution.

Bien entendu, la synthèse de  $a_0(n)$  s'obtient de manière duale comme le montre la figure 2.4.

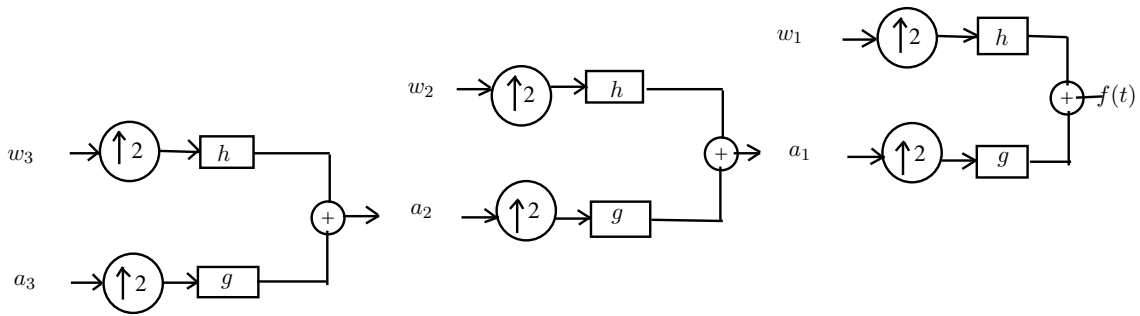


FIG. 2.4 – Synthèse de  $a_0(n)$  à partir des coefficients d'ondelettes pour  $J = 3$  niveaux de résolution.

Ainsi, grâce aux travaux de Mallat, il a été établi une équivalence entre la théorie des ondelettes, l'analyse mutirésolution et la décomposition par bancs de filtres qui est propre au traitement du signal et connue bien avant la théorie des ondelettes.

### 2.1.4 Ondelettes à supports compacts

Notons que pour simplifier la mise en oeuvre de la transformée en ondelettes, des filtres à réponse impulsionnelle finie (RIF) sont choisis pour l'analyse et la synthèse. Ceci revient à imposer à  $\phi(t)$  et  $\psi(t)$  d'être à support compact. Les ondelettes de ce type ont été étudiées par Daubechies [61].

Plus précisément, Daubechies a établi les résultats suivants.

- La base de Haar est la seule base orthonormale mettant en jeu des filtres numériques RIF à phase linéaire.
- Une base d'ondelette à support compact est caractérisée par l'expression de  $|H(\omega)|^2$  en fonction d'un paramètre de régularité.

### 2.1.5 Ondelettes biorthogonales

Il apparaît donc que les ondelettes orthonormales à support compact doivent vérifier des conditions assez strictes. La non linéarité de la phase des filtres d'analyse et de synthèse est un inconvénient notamment dans les applications de compression [62]. Pour contourner ce problème, la contrainte d'orthogonalité est alors relâchée [63]. L'idée est d'utiliser deux bases d'ondelettes biorthogonales :  $(\sqrt{2^j}\psi(2^j t - k))_{(j,k) \in \mathbb{Z}^2}$  et  $(\sqrt{2^j}\tilde{\psi}(2^j t - k))_{(j,k) \in \mathbb{Z}^2}$  telles que :

$$\forall j, j', k, k' \in \mathbb{Z} \quad \int_{\mathbb{R}} \sqrt{2^j}\psi(2^j t - k)\sqrt{2^{j'}}\tilde{\psi}(2^{j'} t - k')dt = \delta_{j-j', k-k'}. \quad (2.13)$$

La fonction  $f(t)$  se décompose sous la forme suivante :

$$f(t) = \sum_{k,j \in \mathbb{Z}^2} w_j(k)\psi(2^j t - k) \quad (2.14)$$

où

$$w_j(k) = 2^j \int_{\mathbb{R}} f(t)\tilde{\psi}(2^j t - k)dt. \quad (2.15)$$

Dans le cas des ondelettes biorthogonales, les filtres de synthèse  $\tilde{h}$  et  $\tilde{g}$  sont reliés aux filtres d'analyse  $h$  et  $g$  de la manière suivante :

$$\begin{aligned} \forall k \in \mathbb{Z} \quad \tilde{h}(k) &= (-1)^k g(-k) \\ \tilde{g}(k) &= (-1)^k h(-k). \end{aligned} \quad (2.16)$$

Par conséquent, la perte d'orthogonalité s'est traduite par le fait que les réponses fréquentielles des filtres primaires et duaux ne sont plus complexes conjuguées les unes des autres. Notons que les ondelettes 9/7 de Cohen-Daubechies-Fouveau comptent parmi les ondelettes biorthogonales les plus connues surtout en codage d'images.

### 2.1.6 Analyse et synthèse par algorithme rapide

Si l'on sépare les parties paires et impaires de  $h$  et  $g$  et si le filtrage se réalise par l'algorithme rapide de la transformée de Fourier, il est possible d'aboutir à un algorithme de décomposition en ondelettes de complexité comparable à celle de la transformée de Fourier [64].

### 2.1.7 Cas de signaux numériques

Les décompositions en ondelettes que nous venons de décrire ne concernent que les signaux analogiques  $f(t)$  de  $L^2(\mathbb{R})$  bien que l'algorithme de calcul des coefficients soit discret. Or dans la pratique, on a affaire à des signaux numériques. Il est donc logique de chercher à étendre la notion de décomposition en ondelettes au cas discret. Cette étude a été menée par Rioul [65] pour définir une base d'ondelettes orthonormales dans  $L^2(\mathbb{Z})$  (dites ondelettes discrètes). Notons que ces ondelettes se déduisent les unes des autres par translation à niveau de résolution donné comme dans  $\ell^2(\mathbb{R})$ . En revanche, le passage de  $j - 1$  à  $j$  ne se fait plus par dilatation. Dans la pratique, comme on manipule souvent des signaux discrets de taille finie, on assimile les  $N$  échantillons aux coefficients d'approximation  $a_0(n)$  à l'étage  $j = 0$ .

### 2.1.8 Extension multidimensionnelle

Il est possible d'étendre la notion de décomposition en ondelettes aux signaux multidimensionnels, en particulier aux images numériques. Dans ce dernier cas, la décomposition de l'image originale  $I_0(m, n)$ ,  $m, n \in \{1, \dots, N\}$  se fait récursivement en  $J$  étages de manière séparable. A chaque étage  $j = 0, \dots, J - 1$ , l'image approximation  $a_j(m, n)$ ,  $m, n \in \{1, \dots, \frac{N}{2^j}\}$  est traitée en 2 temps en analysant successivement les lignes puis les colonnes du résultat intermédiaire. Cette décomposition résulte en une sous-image approximation  $a_J(m, n)$ ,  $m, n \in \{1, \dots, \frac{N}{2^J}\}$  à l'étage  $J$  et 3 sous-images détail  $w_{j,o}(m, n)$  à chaque étage  $j$  orientées horizontalement ( $o = 1$ ), verticalement ( $o = 2$ ) et diagonalement ( $o = 3$ ). Ainsi la transformée en ondelette bidimensionnelle discrète dyadique et séparable (2D S-TOD) possède l'avantage d'être une transformation simple à implémenter et non redondante (le nombre total de coefficients d'approximation et d'ondelettes à chaque étage  $j$  est égal au nombre de pixels de l'image initiale). Notons, néanmoins qu'il existe des décompositions en ondelettes non séparables [66].

### 2.1.9 Transformations en ondelettes $M$ bandes

Une analyse multirésolution en  $M$  bandes de  $L^2(\mathbb{R})$  [67] met en oeuvre une fonction d'échelle  $\psi_0(t) \in L^2(\mathbb{R})$  et  $M - 1$ ,  $M \geq 2$  ondelettes mères  $\psi_m(t) \in L^2(\mathbb{R})$ ,  $m \in \{1, \dots, M - 1\}$  satisfaisant l'équation suivante :

$$\forall m \in \{0, \dots, M - 1\}, \quad \frac{1}{\sqrt{M}} \psi_m\left(\frac{t}{M}\right) = \sum_{\ell=-\infty}^{\ell=+\infty} h_m(\ell) \psi_0(t - \ell) \quad (2.17)$$

où les suites  $\{h_m(\ell)\}_{\ell \in \mathbb{Z}}$  sont réelles de carré sommable.

Les signaux approximation s'obtiennent par une succession de projections orthogonales du signal original sur une suite emboîtée de sous-espaces vectoriels fermés  $V_j$  de  $L^2(\mathbb{R})$ , ( $j =$

$1, \dots, J$ ) admettant comme base orthonormale  $\{M^{\frac{j}{2}}\psi_0(\frac{t}{M^j} - \ell)\}_{\ell \in \mathbb{Z}}$ . Les  $M - 1$  signaux détail à chaque niveau  $j$  représentent des projections orthogonales du signal original sur le sous-espace vectoriel  $W_j^m$  vérifiant  $V_{j-1} = V_j \bigoplus_{m=1}^{M-1} W_j^m$  et admettant comme bases orthonormales  $\{M^{\frac{j}{2}}\psi_m(\frac{t}{M^j} - \ell)\}_{\ell \in \mathbb{Z}}$ ,  $m = 1, \dots, M - 1$ .

Ce type de transformations présente l'avantage de représentations plus fines en fréquence. Elle a aussi l'intérêt d'offrir plus de liberté pour le choix des filtres. Par exemple, il est possible de générer des ondelettes à support compact orthogonales avec des filtres symétriques ce qui n'est pas le cas pour les ondelettes dyadiques [67].

### 2.1.10 Ondelettes de deuxième génération

A partir des années 1995, les transformations en ondelettes (dites de seconde génération) [68] ont commencé à être préférées aux transformations de première génération décrites précédemment. En effet, leur principe est plus intuitif et repose sur le concept du schéma dit de lifting. Comme le montre la figure 2.5, ce schéma comprend les 3 parties suivantes :

- A un étage donné  $j$ , l'ensemble des coefficients  $a_j(n)$  est divisé en coefficients à positions paires et impaires.
- La redondance spatiale présente au niveau du signal approximation est exploitée en prédisant les coefficients à positions paires à partir de ceux à positions impaires grâce à un opérateur de prédiction  $\mathcal{P}$ . L'erreur de prédiction constitue le signal détail  $d_{j+1}(n)$  à l'étage  $j + 1$  donné par :

$$w_{j+1}(n) = a_j(2n) - \mathcal{P}(\bar{\mathbf{a}}_j(n)) \quad (2.18)$$

où le vecteur  $\bar{\mathbf{a}}_j(n)$  contient seulement des coefficients à positions impaires de  $a_j(n)$ .

- Le signal approximation  $a_{j+1}(n)$  à l'étage  $j + 1$  est ensuite calculé en conservant une partie du signal approximation  $a_j(n)$  à l'étage  $j$ , traduite par les coefficients de positions impaires, ainsi qu'une moyenne de l'ensemble du signal traduit par la modification des coefficients de positions impaires par un vecteur  $\bar{\mathbf{d}}_{j+1}(n)$  comportant certains coefficients détail de  $d_{j+1}(n)$  via un opérateur de mise à jour  $\mathcal{U}$  :

$$a_{j+1}(n) = a_j(2n - 1) + \mathcal{U}(\bar{\mathbf{d}}_{j+1}(n)). \quad (2.19)$$

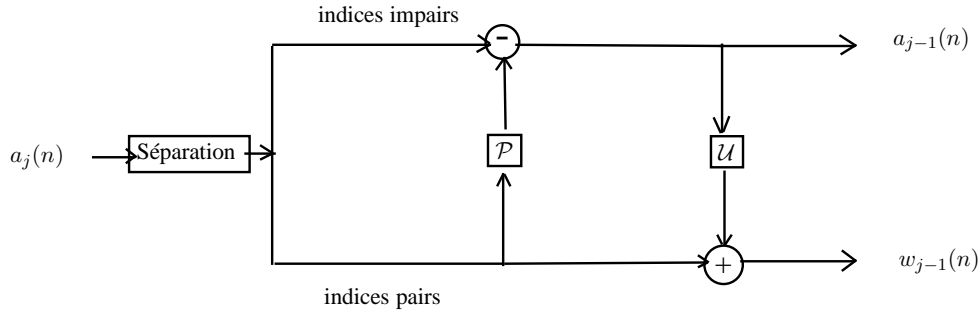


FIG. 2.5 – Structure d’analyse du schéma de lifting.

Cette nouvelle structure de décomposition appelle les commentaires suivants.

- Comme le montre la figure 2.6, la structure intrinsèque de ce schéma de décomposition garantit sa réversibilité sans aucune contrainte sur les opérateurs  $\mathcal{U}$  et  $\mathcal{P}$ . Cette absence de contrainte facilite la conception de ces opérateurs.
- On peut inclure des quantificateurs dans les opérateurs  $\mathcal{P}$  et  $\mathcal{U}$ . Dans ce cas, si le signal original  $a_0(n)$  est à valeurs entières, les coefficients approximation et d’ondelettes sont également à valeurs entières ce qui représente une propriété intéressante dans le contexte de compression sans pertes.
- On peut montrer que toute décomposition en ondelettes biorthogonale associée à des filtres RIF peut être mise en œuvre grâce à un nombre fini de niveaux de lifting suivis en sortie de multiplications par deux constantes non nulles. L’intérêt de représenter le banc de filtres classique sous la forme de schéma de lifting est qu’il est facile à vérifier différentes propriétés utiles des ondelettes. En pratique, un choix très vaste se présente pour les opérateurs  $\mathcal{P}$  et  $\mathcal{U}$  qui peuvent être linéaires ou non linéaires. La transformée 5/3 [69] est l’une des décompositions de deuxième génération les plus performantes et les plus utilisées en compression d’images fixes. Elle se traduit par les équations suivantes :

$$\begin{aligned} w_{j+1}(n) &= a_j(2n) - \lfloor \frac{1}{2}(a_j(2n-1) + a_j(2n+1)) \rfloor \\ a_{j+1}(n) &= a_j(2n-1) + \lfloor \frac{1}{4}w_{j+1}(n-1) + w_{j+1}(n) \rfloor. \end{aligned} \quad (2.20)$$

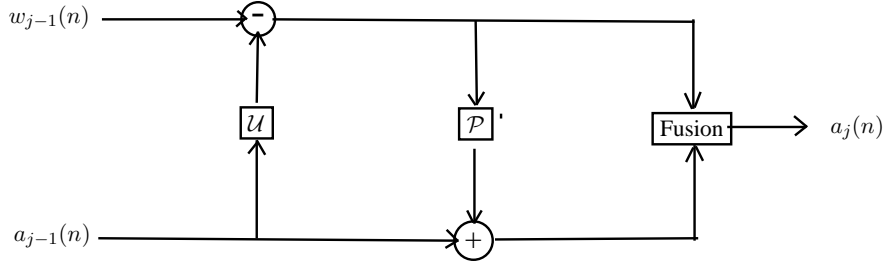


FIG. 2.6 – Structure de synthèse du schéma de lifting

## 2.2 Extensions de la transformation en ondelettes

Malgré ses propriétés intéressantes, la transformée en ondelettes discrète dyadique et séparable (2D S-TOD) possède plusieurs limitations telles que le manque d'invariance par translation et la pauvreté relative en information directionnelle. Ainsi, d'autres décompositions multirésolutions ont-elles été conçues pour remédier à ces problèmes. Dans ce qui suit, nous présentons certaines de ces décompositions les plus utilisées en traitement d'images.

### 2.2.1 Les pyramides orientables

Les pyramides orientables [70, 71] sont des transformations multi-échelles, multi-orientations selon lesquelles l'image est décomposée en  $J$  étages de  $O$  sous-bandes localisées à des orientations différentes. Cette décomposition se fait grâce à la projection de l'image à chaque étage  $j \in \{1, \dots, J\}$  sur un nombre choisi  $O$  de filtres orientés. Ces filtres ont la particularité d'assurer la propriété de l'orientabilité. Des filtres sont dit orientés s'ils forment des copies l'un de l'autre à des orientations différentes. Cette propriété est assurée grâce à la possibilité d'exprimer chaque filtre orienté localisé à une orientation  $o = 1, \dots, O$  comme combinaison linéaire d'un ensemble de filtres de base.

Une propriété très utile de la décomposition en pyramides orientables est ainsi déduite. Notons par  $\mathbf{w}_j(m, n)$ , respectivement  $\mathbf{w}_j^\theta(m, n)$  le vecteur formé par les coefficients  $w_{j,o}(m, n)$  (respectivement  $w_{j,o}^\theta(m, n)$ ) à différentes orientations  $o \in \{1, \dots, O\}$ , à la même position  $(m, n)$  et à la même résolution  $j$  relatif respectivement à une image  $I$  donnée et à sa version  $I_\theta$  tournée d'un angle  $\theta$ . Ces vecteurs sont exprimés par les équations suivantes :

$$\begin{aligned} \mathbf{w}_j(m, n) &= (\mathbf{w}_{j,1}(m, n), \dots, \mathbf{w}_{j,o}(m, n), \dots, \mathbf{w}_{j,O}(m, n))^T \\ \mathbf{w}_j^\theta(m, n) &= (\mathbf{w}_{j,1}^\theta(m, n), \dots, \mathbf{w}_{j,o}^\theta(m, n), \dots, \mathbf{w}_{j,O}^\theta(m, n))^T. \end{aligned} \quad (2.21)$$

Le vecteur  $\mathbf{w}_j^\theta(m, n)$  peut s'écrire simplement en fonction du vecteur  $\mathbf{w}_j(m, n)$  de l'image  $I$  de la manière suivante :

$$\mathbf{w}_j^\theta(m, n) = \mathbf{R}^\theta \mathbf{w}_j(n, m) \quad (2.22)$$

avec  $\mathbf{R}^\theta$  est la matrice orthogonale dont les colonnes sont les fonctions d'interpolation relatives

à l'angle  $\theta$ .

Cette propriété est très utile dans beaucoup d'applications de traitement d'images surtout dans le cas d'invariance par rotation dans la mesure où les propriétés d'une copie tournée d'une image peuvent être facilement déduites à partir de celles de l'image originale. En plus la décomposition en pyramides orientées est une transformation invariante par translation et riche en information directionnelle. Elle permet également une flexibilité pour le choix du nombre d'orientations  $O$ . Toutes ces raisons font des pyramides orientables un outil puissant et largement utilisé en traitement d'images surtout pour la classification [72] et la recherche d'images [73, 74, 48]. Néanmoins, les pyramides orientables présentent l'inconvénient d'être lourdement redondantes. En effet, le nombre de coefficients à la sortie du pyramide est égal à  $(\frac{4O}{3})$  fois le nombre de coefficients initial.

### 2.2.2 La transformée en ondelettes complexe en arbre dual

Pour remédier aux limitations de la 2D S-TOD, Kingsbury [75] a défini la transformée en ondelette 2D complexe en arbre dual.

En effet, Kingsbury a conçu au départ la transformée en ondelettes 2D complexe [75] permet de générer des coefficients détails complexes, cette transformations est alors 2 fois redondante. La transformée en ondelette 2D complexe en arbre dual (2D TOC-AD) est implémentée grâce à deux transformées en ondelettes 2D complexes opérant en parallèle appelées respectivement *arbre primaire* et *arbre dual*.

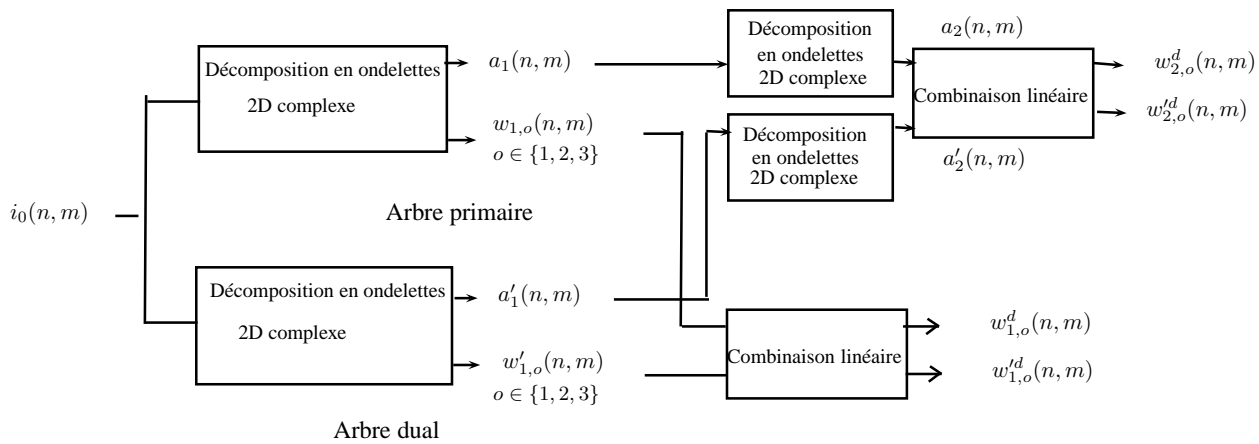


FIG. 2.7 – Transformée en ondelettes complexe en arbre dual à deux étages de décomposition.

Comme le montre la figure 2.7, la décomposition d'une image  $I(m, n)$  par les arbres primaire et dual fournit deux sous-images approximations  $a_1(m, n)$  et  $a'_1(m, n)$  à coefficients complexes ainsi que 3 sous-images de détails orientés horizontalement, diagonalement et verticalement à la sortie de chaque arbre :  $w_{1,o}(m, n)$ , respectivement,  $w'_{1,o}(m, n)$ ,  $o \in \{1, 2, 3\}$ . Kingsbury [75] a montré que les 6 sous-bandes  $w_{1,o}^d(m, n)$  et  $w_{1,o}'^d(m, n)$ ,  $o \in \{1, 2, 3\}$ , résultant de la combinaison linéaire des signaux  $w_{1,o}(m, n)$  et  $w'_{1,o}(m, n)$ , pour chaque orientation  $o \in \{1, 2, 3\}$ , sont localisés dans les 6 directions suivantes ( $\pm 15$ ,  $\pm 45$  et  $\pm 75$ ) degrés :

$$\begin{aligned} w_{1,o}^d(m, n) &= \frac{1}{\sqrt{2}}(w_{1,o}(m, n) + w'_{1,o}(m, n)) \\ w_{1,o}'^d(m, n) &= \frac{1}{\sqrt{2}}(w_{1,o}(m, n) - w'_{1,o}(m, n)). \end{aligned} \quad (2.23)$$

Ainsi, cette décomposition permet-elle une meilleure sélectivité directionnelle que la S-TOD au prix d'un facteur de redondance égal à 4. Elle a aussi l'avantage d'être invariante par translation. Comme le montre la figure 2.7, cette décomposition est itérée d'une manière récursive sur  $J$  étages. On se retrouve, ainsi, avec 2 sous-images approximation  $a_j(n, m)$  et  $a'_j(n, m)$  à l'étage  $J$  et 6 sous-images détail orientées dans six directions différentes ( $\pm 15$ ,  $\pm 45$  et  $\pm 75$ ) degrés à chaque étage  $j \in \{1, \dots, J\}$ .

La transformée en ondelette en arbre dual présente l'avantage d'être invariante par translation et d'être plus riche en information directionnelle que la 2D S-TOD puisqu'elle permet d'obtenir des détails dans six orientations différentes ce qui justifie son utilisation pour des applications comme la classification et la recherche d'images [76, 77, 78, 79].

### 2.2.3 La transformée en ondelette bidimensionnelle $M$ bandes en arbre dual

Comme le montre les transformées précédentes, réaliser un compromis entre un minimum de redondance et une richesse en information directionnelle consiste un véritable défi. C'est dans ce contexte que la transformée en ondelette 2D  $M$ -bandes en arbre dual (2D TOM-AD) a été introduite par Chaux *et. al* [80]. Elle a pour but de combiner les avantages des transformations  $M$  bandes à savoir la flexibilité du choix de filtres et la sélectivité en fréquence ainsi que la directionnalité et l'invariance par translation assurées par la décomposition en arbre dual. Comme le montre la figure 2.8, la 2D TOM-AD est basée sur la décomposition d'une image originale  $I(p, q)$ ,  $p, q \in \{1, \dots, N\}$  par deux transformations  $M$  bandes bidimensionnelles séparables associées respectivement en un arbre primaire et un arbre dual. Les bancs de filtres  $\tilde{h}_0, \dots, \tilde{h}_M$  et  $h'_0, \dots, h'_M$  associés respectivement à l'arbre primaire et à l'arbre dual forment des paires de Hilbert. La décomposition de l'image  $I(p, q)$  par les arbres primaire et dual résulte en deux images approximations  $a_1(p, q)$  et  $a'_1(p, q)$  ainsi que  $M - 1$  sous-bandes détail  $w_{1,m,m'}(p, q)$ , respectivement,  $w'_{1,m,m'}(p, q)$ ,  $m, m' \in \{1, \dots, M - 1\}$ ,  $(m, m') \neq (0, 0)$  à la sortie de chaque arbre. Pour obtenir des sous-bandes directionnelles, une combinaison linéaire des



signaux  $w_{1,m,m'}(p, q)$  et  $w'_{1,m,m'}(p, q)$  est effectuée de la manière suivante :

$$\begin{aligned} w_{1,m,m'}^d(p, q) &= \frac{1}{\sqrt{2}}(w_{1,m,m'}(p, q) + w'_{1,m,m'}(p, q)) \\ w_{1,m,m'}^d(p, q) &= \frac{1}{\sqrt{2}}(w_{1,m,m'}(p, q) - w'_{1,m,m'}(p, q)). \end{aligned} \quad (2.24)$$

Les sous-bandes résultantes  $w_{1,m,m'}^d$  et  $w'_{1,m,m'}^d$  sont orientées selon  $2(M^2 - 1)$  directions différentes.

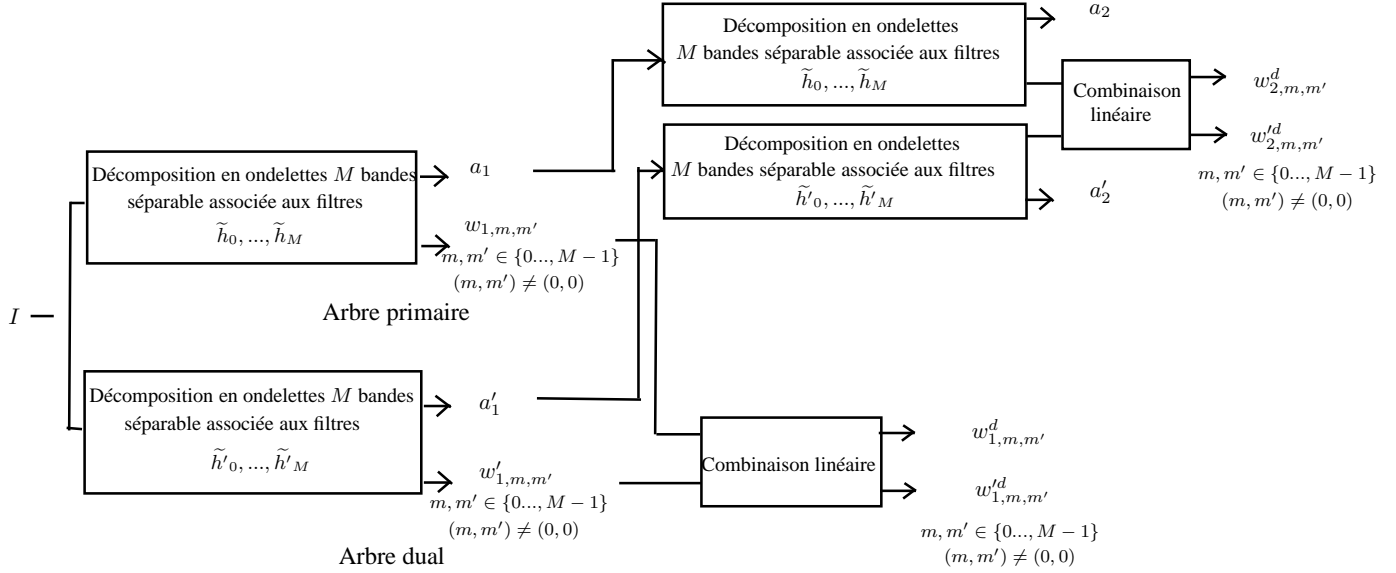


FIG. 2.8 – Transformée en ondelettes  $M$  bandes en arbre dual à deux étages de décomposition.

Comme le montre la figure 2.8, cette décomposition est itérée d'une manière récursive sur  $J$  étages. On obtient,  $2$  sous-images approximation  $a_J(n, m)$  et  $a'_J(n, m)$  et  $2(M^2 - 1)$  sous-images détail orientées dans  $2(M^2 - 1)$  directions différentes.

Si la 2D TOC-AD caractérise seulement 6 directions différentes, la 2D TOM-AD présente l'avantage de donner à l'utilisateur le libre choix du nombre de directions en ajustant le paramètre  $M$  et ceci, indépendamment du facteur de redondance qui demeure toujours égal à deux. Le tableau 2.1 résume les propriétés des différentes décompositions multirésolutions déjà citées. On remarque que la 2D TOM-AD réalise un meilleur compromis entre sélectivité directionnelle et facteur de redondance. Cette dernière a aussi l'avantage d'être invariante par translation. Enfin, elle présente l'intérêt d'offrir plus de liberté pour le choix des filtres par rapport à la 2D TOC-AD comme le fait d'utiliser des ondelettes à support compact orthogonales avec des filtres symétriques. A cet égard, elle a été utilisée avec succès en débruitage d'image [80], elle serait d'un intérêt considérable pour caractériser les images dans le contexte de recherche d'image.

TAB. 2.1 – Propriétés des différentes décompositions multirésolutions,  $O$  : nombre de filtres orientés,  $M$  : paramètre de la décomposition  $M$  bandes.

Transformation	Nombre d'orientations	Facteur de redondance
S-TOD	3	1
TOC-AD	6	2
Pyramides orientables	$O$	$4O/3$
TOM-AD	$2(M^2 - 1)$	2

## 2.3 Intérêt de la transformation en ondelettes en compression

### 2.3.1 Principe de la compression

De nos jours, des plateformes terrestres ou aériennes de télédétection ainsi que des satellites à haute résolution fournissent des images mono et multispectrales très volumineuses (une image SPOT5 monochromatique contient  $12000 \times 12000$  pixels codés sur 8 bits par pixel). Ces volumes élevés de données doivent satisfaire les contraintes d'espace de stockage et des bandes passantes des canaux de transmission. Pour tenir compte de ces contraintes, il n'y a qu'une solution : coder les images afin de réduire leur taille et ce, en réduisant la redondance des données d'une image. De nombreuses méthodes de compression sont envisageables [81] comme la transformée en cosinus discrète retenue dans la fameuse norme JPEG [82]. Cependant certaines applications nécessitent éventuellement un codage sans pertes de l'image comme l'imagerie médicale ou satellitaire et même une reconstruction progressive pour la téléconsultation. Pour satisfaire les exigences de ces applications, les transformations multirésolutions sont préconisées. Le principe de compression d'images dans le domaine des ondelettes repose sur le fait que la plupart des coefficients d'ondelettes sont généralement concentrés autour de 0. Ainsi, l'information utile est portée par un nombre limité de coefficients ce qui permet une compression plus efficace des données transformées : les coefficients non significatifs pour l'oeil humain sont alors éliminés ou reconstruits avec moins de précision [62, 83, 84]. Dans ce cadre, la parcimonie des représentations permet d'assurer un taux de compression élevé.

D'autre part, l'échelonnabilité assurée par une décomposition multirésolution est très utile dans le contexte de la transmission progressive largement utilisée pour des applications comme la téléconsultation ou la diffusion vers des destinataires connectés à des canaux ayant des largeurs de bande différentes. Dans ce dernier cas, l'émetteur envoie à chaque destinataire une partie de l'information codée correspondante au niveau de résolution satisfaisant à ses exigences et autorisé par la bande passante de son canal de transmission.

Enfin, certaines transformations multirésolutions permettent une reconstruction parfaite de

l'image transformée comme le schéma de lifting 5/3 [69]. Ces transformations sont alors exploitées pour le codage sans pertes [85, 86].

### 2.3.2 La norme JPEG2000

La norme JPEG2000 [87, 88] est une norme de compression d'images fixes qui utilise le codage par transformation, en particulier les transformations en ondelettes. La norme JPEG2000 a pour but de satisfaire des applications dont les exigences ne sont pas atteintes par les normes précédentes comme JPEG. En effet, la norme JPEG2000 comporte plusieurs fonctionnalités exigées par ces applications. Parmi les fonctionnalités les plus importantes, JPEG2000 assure :

- une meilleure performance en bas débit ;
- une compression avec et sans pertes, le mode de compression avec pertes utilise la transformation 9/7 [62], tandis que le mode sans pertes utilise la décomposition 5/3 [69] ;
- un codage évolutif échelonnable ;
- un codage par régions d'intérêt qui permet à l'utilisateur de définir certaines régions privilégiées pour l'application qui doivent être codées et transmises avec une meilleure qualité et moins de pertes que le reste de l'image ;
- une robustesse aux erreurs améliorée grâce à une correction efficace des erreurs.

En plus de la partie I qui est la partie principale de la norme JPEG2000, celle-ci envisage d'autres parties (II à VII). La partie II de la norme autorise le recours à n'importe quelle transformation multi-échelle. Quatre schémas de codage ont été fournis par la norme : la décomposition pyramidale, la décomposition SPACL de l' Université d'Arizona, la décomposition FBI (Federal Bureau of Investigation) et les paquets d'ondelettes [89].

## 2.4 Conclusion

Dans ce chapitre, nous avons rappelé les principes de la décomposition en ondelettes et l'analyse multirésolution. Nous avons présenté les décompositions multirésolutions les plus utilisées en traitement d'images ainsi que leur propriétés et nous avons fini par expliquer l'intérêt des transformations en ondelettes en compression d'images. Dans le chapitre suivant, nous proposons de présenter nos premières contributions dans le cadre d'indexation d'images multispectrales dans le domaine transformée en ondelettes.

## Chapitre 3

# Contributions à l'extraction de descripteurs d'images multispectrales

L'étape d'extraction des signatures est une partie primordiale dans notre travail. En effet, les performances d'un système de recherche d'images par le contenu sont fortement liées à l'habileté des descripteurs choisis à décrire le plus "fidèlement" possible le contenu de l'image. Dans ce chapitre, nous présentons nos deux principales contributions pour la description par le contenu d'images multispectrales à partir de représentations multirésolutions.

Pour ce faire, nous structurons ce chapitre comme suit. Dans la section 3.1, nous commençons par présenter un état de l'art des différents travaux d'extraction de signatures à partir d'images monochromes et multispectrales dans le domaine transformé en ondelettes. Nous focalisons sur leurs avantages ainsi que leurs limitations qui nous ont incités à concevoir une nouvelle stratégie pour définir des descripteurs. Plus précisément, notre première contribution qui fait l'objet de la section 3.2, consiste à établir des modèles statistiques aptes à refléter les propriétés statistiques des coefficients d'ondelettes d'images multicomposantes. Si dans un premier temps, nous nous sommes limités à une décomposition en ondelettes dyadique séparable (S-TOD) appliquée indépendamment à chaque composante spectrale, dans la section 3.4, en revanche, nous présentons notre seconde contribution qui consiste à utiliser des décompositions multirésolutions plus sophistiquées ayant des propriétés plus intéressantes que la S-TOD. Enfin, nous achevons ce chapitre en tirant quelques conclusions dans la section 3.5.

## 3.1 Signatures dans le domaine transformé en ondelettes

### 3.1.1 Motivations

Pour accéder par le contenu à une base d'images monochromes ou multispectrales, stockées sous une forme codée, une approche intuitive consiste à décoder les images de cette collection avant d'en extraire des signatures. Cette étape de décodage exige des ressources supplémentaires en terme de calcul et de mémoire de stockage. Pour éviter ce gaspillage de ressources, plusieurs travaux de recherche d'images par le contenu opérant directement dans le domaine codé ont été conçus. Ils ont pour principe d'extraire les vecteurs des signatures à partir de la représentation codée des images [90, 91, 92]. L'extraction de signatures à partir d'une représentation de l'image dans le domaine transformé en ondelettes se situe dans ce cadre. A ce niveau, éviter la décodage n'est pas la seule motivation. En effet, les transformations en ondelettes présentent l'avantage de générer des représentations parcimonieuses dans la mesure où elles concentrent le maximum d'énergie sur un faible nombre de coefficients à partir desquels les signatures sont calculées. La prise en compte de la compacité des représentations multirésolutions devrait réduire la complexité opératoire de l'étape d'extraction. Enfin, comme on l'a vu dans le chapitre 2, les représentations échelle-espace permettent une analyse du signal à différentes orientations et échelles et les signatures issues de ces coefficients permettraient une meilleure caractérisation du contenu visuel de l'image surtout dans le cas d'images texturées et ceci en assurant une analyse proche de celle réalisée par le système visuel humain [42, 93, 94].

### 3.1.2 Cas d'images monochromes

Dans le cas d'images monochromes, les descripteurs sont calculés directement à partir d'une représentation de l'image dans le domaine transformé en ondelettes de première génération [95, 96, 149, 46] ou de seconde génération [97]. En général, les signatures sont définies soit à partir du flux binaire du codage des coefficients, soit à partir des coefficients de la transformée en ondelettes considérée. Dans ce qui suit, nous allons passer brièvement en revue les principales méthodes rapportées dans la littérature.

#### Signatures issues du flux binaire des coefficients

Comme on l'a vu dans le chapitre 2, la majorité des coefficients d'ondelettes d'une image ont des valeurs proches de zéro et sont donc non significatifs. Cette propriété est exploitée dans le contexte de la compression d'images par plusieurs algorithmes de codage progressif tels que l'algorithme EZW (Embedded Zerotree Wavelet coding)[98], l'algorithme SPIHT (Set Partitioning In Hierarchical Trees) [99, 85], l'algorithme EBCOT [100] utilisé pour la norme de compression JPEG2000 et l'algorithme EZBC [101]. Ces algorithmes ont l'avantage

de produire des représentations de l'image échelonnables selon la résolution et/ou selon la précision. L'échelonnabilité en résolution permet de reconstruire progressivement l'image à des résolutions de plus en plus fines tandis que l'échelonnabilité en précision permet de décoder l'image à la même résolution spatiale mais avec des niveaux croissants de précision. De manière générale, les algorithmes de codage binaire des coefficients reposent sur le codage des coefficients d'ondelettes significatifs. C'est la raison pour laquelle l'histogramme des bits correspondants aux coefficients significatifs est souvent utilisé comme descripteur comme pour l'algorithme EZW [102], l'algorithme SPHIT [103] et l'algorithme ZTE [104].

### Signatures issues des coefficients d'ondelettes

Malgré l'existence de méthodes exploitant la structure spatiale des coefficients d'ondelettes [105], la stratégie prépondérante consiste à décrire les propriétés statistiques des coefficients d'ondelettes dans la mesure où ces propriétés peuvent facilement se traduire par des modèles. Dans ce cas, on peut distinguer 2 types de descripteurs : locaux et globaux.

– Descripteurs locaux :

Ils décrivent les propriétés statistiques locales des coefficients d'ondelettes, en d'autres termes, les relations spatiales entre les différentes valeurs des coefficients d'ondelettes. Dans ce cadre, le descripteur local le plus connu est la matrice de co-occurrence calculée dans le domaine transformée en ondelettes [149, 106]. Plusieurs descripteurs locaux peuvent être calculés à partir des éléments de la matrice de co-occurrence tels que l'entropie et l'énergie locale, l'homogénéité locale et le contraste [106].

– Descripteurs globaux :

Ils visent à décrire les propriétés statistiques globales des coefficients d'ondelettes tels que l'histogramme des coefficients d'ondelettes [149], la moyenne [106], l'énergie globale [149, 45, 44] et l'écart-type [149, 106]. Une autre solution est de choisir un modèle paramétrique de la distribution statistique des coefficients d'ondelettes. A ce niveau, il est d'usage de partitionner les coefficients en sous-ensembles (typiquement, des sous-bandes) et d'assigner un modèle de loi de probabilité pour chaque sous-ensemble. Une fois estimés, les paramètres de ce modèle sont considérés comme les signatures de ces sous-ensembles. Adopter un modèle statistique se justifie par plusieurs raisons dont les plus importantes sont les suivantes.

- La concentration des valeurs des coefficients d'ondelettes autour de zéro peut facilement se traduire par différents modèles statistiques appropriés.
- Ces modèles de distributions des coefficients restent valables pour la plupart des images naturelles ce qui n'est pas le cas pour la représentation spatiale.
- Les modèles préconisés sont associés à des paramètres faciles à estimer à partir des données.

Notons que le choix du modèle est guidé par deux facteurs importants : l’aptitude à refléter les propriétés statistiques du sous-ensemble de coefficients considéré et les propriétés intrinsèques des paramètres extraits dans le contexte de la description par le contenu. En effet, le modèle paramétrique doit être approprié dans le sens où il doit traduire les propriétés statistiques des coefficients d’ondelettes. Ainsi, est-il possible de modéliser les coefficients d’ondelettes à chaque étage et dans chaque orientation. A cet égard, plusieurs modèles statistiques des sous-bandes d’ondelettes ont été proposés dans la littérature. Les plus connus sont la Distribution Gaussienne Généralisée (DGG) [149, 46, 107], la Distribution Gamma (DG) [47] et le mélange gaussien [108]. Un autre type de modélisation peut être envisagé pour prendre en compte la dépendance qui existe entre les coefficients entre des échelles successives dans une orientation donnée [98]. Dans ce cas, certains travaux font intervenir une décorrélation des coefficients d’ondelettes à différentes échelles en utilisant une méthode de séparation de sources [109]. D’autres utilisent des modèles plus compliqués pour modéliser ces dépendances tels que les modèles de Markov [48] ou les modèles hiérarchiques de corrélation [110].

Il est également désirable que les paramètres du modèle choisi, qui représentent les signatures de l’image, soient invariants par rapport à certaines “perturbations” de l’image telles que la translation, la rotation ou le changement d’échelle. Pour assurer l’invariance par rapport à ces “perturbations”, un degré de liberté supplémentaire s’offre par rapport à l’indexation dans le domaine spatial et réside dans le recours à d’autres décompositions multirésolutions capables de générer des coefficients multirésolutions d’une image “perturbée” se déduisant facilement de ceux de l’image originale. Par exemple, dans le souci d’assurer une invariance par rotation, il est possible d’appliquer des décompositions multirésolutions ayant une plus grande sélectivité directionnelle que la transformée en ondelettes dyadique comme les pyramides orientables [70] utilisées en indexation dans [72, 48] et la transformée en ondelettes complexe en arbre dual [75] adoptée en indexation dans [113, 77].

## Commentaires

A ce niveau, il s’avère utile de discuter des avantages et inconvénients des deux principales stratégies de calculs des descripteurs. En effet, la première approche d’extraction des signatures directement à partir du flux binaire permet d’éviter le décodage du flux binaire avant d’en extraire les descripteurs. Elle a aussi pour avantage de ne raisonner que sur les bits correspondants aux coefficients les plus significatifs ce qui réduit la complexité opératoire. Toutefois, dans le cas de codage avec pertes, l’extraction de signatures directement sur les coefficients d’ondelettes (avant leur quantification) a pour avantage d’éviter la perte éventuelle d’information. C’est cette approche que nous retiendrons dans le cadre de la thèse. Dans ce

cas, les méthodes globales et locales se complètent. Si les dernières ont l'inconvénient de ne donner qu'une idée statistique globale sur la distribution des coefficients d'ondelettes dans une sous-bande donnée, elles ont l'avantage d'être plus faciles à mettre en œuvre, d'engendrer moins de complexité opératoire et sont plus adaptées à l'extraction de signatures invariantes par rapport à certaines modifications de l'image telles que la rotation et la translation.

Dans ce qui suit, nous allons dresser un bref panorama des techniques d'extraction de descripteurs dans les images multispectrales.

### 3.1.3 Cas d'images multispectrales

Une image multispectrale est une image contenant différents canaux spectraux correspondants à des données à différentes longueurs d'ondes du spectre électromagnétique. Les images couleurs constituent un cas particulier d'images multispectrales sur 3 canaux appartenant au spectre visible (rouge, vert, bleu) caractérisés par des longueurs d'onde entre  $0.4\mu\text{m}$  et  $0.7\mu\text{m}$ . Les images multispectrales peuvent aussi contenir des canaux dans le spectre de l'infrarouge caractérisé par des longueurs au-delà de  $0.7\mu\text{m}$  et invisibles pour l'être humain. Les images multispectrales ont été essentiellement développées en imagerie satellitaire optique, domaine dans lequel se situe notre recherche mais de nos jours, elles sont également présentes dans plusieurs applications telles que l'imagerie médicale IRM [114]. D'énormes bases d'images satellitaires telles que les archives d'images de satellites SPOT et la base d'images Microsoft TerraServer [8] contiennent des millions d'images satellitaires multispectrales produites quotidiennement par des satellites à haute résolution tels que IKONOS, SPOT5, ALOS, GeoEye-1, WorldView-2 [3] et TerraSAR-X [4]. Ces bases sont régulièrement consultées par les scientifiques et experts en imagerie. Ainsi, est-il nécessaire de mettre en place des systèmes de recherche d'images par le contenu permettant de rechercher des informations pertinentes à partir de ces bases. Dans ce qui suit, nous nous intéressons à l'extraction de descripteurs à partir de ces images multispectrales. Dans la littérature, les travaux d'extraction de signatures dans le cas de ces images peuvent être classés en deux familles principales selon que l'on traite séparément ou conjointement les composantes spectrales. On parlera respectivement d'approche univariée ou multivariée.

#### Approche univariée

Pour extraire des signatures à partir d'images multispectrales, l'approche la plus intuitive est l'approche univariée qui consiste à extraire des descripteurs à partir des coefficients d'ondelettes relatifs à chaque composante (ou canal) spectral indépendamment des autres canaux [115]. Lors de la procédure de recherche, les vecteurs des signatures des différents canaux de l'image requête sont comparés à leurs homologues associés aux images de la base comme le montre le schéma de la figure 3.1. Dans certains travaux, le vecteur des signatures représente



la moyenne pondérée des vecteurs descripteurs calculés sur toutes les composantes. Cette approche a l'avantage d'être simple à mettre en œuvre ainsi que de permettre le recours à des méthodes d'extraction des signatures dans le cas d'images monochromes, citées précédemment. Néanmoins, l'approche univariée présente l'inconvénient de prendre seulement en considération l'information relative à chaque canal, tout en ignorant la dépendance qui existe réellement entre les différentes composantes (dépendance inter-canaux). Pour remédier à ce problème, une approche multivariée est recommandée.

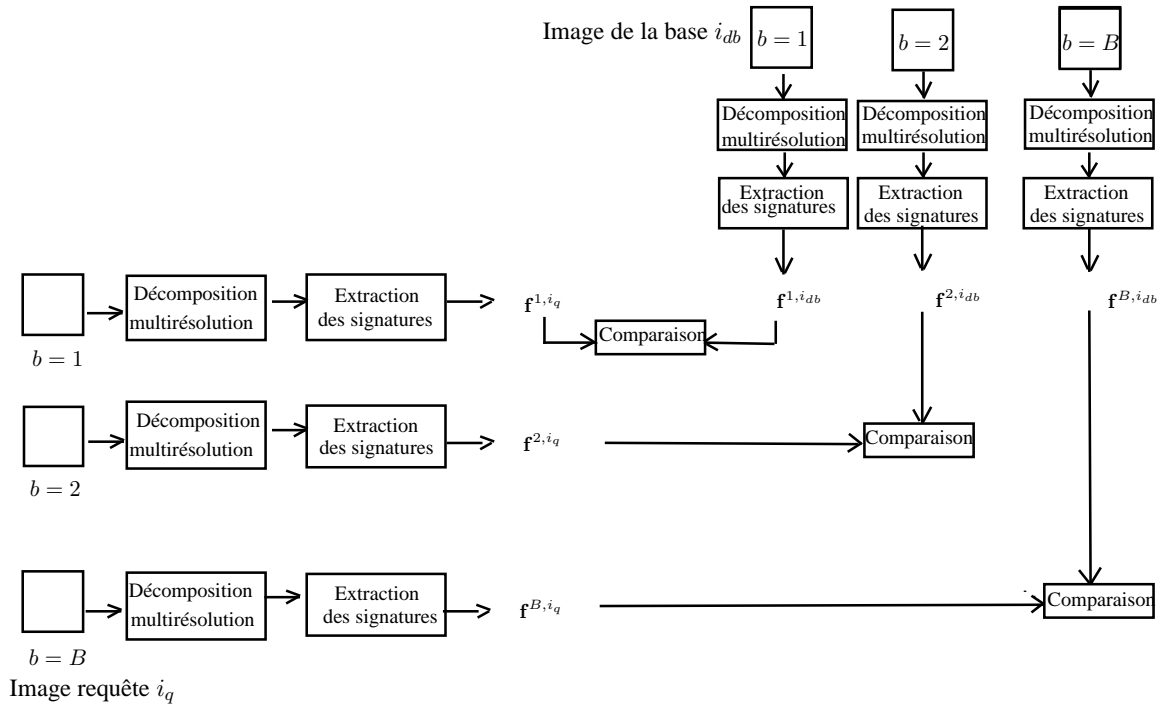


FIG. 3.1 – Approche univariée pour l'indexation d'images multispectrales codées dans le domaine des ondelettes.

### Approche multivariée

Dans la pratique, une dépendance (parfois forte selon les images) existe entre les différentes composantes d'une image multispectrale et éventuellement entre les coefficients d'ondelettes correspondants. En effet, les différents canaux d'une image multispectrale correspondent à l'acquisition de la même scène par des capteurs opérant à différentes longueurs d'ondes du spectre électromagnétique. Ainsi, les coefficients d'ondelettes des différentes composantes devraient avoir des propriétés statistiques communes et présenter une dépendance statistique. Notons que cette dépendance a été en effet exploitée dans plusieurs domaines du traitement d'images multispectrales. C'est le cas de la détection et l'extraction de données à partir d'images LANDSAT et SPOT4 [116], pour la compression d'images [117, 118], le débruitage [119]. En revanche, dans le contexte de la recherche par le contenu et au moment du démarrage

de nos travaux de thèse début 2006, peu de travaux ont été consacrés à la prise en compte de la dépendance inter-spectrale lors de l'extraction des signatures dans le domaine transformé en ondelettes. Sauf erreur de notre part, il s'agit de travaux sur la recherche d'images SPOT [120]. Toutefois, ces deux dernières années, quelques travaux ont été entrepris dans ce sens comme nous allons le voir dans le bref état de l'art qui suit. Schématiquement, ces travaux se divisent en 2 catégories.

– Réduction des dépendances inter-spectrales

Cette approche a pour but de réduire les dépendances inter-canaux via un pré-traitement des composantes de coefficients d'ondelettes puis d'adopter une stratégie univariée. A cet égard, il est possible de recourir à une méthode de séparation de sources (analyse en composantes indépendantes) ou à une décorrélation par analyse en composantes principales avant d'extraire les signatures à partir de chaque composante résultante selon une stratégie univariée [109, 121]. En plus de sa simplicité, cette méthode présente l'avantage de pouvoir appliquer n'importe quelle méthode d'indexation univariée citée précédemment. Néanmoins, les algorithmes de séparation de sources considèrent que les sources sont linéairement mélangées ce qui n'est pas toujours le cas pour les coefficients d'ondelettes relatifs à différents canaux. L'imprécision qui en résulte a été la motivation pour l'approche suivante.

– Prise en compte directe des dépendances inter-spectrales

Il s'agit d'éviter l'étape de prétraitement et de tenir pleinement compte de cette dépendance. Un premier travail dans ce sens a été proposé par Newsam et Kamath dans le cas d'images SPOT [115]. Plus précisément, l'image de l'indice de végétation est calculée en combinant les composantes rouge et proche infrarouge puis cette image résultante est décomposée par une transformation en ondelettes afin d'extraire des descripteurs locaux de ses coefficients (comme ceux issus de la matrice de co-occurrence de chaque sous-bande). Une seconde alternative consiste à choisir un modèle multivarié pour modéliser la distribution conjointe du vecteur des coefficients d'ondelettes des sous-bandes à la même échelle et orientation dominante de *toutes* les composantes spectrales. A notre connaissance, le premier travail entrepris dans cette optique [120] a consisté à choisir la Distribution Gaussienne Généralisée Multivariée (DGGM) comme modèle conjoint des coefficients d'ondelettes des différentes composantes d'images SPOT3 et SPOT4. Récemment, un regain d'intérêt s'est manifesté pour la DGGM dans [122] mais dans le cas d'images couleurs. Rappelons que la densité de probabilité d'une DGGM pour des vecteurs de  $B$  composantes est donnée par :

$$\forall \mathbf{x} \in \mathbb{R}^B \quad f(\mathbf{x}) = \frac{\Gamma(B/2)}{\pi^{B/2} \Gamma(B/2s) 2^{B/2s}} \frac{s}{|\mathbf{R}|^{1/2}} \times \exp(-1/2 \mathbf{x} \mathbf{R}^{-1} \mathbf{x}^t) \quad (3.1)$$

où  $\mathbf{R}$  est la matrice de covariance du vecteur  $\mathbf{x}$  et  $\Gamma(\cdot)$  est la fonction gamma. Ce modèle possède l'avantage de refléter certaines propriétés statistiques des coefficients d'onde-

lettres à savoir l'unimodalité de la distribution empirique ainsi que la concentration des valeurs autour du vecteur nul. En plus les paramètres  $\mathbf{R}$  et  $s$  peuvent être estimés en utilisant respectivement la méthode des moments.

Ce bref état de l'art indique que plusieurs pistes peuvent être explorées dans le cadre de cette thèse.

- En effet, si dans un premier temps la DGGM a été considérée, il importe de voir si d'autres modèles de loi conjointe peuvent être utilisés et s'ils peuvent surpasser la DGGM dans le cadre de notre application.
- Si concernant les images monochromes, les dépendances inter-échelle des coefficients ont été exploités, il n'en reste pas moins que cet aspect n'a pas été pris en compte dans le cas d'images multispectrales.
- Il en est de même pour l'invariance des signatures par rapport à certaines "perturbations" des images multicomposantes.

A notre connaissance, il n'y pas de travaux qui ont simultanément pris en considération ces trois axes. Dans ce travail, nous visons à extraire à partir des coefficients d'ondelettes d'images multispectrales, des signatures satisfaisant ces trois conditions séparément (dans un souci de clarté de la présentation) puis d'une manière simultanée. Plus précisément, comme le montrent les deux étapes indiquées dans la figure 3.2, nos contributions résident dans :

- la conception de nouveaux modèles des coefficients d'ondelettes basés sur la théorie des copules dans le cas d'une décomposition en ondelettes conventionnelle appliquée à chaque composante spectrale ;
- la généralisation de ces modèles pour d'autres décompositions multirésolutions appliquées conjointement aux composantes.

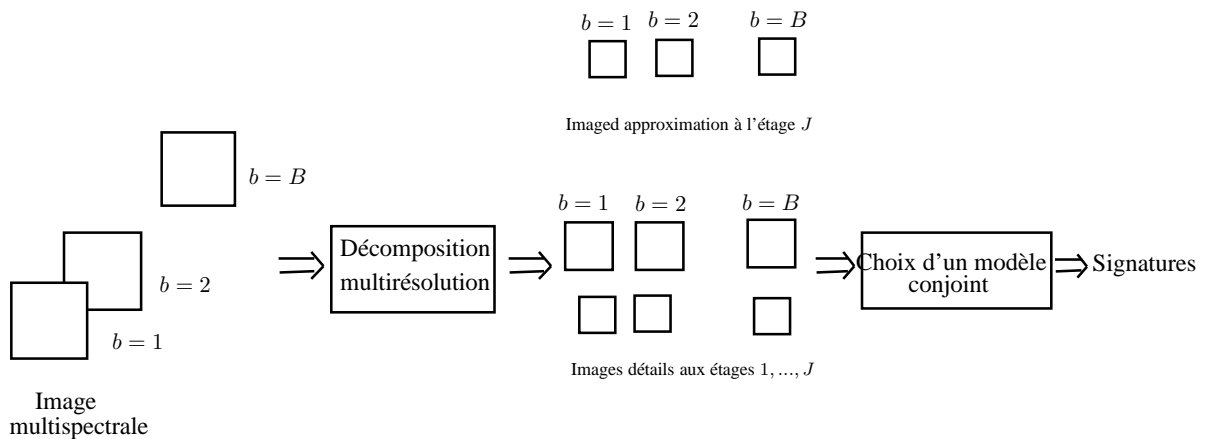


FIG. 3.2 – Différentes parties à considérer dans l'étape d'extraction des signatures.

## 3.2 Contribution 1 : Modèle multivarié inter-spectral

Dans toute cette section, nous adoptons comme décomposition, la décomposition en ondelettes de première ou de seconde génération appliquée séparément à chaque composante d'une image multispectrale  $(I_0^{(b)})_{b=1}^B$  ayant  $B$  composantes ( $B \in \mathbb{N}^*$ ). Pour une meilleure clarté de la description de notre contribution, nous allons ne plus mentionner l'indice de la direction  $o$ . Autrement dit, un coefficient d'ondelettes  $w_{j,o}^{(b)}(m, n)$  sera noté  $w_j^{(b)}(m, n)$ .

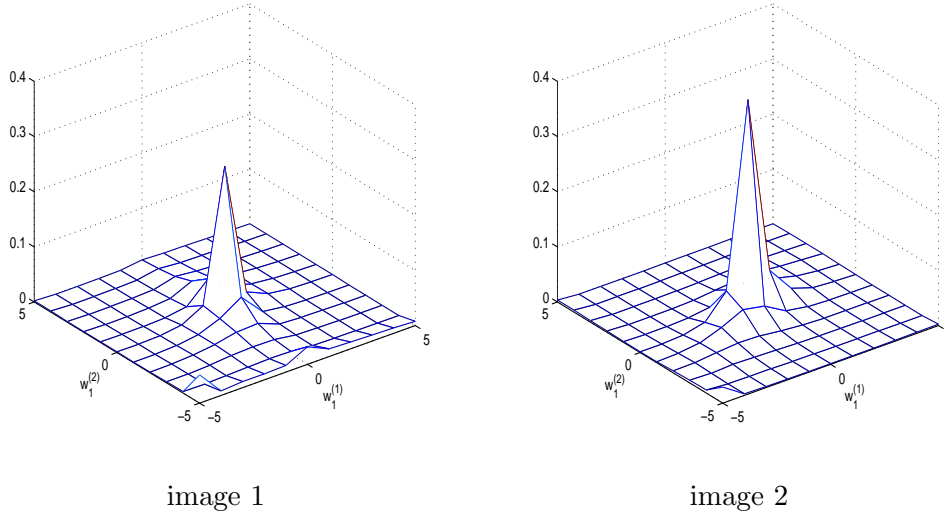


FIG. 3.3 – Histogrammes conjoints empiriques des coefficients d'ondelettes des deux premières composantes de deux images SPOT4 différentes, pour  $j = 1$ .

### 3.2.1 Motivations

Comme on l'a vu dans la section précédente, l'une des méthodes pour l'extraction de signatures à partir de la représentation d'une image multispectrale dans le domaine des ondelettes repose sur la modélisation de la distribution conjointe des coefficients multirésolutions  $w_j^{(b)}$  obtenus pour toutes les composantes spectrales  $b = 1, \dots, B$ . En d'autres termes, il s'agit de trouver un modèle de la densité conjointe  $f_{\mathbf{w}_j}$  du vecteur  $\mathbf{w}_j$  défini par :

$$\mathbf{w}_j(m, n) \triangleq (w_j^{(1)}(m, n), \dots, w_j^{(B)}(m, n))^{\top} \quad (3.2)$$

et qui traduit le plus "fidèlement" possible les propriétés statistiques du vecteur  $\mathbf{w}_j(m, n)$  y compris la dépendance importante qui existe entre les différentes composantes  $w_j^{(b)}$ . La figure 3.3 montre les histogrammes empiriques conjoints des coefficients d'ondelettes issus d'un étage de décomposition des deux premières composantes de deux images SPOT4 différentes. On remarque que ces histogrammes peuvent prendre des formes assez complexes et variables d'un image à l'autre. Ainsi leurs approximation par des distributions DGGM illustrées à la

figure 3.4 peut s'avérer inexacte pour refléter les propriétés statistiques des coefficients d'ondelettes. Pour pallier cette insuffisance, nous allons recourir à des modèles s'appuyant sur la théorie des copules que nous rappelons dans ce qui suit.

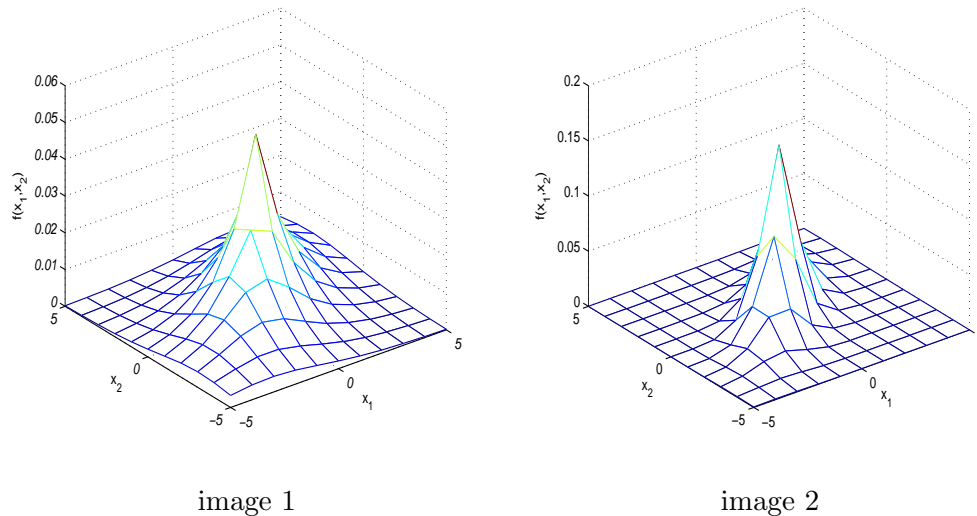


FIG. 3.4 – Densités gaussiennes généralisées bivariées ( $B = 2$ ) de paramètres respectifs  $\rho_1 = 0.77$ ,  $s_1 = 0.54$  et  $\rho_2 = 0.8$ ,  $s_2 = 0.73$  modélisant les distributions conjointes des coefficients d'ondelettes des deux premières composantes des deux images SPOT4 précédentes, pour  $j = 1$ , où  $\rho_1$ , (resp.  $\rho_2$ ) est le coefficient de corrélation linéaire entre les deux composantes de l'image 1 (resp. l'image 2).

### 3.2.2 Bref rappel sur la théorie des copules

#### Importance de la théorie des copules

Dans la pratique, l'étude de différents systèmes de nature variée (biologiques, chimiques, financiers, commerciaux, sociaux) indique l'existence de relations de dépendance entre les divers phénomènes qui les régissent. Dans le domaine du traitement d'images, les dépendances ne manquent pas : dépendances entre pixels d'une même image, entre canaux d'une image multispectrale, entre les images d'une même scène prises à des instants différents. Ces dépendances peuvent être judicieusement exploitées pour des applications comme la compression, la classification d'images, l'imagerie 3D et la recherche d'images. Les phénomènes de dépendance sont généralement formalisés mathématiquement par des variables aléatoires dépendantes entre elles. La covariance linéaire est l'une des mesures de dépendance les plus connues en statistique. Néanmoins, celle-ci ne donne qu'une idée globale sur la similarité entre variables aléatoires sans décrire la structure complète de dépendance entre elles. En plus, elle

est limitée au cas de couple de 2 variables aléatoires (covariance entre 2 variables aléatoires). La théorie des copules [123] (*copula* en latin signifie lien, chaîne ou union) est une théorie ayant pour but de décrire la totalité de l'information concernant la dépendance entre plusieurs variables aléatoires et ce, indépendamment de leurs propres lois marginales [124]. Dans ce qui suit, nous présentons un bref rappel sur la théorie des copules.

### Définition d'une copule

Notons par  $B$  un entier positif ( $B > 1$ ). Une copule  $C$  est définie comme étant la fonction de répartition conjointe de  $B$  variables aléatoires  $X_1 \dots X_B$  ayant des distributions marginales uniformes dans l'intervalle  $[0, 1]$ . Lorsque la copule  $C$  est continue et différentiable, on peut définir la densité de copule  $c$  par :

$$\forall (x_1, \dots, x_B) \in \mathbb{R}^B, \quad c(x_1, \dots, x_B) \triangleq \frac{\partial^B C(x_1, \dots, x_B)}{\partial x_1 \dots \partial x_B}. \quad (3.3)$$

Notons que d'après cette définition, une copule  $C$  décrit les dépendances mutuelles du  $B$ -uplet  $(X_1 \dots X_B)$  et par conséquent, il existe une grande variété d'expressions de copules qui traduisent un grand nombre de structures de dépendance différentes. Cette définition est utile également dans le cas de  $B$ -uplets quelconques de variables aléatoires  $(X_1 \dots X_B)$  n'ayant pas forcément des lois marginales uniformes. Cela se traduit mathématiquement par le théorème suivant dû à Sklar [125].

### Théorème de Sklar

Soient  $X_1, \dots, X_B$ ,  $B$  variables aléatoires quelconques de fonction de distribution conjointe  $F$  et admettant respectivement  $F_{X_1}, \dots, F_{X_B}$  comme fonctions de répartition marginales. Il existe une unique copule  $C$  telle que :

$$\forall (x_1, \dots, x_B) \in \mathbb{R}^B \quad F(x_1, \dots, x_B) = C(F_{X_1}(x_1), \dots, F_{X_B}(x_B)). \quad (3.4)$$

Dans cette équation, la copule  $C$  est unique et indépendante des distributions marginales. Elle reflète la dépendance mutuelle entre les variables marginales  $X_1, \dots, X_B$ .

Dans le cas où  $C$  est continue et différentiable, on peut également exprimer la fonction de densité de probabilité conjointe  $f$  en fonction des densités marginales  $f_{X_1}, \dots, f_{X_B}$  et la densité de copule  $c$  :

$$\forall (x_1, \dots, x_B) \in \mathbb{R}^B \quad f(x_1, \dots, x_B) = c(F_{X_1}(x_1), \dots, F_{X_B}(x_B)) \prod_{b=1}^B f_{X_b}(x_b). \quad (3.5)$$

L'importance de ce théorème réside dans les points suivants.

- La copule  $C$  ou sa densité  $c$  traduit d'une manière *complète et unique* la structure de dépendance mutuelle qui existe entre les variables aléatoires  $X_1, \dots, X_B$  contrairement à d'autres mesures de dépendance comme la covariance qui est restreinte à la dépendance linéaire entre deux variables aléatoires.
- La copule  $C$  traduit la dépendance entre n'importe quel nombre de variable aléatoires ( $B > 2$ ). Elle n'est pas limité au cas bivarié ( $B = 2$ ).
- Contrairement à d'autres distributions multivariées telles que la DGGM, le théorème de Sklar exprime séparément les distributions marginales des variables  $X_1, \dots, X_B$  et les dépendances entre elles traduites par la copule. En plus, la copule  $C$  est complètement indépendante des distributions marginales. Ainsi, l'estimation des distributions marginales et celle de la copule peuvent-elles se faire séparément.

Grâce au théorème de Sklar, la théorie des copules a été introduite avec succès dans diverses applications, d'abord, dans le domaine des finances à la fin des années 90 pour l'évaluation du risque financier [126, 127]. Ces dernières années, les copules commencent à être utilisées dans le domaine du traitement d'images. Dans [128], les copules permettent de modéliser la dépendance entre les coefficients d'ondelettes relatifs à différents canaux spectraux pour la *fusion* et la *restauration* d'images. D'autre part, les copules sont également employées dans [129] pour la *classification* d'images multispectrales. Tout récemment, en 2009, les copules sont utilisées dans [130] pour la conception d'un modèle multivarié traduisant la distribution de coefficients d'ondelettes correspondants à différents étages et différentes orientations et ceci pour la *classification* d'images texturées monochromatiques. Les copules sont également exploitées dans [131] pour le *recalage* d'images dans la mesure où elles permettent la modélisation de la dépendance entre images prises à différents points de vue et/ou par différents capteurs. Enfin, on a recours aux copules pour mesurer l'information mutuelle entre images hyperspectrales [132] ou radar [133]. Notons qu'au moment du démarrage de cette thèse, peu d'applications en traitement d'images ont fait emploi des copules. L'originalité de notre démarche repose sur le recours à cet outil pour prendre en compte les dépendances entre coefficients.

### 3.2.3 Principes des modèles proposés

Dans cette partie, on considère les coefficients d'ondelettes à un même étage de décomposition donné  $j \in \{1, \dots, J\}$ . En particulier, on cherche à modéliser la distribution conjointe du vecteur  $\mathbf{w}_j(m, n) = (w_j^{(1)}(m, n), \dots, w_j^{(B)}(m, n))^\top$  formé par les coefficients d'ondelettes relatifs à toutes les composantes  $b = 1, \dots, B$  à un étage donné  $j$ . Pour ceci, on a recours au théorème de Sklar. Ainsi la distribution conjointe  $F_{\mathbf{w}_j}$  peut-elle s'écrire de la manière suivante :

$$\forall (x_1, \dots, x_B) \in \mathbb{R}^D \quad F_{\mathbf{w}_j}(x_1, \dots, x_B) = C_j(F_{\mathbf{w}_j^{(1)}}(x_1), \dots, F_{\mathbf{w}_j^{(B)}}(x_B)) \quad (3.6)$$

où  $F_{\mathbf{w}_j^{(b)}}$  représentent les distributions marginales des coefficients d'ondelettes de chaque composante  $b$  et  $C_j$  la copule traduisant leur dépendance mutuelle. De la même manière, dans le cas où  $C_j$  est continue et différentiable, la densité de probabilité conjointe  $f_{\mathbf{w}_j}$  peut s'écrire sous la forme suivante :

$$f_{\mathbf{w}_j}(x_1, \dots, x_B) = f_{w_j^{(1)}}(x_1) \dots f_{w_j^{(B)}}(x_B) \times c_j(F_{w_j^{(1)}}(x_1), \dots, F_{w_j^{(B)}}(x_B)), \quad (3.7)$$

où  $f_{\mathbf{w}_j^{(b)}}$  représente la densité de probabilité marginale des coefficients d'ondelettes de la composante  $b$  et  $c_j$  la densité de la copule  $C_j$  à l'échelle  $j$ .

D'après ces équations, il est facile de déduire que modéliser la distribution conjointe  $F_{\mathbf{w}_j}$  (ou la densité conjointe  $f_{\mathbf{w}_j}$ ) du vecteur  $\mathbf{w}_j$  revient simplement à trouver un modèle paramétrique pour les distributions marginales  $F_{w_j^{(b)}}$  et une copule  $C_j$ .

Le vecteur des signatures de l'image correspondant à un étage de décomposition  $j$  est formé par la concaténation des paramètres des distributions marginales et ceux de la copule choisie.

Le défi est de choisir des modèles paramétriques qui satisfont les conditions suivantes.

- Les modèles doivent traduire le plus "fidèlement" possible les distributions réelles des coefficients.
- Les paramètres des modèles doivent être facilement estimés à partir des coefficients observés.

Dans ce qui suit, nous allons présenter les modèles que nous avons préconisés pour  $f_{w_j^{(b)}}$  et  $C_j$ .

### 3.2.4 Choix des distributions marginales

Comme le montre la figure 3.5, les histogrammes des lois marginales des coefficients d'ondelettes de chaque composante  $b$  sont monomodaux et localisés autour de 0.

La figure 3.6 montre également que les histogrammes des lois marginales des valeurs absolues des coefficients d'ondelettes de chaque composante  $b$  sont localisés autour de 0.

Ainsi pour modéliser les densités de probabilité marginales  $f_{w_j^{(b)}}$  des coefficients détail à une composante  $b$  et un étage  $j$ , deux distributions sont envisagées dans ce travail : la Distribution Gaussienne Généralisée Univariée (DGG) et la Distribution Gamma (DG) :

- La Distribution Gaussienne Généralisée (DGG) :

Elle est caractérisée par la fonction de densité de probabilité  $f_{w_j^{(b)}}^{\text{DGG}}$  définie pour tout  $x \in \mathbb{R}$  par :

$$f_{w_j^{(b)}}^{\text{DGG}}(x) = \frac{\beta_j^{(b)}}{2\alpha_j^{(b)}\Gamma(1/\beta_j^{(b)})} e^{-(|x|/\alpha_j^{(b)})^{\beta_j^{(b)}}} \quad (3.8)$$

où  $\Gamma(z) \triangleq \int_0^{+\infty} t^{z-1} e^{-t} dt$ ,  $\alpha_j^{(b)}$  et  $\beta_j^{(b)}$  correspondent respectivement aux paramètres d'échelle et de forme. Le paramètre  $\alpha_j^{(b)}$  reflète la largeur du sommet de la distribution



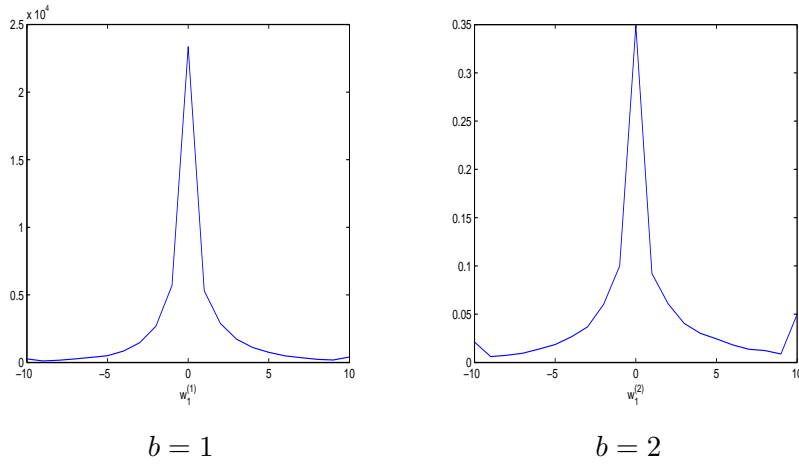


FIG. 3.5 – Histogrammes empiriques des coefficients d’ondelettes des 2 premières composantes d’une image SPOT4 pour  $j = 1$ .

en traduisant l’écart-type (plus  $\alpha_j^{(b)}$  est grand plus le sommet est large). Le paramètre  $\beta_j^{(b)}$  est inversement proportionnel à la vitesse de décroissance des arêtes du pic. La figure 3.7 montre les densités gaussiennes généralisées modélisant les histogrammes des coefficients d’ondelettes représentés par la figure 3.5. On remarque que cette loi reflète le caractère monomodal des histogrammes empiriques des  $B$  composantes considérées séparément.

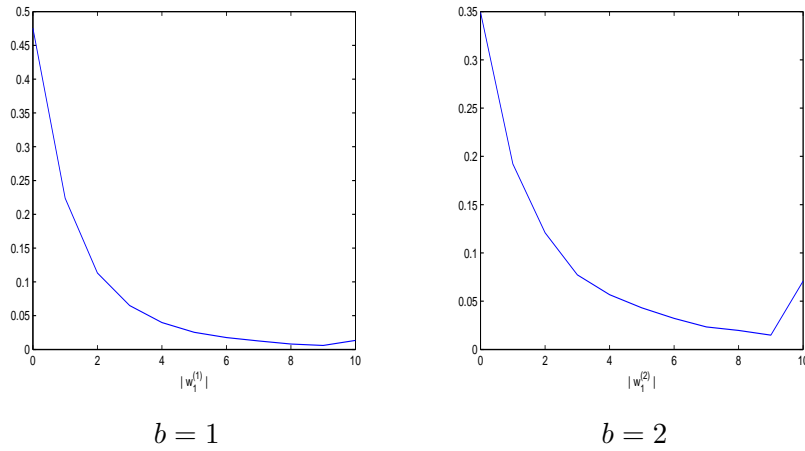


FIG. 3.6 – Histogrammes empiriques des valeurs absolues des coefficients d’ondelettes des 2 premières composantes d’une image SPOT4 pour  $j = 1$ .

– La Distribution gamma (DG) :

Elle caractérise plutôt la distribution des valeurs absolues des coefficients d’ondelettes  $w_j^{(b)}(m, n)$ , elle correspond à la fonction de densité de probabilité  $f_{w_j^{(b)}}^{\text{DG}}$  définie pour tout  $x \in \mathbb{R}^+$  par :

$$f_{w_j^{(b)}}^{\text{DG}}(x) = x^{\gamma_j^{(b)}-1} \frac{e^{-\delta_j^{(b)}}}{\delta_j^{(b)\gamma_j^{(b)}} \Gamma(\gamma_j^{(b)})} \quad (3.9)$$

où les paramètres  $\gamma_j^{(b)}$  et  $\delta_j^{(b)}$  correspondent respectivement aux paramètres d’échelle et de forme. La figure 3.8 montre les densités gamma modélisant les histogrammes des coefficients d’ondelettes représentés par la figure 3.6. Cette loi présente, en effet, un pic proche de zéro, ce qui reflète le fait que la plupart des valeurs absolues des coefficients d’ondelettes sont faibles.

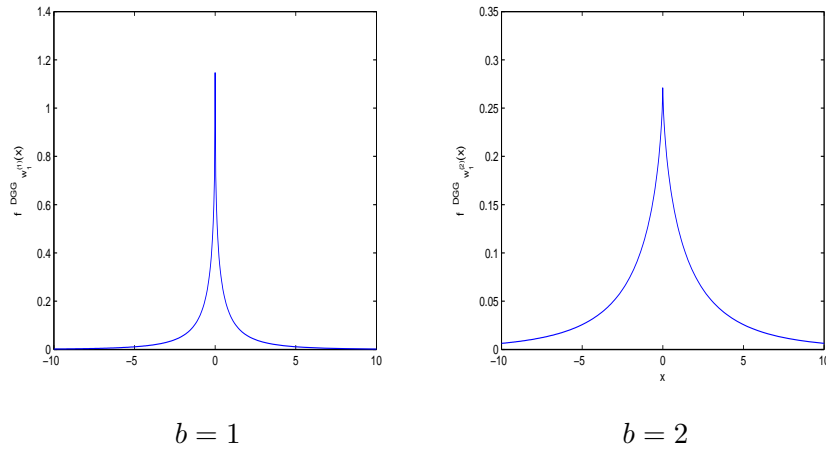


FIG. 3.7 – Densités gaussiennes généralisées de paramètres respectifs  $\beta_1 = 0.19$ ,  $\alpha_1 = 0.47$  et  $\beta_2 = 0.15$ ,  $\alpha_2 = 0.45$  modélisant les coefficients d’ondelettes horizontales des 2 premières composantes de l’image SPOT4 précédente pour  $j = 1$ .

Dans les deux cas (DGG et DG), les paramètres  $\alpha_j^{(b)}$  et  $\beta_j^{(b)}$  ou  $\gamma_j^{(b)}$  et  $\delta_j^{(b)}$  sont estimés grâce à une estimation au sens du maximum de vraisemblance.

### 3.2.5 Familles de copules proposées

Dans la littérature, un vaste choix de familles de copules paramétriques se présente. Ainsi le choix de la meilleure famille de copules qui modélise “le mieux” la dépendance inter-canal est un véritable défi.

– La copule gaussienne :

Vu sa simplicité à mettre en œuvre, la première intuition [134, 135] que nous avons eue pour modéliser les dépendances inter-canal est d’avoir recours à la copule gaussienne qui fait partie des copules elliptiques et qui est donnée par :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in [0, 1]^B, \quad C_j(x_1, \dots, x_B) = \varphi_{\mathbf{R}_j}(\phi^{-1}(x_1), \dots, \phi^{-1}(x_B)) \quad (3.10)$$

où  $\mathbf{R}_j$  est la matrice de corrélation associé du vecteur  $\mathbf{w}_j$ ,  $\varphi_{\mathbf{R}_j}$  est la fonction de répartition de la gaussienne  $B$ -variée  $\mathcal{N}(\mathbf{0}, \mathbf{R}_j)$  et  $\phi^{-1}$  est la fonction de répartition inverse de la distribution normale univariée  $\mathcal{N}(0, 1)$ . La densité de copule gaussienne est exprimée de la manière suivante :

$$\forall \mathbf{x} = (x_1, \dots, x_B)^T \in \mathbb{R}_j^B, \quad c_j(\mathbf{x}) = |\mathbf{R}_j|^{-1/2} \exp\left[-\frac{\check{\mathbf{x}}^T (\mathbf{R}_j^{-1} - \mathbf{I}) \check{\mathbf{x}}}{2}\right] \quad (3.11)$$

où  $\check{\mathbf{x}} = (\phi^{-1}(x_1), \dots, \phi^{-1}(x_B))$ .

En plus de sa simplicité, l’avantage de choisir une copule gaussienne consiste à estimer

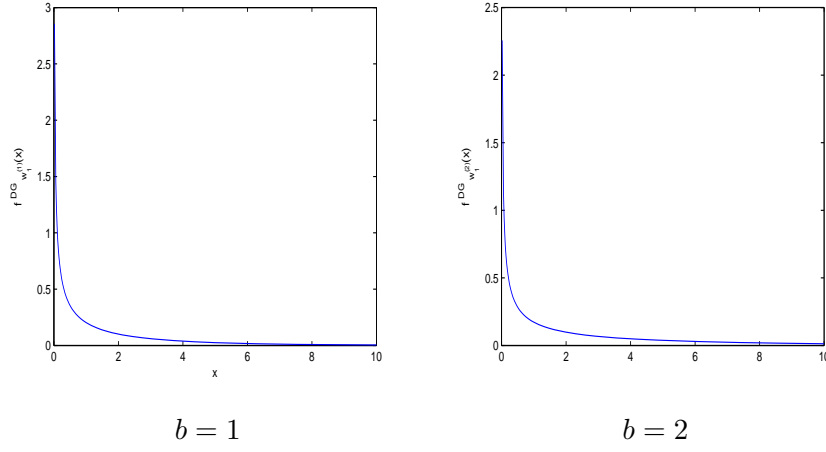


FIG. 3.8 – Densités gamma de paramètres respectifs  $\gamma_1 = 0.28$ ,  $\delta_1 = 4.29$  et  $\gamma_2 = 0.4$ ,  $\delta_2 = 3.32$  modélisant les valeurs absolues des coefficients d’ondelettes horizontales des 2 premières composantes de l’image SPOT4 précédente pour  $j = 1$ .

facilement la matrice  $\mathbf{R}_j$  par de simples moyennes empiriques de la manière suivante :

$$\widehat{R}_j(b_1, b_2) = 1 \quad \text{si } b_1 = b_2$$

$$R_j(b_1, b_2) = \frac{\sum_{m=1}^{m=N_j} \sum_{n=1}^{n=N_j} w_j^{(b_1)}(m, n) w_j^{(b_2)}(m, n)}{\sqrt{\sum_{m=1}^{m=N_j} \sum_{n=1}^{n=N_j} w_j^{(b_1)}(m, n)^2} \sqrt{\sum_{m=1}^{m=N_j} \sum_{n=1}^{n=N_j} w_j^{(b_2)}(m, n)^2}} \quad \text{si } b_1 \neq b_2 \quad (3.12)$$

La figure 3.9 représente les densités de probabilité  $f_{\mathbf{w}_j}$  modélisant les histogrammes conjoints des coefficients d’ondelettes apparaissant dans la figure 3.3. Les densités marginales sont modélisées par des DGG, tandis qu’une copule gaussienne traduit les dépendances inter-canal. On remarque que les densités obtenues permettent une meilleure modélisation des histogrammes empiriques que les DGGM affichées dans la figure 3.4.

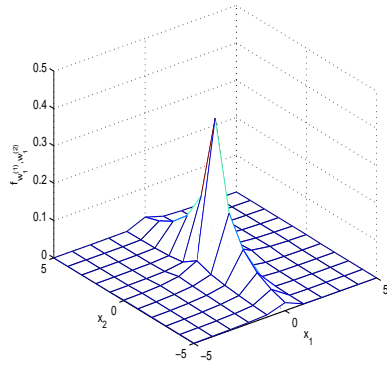


image 1

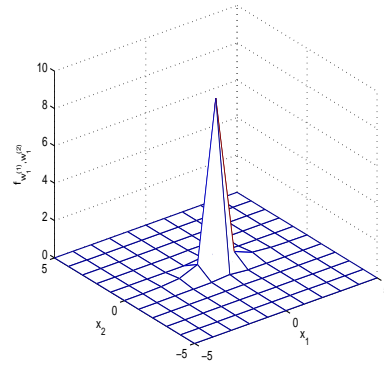


image 2

FIG. 3.9 – Densités de probabilité  $f_{\mathbf{w}_j}$  résultants de la modélisation des distributions conjointes des coefficients d'ondelettes des deux premières composantes ( $B = 2$ ) de deux images SPOT4 par des DGG et des copules gaussiennes de paramètres respectifs  $\rho_1 = 0.74$  et  $\rho_2 = 0.78$ ,  $j = 1$ , où  $\rho_1$ , respectivement  $\rho_2$  est le coefficient de corrélation linéaire entre les deux composantes.

**Vecteur des signatures** Le vecteur des signatures  $\mathbf{f}_j$  à l'étage  $j$  est formé en plus des paramètres des distributions marginales par les  $B(B-1)/2$  éléments non diagonaux supérieurs de la matrice  $\mathbf{R}_j$  qui ne sont autres que les coefficients de corrélation linéaire entre chaque couple de composantes de coefficients d'ondelettes.

A cet égard, on peut affirmer que la copule gaussienne présente l'inconvénient de traduire la structure de dépendance inter-canal à travers uniquement la corrélation linéaire entre chaque paire de composantes.

– La copule de Student :

La dépendance inter-canal entre coefficients d'ondelettes peut avoir d'autres propriétés intéressantes qui ne sont pas traduites par la simple matrice  $\mathbf{R}_j$ . Par exemple, si certains coefficients d'ondelettes relatifs à une composante  $b_1$  ont des valeurs à forte amplitude, il est fort probable que les coefficients d'ondelettes à la même position d'une autre composante  $b_2$  aient aussi de fortes amplitudes. Cette propriété n'est en fait pas reflétée par la copule gaussienne. Ceci nous a motivé à tester la copule de Student donnée par :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in [0, 1]^B, C_j(\mathbf{x}) = \varphi_{\Sigma_j, \nu_j}^t(\phi_{\nu_j}^{t-1}(x_1), \dots, \phi_{\nu_j}^{t-1}(x_B)) \quad (3.13)$$

où  $\varphi_{\Sigma_j, \nu_j}^t$  est la fonction de répartition de la distribution  $t$   $B$ -variée de degré de liberté  $\nu_j$  et de matrice de dispersion  $\Sigma_j$ . Celle ci s'obtient à partir de la matrice de corrélation  $\mathbf{R}_j$  comme suit :

$$\Sigma_j = \frac{\nu_j - 2}{\nu_j} \mathbf{R}_j \quad (3.14)$$

et  $\phi_{\nu_j}^{t-1}$  est la fonction de distribution cumulative inverse de la distribution  $t$  univariée

de degré de liberté  $\nu_j$ . La densité de la copule de Student s'écrit alors :

$$\forall (x_1, \dots, x_B) \in \mathbb{R}_j^B \quad c_j(x_1, \dots, x_B) = \frac{\varphi_{\Sigma_j, \nu_j}^t(\phi_{\nu_j}^{t-1}(x_1), \dots, \phi_{\nu_j}^{t-1}(x_B))}{\prod_{b=1}^B \varphi_{\Sigma_j, \nu_j}^t(\phi_{\nu_j}^{t-1}(x_b))} \quad (3.15)$$

Ainsi, la copule de Student représente une généralisation de la copule gaussienne. En effet, en plus de traduire la dépendance globale grâce à la matrice de dispersion  $\Sigma_j$  proportionnelle à la matrice de covariance  $\mathbf{R}_j$ , elle a l'habilité d'exprimer la dépendance entre événements extrêmes appelée aussi dépendance de queue grâce au paramètre  $\nu_j$ . Dans le cas bivarié, ceci se traduit mathématiquement pour deux composantes  $w_j^{(b_1)}$  et  $w_j^{(b_2)}$  en définissant les quantités suivantes :

$$\begin{aligned} \lambda_L &= \lim_{\alpha \rightarrow 0^+} \Pr[w_j^{(b_1)} \leq F_{w_j^{(b_1)}}^{-1}(\alpha) | w_j^{(b_2)} \leq F_{w_j^{(b_2)}}^{-1}(\alpha)], \\ \lambda_U &= \lim_{\alpha \rightarrow 1^-} \Pr[w_j^{(b_1)} > F_{w_j^{(b_1)}}^{-1}(\alpha) | w_j^{(b_2)} > F_{w_j^{(b_2)}}^{-1}(\alpha)], \\ \lambda_L &= \lambda_U = 2\varphi_{\nu_j+1}^t\left(\sqrt{\nu_j+1} \sqrt{\frac{1-\Sigma_j(b_1, b_2)}{1+\Sigma_j}}\right). \end{aligned} \quad (3.16)$$

$\lambda_U$  représente la dépendance de queue supérieure tandis que  $\lambda_L$  représente la dépendance de queue inférieure. Exprimées en fonction du paramètre  $\nu_j$ , elles se traduisent physiquement, dans notre cas, par la dépendance entre les deux composantes  $w_j^{(b_1)}$  et  $w_j^{(b_2)}$  au niveau des valeurs extrêmes. En d'autres termes,  $\lambda_U$  et  $\lambda_L$  traduisent les probabilités que les composantes  $w_j^{(b_1)}$  et  $w_j^{(b_2)}$  aient simultanément des valeurs positives ou négatives élevées. Dans le cas où  $B > 2$ , on a affaire à une copule de Student  $B$ -variée caractérisée par une matrice de dispersion  $\Sigma_j$  de taille  $B \times B$  qui traduit la corrélation linéaire entre les  $B$  composantes  $w_j^{(b)}$  et un paramètre unique  $\nu_j$  qui exprime la dépendance entre valeurs extrêmes pour chaque paire de composantes  $w_j^{(b_1)}$  et  $w_j^{(b_2)}$ ,  $b_1, b_2 \in \{1, \dots, B\}$  et  $b_1 \neq b_2$ .

**Vecteur des signatures** Dans ce cas, le vecteur des signatures  $\mathbf{f}_j$  à l'étage  $j$  est formé en plus des paramètres des distributions marginales par les  $B(B-1)/2$  éléments non diagonaux supérieurs de la matrice  $\Sigma_j$  associés aux corrélations linéaires entre les différents couples de composantes de coefficients d'ondelettes et le paramètre  $\nu_j$  qui traduit la dépendance entre les différents couples de composantes au niveau des coefficients d'ondelettes d'amplitudes extrêmes.

D'après ce qui précède, on remarque que la copule de Student est symétrique au niveau des valeurs extrêmes supérieures et inférieures ce qui n'est pas toujours le cas pour les coefficients d'ondelettes qui peuvent présenter des structures différentes de dépendances au niveau des coefficients à amplitudes élevées positives et ceux à amplitudes élevées négatives. Ainsi, une autre classe de familles de copules est préconisée dans le cadre de notre travail, à savoir les copules archimédiennes [136].

– Les copules archimédiennes :

Une copule archimédienne s'écrit sous la forme générale suivante :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in [0, 1]^B,$$

$$C_j(x_1, \dots, x_B) = \begin{cases} \varphi^{-1}(\varphi(x_1), \dots, \varphi(x_B)) & \text{si } \sum_{b=1}^B (\varphi(x_b)) \leq \varphi(0) \\ 0 & \text{sinon} \end{cases} \quad (3.17)$$

où  $\varphi$  est une fonction continue, strictement croissante de  $[0, 1]$  vers  $[0, +\infty[$  vérifiant  $\varphi(1) = 0$  et appelée fonction génératrice de la copule. La densité d'une copule archimédienne est exprimée en fonction de la fonction génératrice  $\varphi$  de la manière suivante :

$$c_j(\mathbf{x}) = (\varphi^{-1})^{(B)}(\varphi(x_1) + \dots + \varphi(x_B)) \prod_{b=1}^B \varphi'(x_b) \quad (3.18)$$

où  $(\varphi^{-1})^{(B)}$  est la dérivée d'ordre  $B$  de  $\varphi^{-1}$  et  $\varphi'$  est la dérivée de  $\varphi$ .

La copule de Gumbel et la copule de Clayton sont 2 exemples de copules archimédiennes correspondant respectivement aux fonctions génératrices suivantes :

$$\forall t \in [0, 1] \quad \begin{aligned} \varphi_{\text{Gumbel}} &= \exp(-t^{\frac{1}{\theta}}) \\ \varphi_{\text{Clayton}} &= t^\theta - 1. \end{aligned} \quad (3.19)$$

– **La copule de Gumbel :**

Elle est donnée par :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in [0, 1]^B, \quad C_j(\mathbf{x}) = \exp\left\{-\left(\sum_{b=1}^B (-\ln x_b)^{\theta_j}\right)^{\frac{1}{\theta_j}}\right\} \quad (3.20)$$

où  $\theta_j \geq 1$  est le paramètre de la copule. Malheureusement, la dérivée d'ordre  $B$  de l'inverse de la fonction génératrice de la copule de Gumbel n'admet pas une forme explicite pour exprimer la densité de la copule de Gumbel dans le cas général  $B > 2$ . Dans le cas particulier  $B = 2$ , cette dernière est donnée par :

$$\forall \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2,$$

$$c_j(\mathbf{x}) = [\ln(x_1 + x_2)]^{-1+\theta_j} \{-1 + \theta_j + [-\ln(x_1)]^{\theta_j} + [-\ln(x_2)]^{\theta_j}\}^{1/\theta_j}$$

$$\{[-\ln(x_1)]^{\theta_j} + [-\ln(x_2)]^{\theta_j}\}^{-2+1/\theta_j} \frac{1}{\exp\{[-\ln(x_1)]^{\theta_j} + [-\ln(x_2)]^{\theta_j}\}^{1/\theta_j} x_1 x_2}. \quad (3.21)$$

La copule de Gumbel permet de caractériser une forte dépendance de queue supérieure tandis que la dépendance de queue inférieure est nulle ce qui se traduit dans notre cas par une forte dépendance entre les coefficients d'ondelettes à valeurs significatives (à forte amplitude) positives. En effet les dépendances de queue supérieure et inférieure se réduisent à :

$$\begin{aligned}\lambda_U &= 2 - 2^{\frac{1}{\theta_j}} \\ \lambda_L &= 0.\end{aligned}\tag{3.22}$$

– **La copule de Clayton :**

La copule de Clayton est donnée par :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in [0, 1]^B, \quad C_j(\mathbf{x}) = \left\{ \left( \sum_{b=1}^B x_b^{-\theta_j} \right) - B + 1 \right\}^{-\frac{1}{\theta_j}}\tag{3.23}$$

où  $\theta_j > 0$  est le paramètre de la copule. Elle est associée à la densité  $c_j$  :

$$\forall \mathbf{x} = (x_1, \dots, x_B) \in \mathbb{R}^B, \quad c_j(\mathbf{x}) = \prod_{b=1}^B \{1 + (B-1)\theta_j\} x_b^{-\theta_j+1} \left[ \left( \sum_{b=1}^B x_b^{-\theta_j} \right) - B + 1 \right]^{-(\theta_j^{-1}+B)}.\tag{3.24}$$

Contrairement à la copule de Gumbel, la copule de Clayton caractérise une forte dépendance de queue inférieure tandis que la dépendance de queue supérieure est nulle ce qui se traduit dans notre cas par une forte dépendance entre les coefficients d'ondelettes à valeurs significatives (à grande amplitude) négatives. Les dépendances de queue supérieure et inférieure sont, en effet données par :

$$\begin{aligned}\lambda_U &= 0 \\ \lambda_L &= 2^{\frac{1}{\theta_j}}.\end{aligned}\tag{3.25}$$

### 3.2.6 Sélection de la famille de copules

D'après ce qui précède, on remarque que le choix de la famille de copules qui modélise le mieux les dépendances inter-canaux constitue un véritable défi. Pour choisir la meilleure famille de copules parmi un ensemble prédéfini de  $L$  familles de copules paramétriques  $\mathcal{C} = \{C_l; \theta_l\}_{l=1}^L$ , où  $\theta_l$  représente le paramètre de la copule, l'approche la plus intuitive consiste à tester plusieurs familles de copules sur les réalisations de coefficients d'ondelettes disponibles et à retenir celle qui donne les meilleures performances lors de la procédure de recherche. Cette approche présente l'inconvénient d'être gourmande en termes de calculs limitant drastiquement sa portée pratique. D'autre part, elle suppose que la dépendance inter-canaux correspondante à toutes les images de la base est modélisée à tous les étages  $j = 1, \dots, J$  par la même famille de copules ce qui n'est pas forcément le cas. Ainsi, d'autres méthodes de sélection peuvent être envisagées. Nous pouvons les classer en deux grandes catégories.

– Méthodes nécessitant une estimation préalable des paramètres :

Les paramètres de chaque famille candidate sont d'abord estimés puis un critère d'adéquation entre la distribution empirique et le modèle tel que le critère de maximum de vraisemblance ou un test d'adéquation est appliqué pour retenir la meilleure famille [137, 138, 139].



– Méthodes ne nécessitant pas une estimation préalable des paramètres :

Pour éviter une éventuelle estimation préalable des paramètres des copules à tester qui risque d’être coûteuse en calculs, une approche bayésienne pour la sélection de la meilleure famille de copules à été proposée dans [140] dans le cas bivarié  $B = 2$ . Dans l’annexe A, nous présentons une généralisation de cette méthode au cas multivarié  $B \geq 2$  [141]. Elle a pour but de sélectionner la meilleure famille de copules  $C_{l^*, \theta_{l^*}}$  dans  $\mathcal{C} = \{C_l, \theta_l\}_{l=1}^L$  sans estimation préalable de l’ensemble des  $L$  paramètres  $\{\theta_l\}_{l=1}^L$ . Une fois la famille de copule  $C_{l^*, \theta_{l^*}}$  sélectionnée, le paramètre  $\theta_{l^*}$  peut être estimé à partir de réalisations du vecteur de coefficients d’ondelettes  $\mathbf{w}_j$  par une approche par maximum de vraisemblance ou des moments.

**Vecteur des signatures :** D’après ce qui précède, on peut déduire que les coefficients d’ondelettes d’une image de la base  $i_{db}$  à un étage  $j \in \{1, \dots, J\}$  sont caractérisés par le vecteur descripteur  $\mathbf{f}_j^{i_{db}}$  formé par les éléments suivants.

- Les descripteurs “marginiaux” correspondant à chaque composante  $b$  et qui sont soit les paramètres  $\alpha_j^{(b), i_{db}}$  et  $\beta_j^{(b), i_{db}}$  dans le cas où les distributions marginales  $F_{w_j^{(b), i_{db}}}$  sont modélisées par des DGG soit par les paramètres  $\gamma_j^{(b), i_{db}}$  et  $\delta_j^{(b), i_{db}}$  quand la DG est choisie.
- Le paramètre vecteur  $\bar{\theta}_{l_j^*, i_{db}}$  de la copule  $C_{l_j^*, i_{db}; \theta_{l_j^*, i_{db}}}$  traduisant les dépendances inter-spectrales.

Le tableau 3.1 récapitule les différents choix possibles que nous avons envisagés.

TAB. 3.1 – Signatures issues des modèles multivariés proposés.

	DGG	DG
ignorée	$\alpha_j^{(b), i_{db}}, \beta_j^{(b), i_{db}}$	$\gamma_j^{(b), i_{db}}, \delta_j^{(b), i_{db}}$
Gaussienne	$\alpha_j^{(b), i_{db}}, \beta_j^{(b), i_{db}}, \mathbf{R}_j$	$\gamma_j^{(b), i_{db}}, \delta_j^{(b), i_{db}}, \mathbf{R}_j$
Student	$\alpha_j^{(b), i_{db}}, \beta_j^{(b), i_{db}}, \Sigma_j, \nu_j$	$\gamma_j^{(b), i_{db}}, \delta_j^{(b), i_{db}}, \Sigma_j, \nu_j$
Clayton	$\alpha_j^{(b), i_{db}}, \beta_j^{(b), i_{db}}, \theta_j$	$\gamma_j^{(b), i_{db}}, \delta_j^{(b), i_{db}}, \theta_j$
Gumbel	$\alpha_j^{(b), i_{db}}, \beta_j^{(b), i_{db}}$	$\gamma_j^{(b), i_{db}}, \delta_j^{(b), i_{db}}, \theta_j$

On remarque que pour une décomposition en ondelettes à  $J$  étages, les descripteurs sont calculées pour chaque étage  $j$  séparément. Ainsi la dépendance inter-échelle est complètement ignorée. Pour pallier cette insuffisance, notre deuxième contribution a consisté à élaborer des modèles plus sophistiqués que ce que nous venons de dériver.

### 3.3 Modèles prenant en compte les dépendances inter-échelles et inter-spectrales

#### 3.3.1 Motivation

Des dépendances importantes appelées dépendances inter-échelles existent également entre les coefficients d'ondelettes à différents étages de décomposition. En effet, comme le montre la figure 3.10, un coefficient d'ondelettes  $w_j^{(b)}(m, n)$  à un étage  $j > 1$  est fortement dépendant de ses quatre coefficients-fils à l'étage le plus fin  $j - 1$  :  $w_{j-1}^{(b)}(2m, 2n)$ ,  $w_{j-1}^{(b)}(2m + 1, 2n)$ ,  $w_{j-1}^{(b)}(2m, 2n + 1)$  et  $w_{j-1}^{(b)}(2m + 1, 2n + 1)$ . Ceci peut se traduire également par le fait qu'un coefficient  $w_j^{(b)}(m, n)$ ,  $0 < j < J$  dépend de ses  $J - j$  ascendants (ou parents) aux échelles les plus grossières  $j + p$ ,  $p \in \{1, \dots, J - j\}$  qui correspondent aux coefficients  $w_{j+p}^{(b)}(\lfloor \frac{m}{2^p} \rfloor, \lfloor \frac{n}{2^p} \rfloor)$ . Plus précisément, une dépendance importante existe entre un vecteur  $\mathbf{w}_j(m, n)$  et ses parents  $\mathbf{w}_{j+p}(\lfloor \frac{m}{2^p} \rfloor, \lfloor \frac{n}{2^p} \rfloor)$  aux étages  $j + p$ ,  $p = 1, \dots, J - j$ .

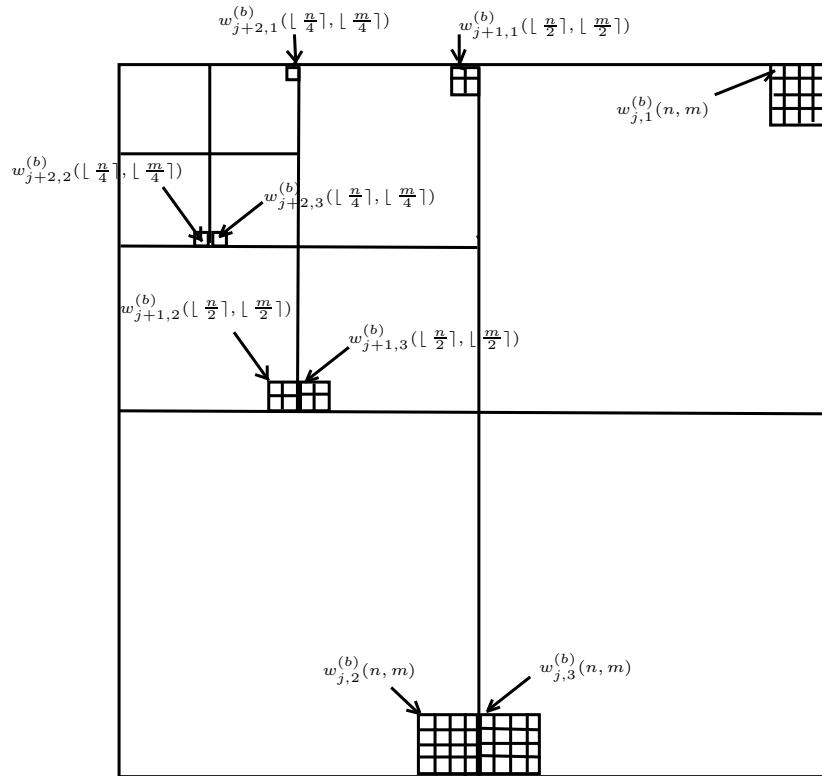


FIG. 3.10 – Dépendances inter-échelle entre coefficients d'ondelettes.

A notre connaissance, peu de travaux considèrent en même temps les dépendances inter-canaux et inter-échelles lors de l'extraction de signatures. Nous allons proposer de le faire en recourant encore une fois à la théorie des copules [142].

### 3.3.2 Principe

Soit  $1 \leq j_u \leq J$  un étage choisi par l'utilisateur, nous cherchons à exprimer la fonction de répartition conjointe  $F_{\tilde{\mathbf{w}}_{j_u}}$  du vecteur  $\tilde{\mathbf{w}}_{j_u}(m, n)$  de dimension  $d_{j_u} = B(J - j_u + 1)$  défini par :

$$\tilde{\mathbf{w}}_{j_u}(m, n) = (\dots, \mathbf{w}_{j_u+p}^\top(\lfloor \frac{m}{2^{j_u+p}} \rfloor, \lfloor \frac{n}{2^{j_u+p}} \rfloor), \dots)^\top \quad (3.26)$$

où  $p = 0, \dots, J - j_u$ . La difficulté réside dans le fait que le vecteur aléatoire  $\tilde{\mathbf{w}}_{j_u}(m, n)$  est formé de sous-vecteurs aléatoires  $\mathbf{w}_{j_u+p}^\top(\lfloor \frac{m}{2^{j_u+p}} \rfloor, \lfloor \frac{n}{2^{j_u+p}} \rfloor)$  de dimension  $B$ . Heureusement, cette difficulté peut être résolue grâce à une généralisation récente du théorème de Sklar [143] expliqué dans l'annexe B. Ainsi, comme le montre la figure 3.11, la fonction de distribution  $F_{\tilde{\mathbf{w}}_{j_u}}$  peut-elle être exprimée pour tout vecteur  $(x_1, \dots, x_{d_{j_u}})^\top$  dans  $\mathbb{R}^{d_{j_u}}$  de la manière suivante :

$$F_{\tilde{\mathbf{w}}_{j_u}}(x_1, \dots, x_{d_{j_u}}) = \tilde{C}_{j_u} \left( \bar{C}_{j_u} (F_{w_{j_u}^{(1)}}(x_1), \dots, F_{w_{j_u}^{(B)}}(x_B)), \dots, \bar{C}_J (F_{w_J^{(1)}}(x_{d_{j_u}-B+1}), \dots, F_{w_J^{(B)}}(x_{d_{j_u}})) \right) \quad (3.27)$$

où  $p = 0, \dots, J - j_u$ . Dans cette équation, 3 entités principales sont mises en jeu :

- $F_{w_{j_u+p}^{(b)}}$  les distributions marginales de  $w_{j_u+p}^{(b)}$ ,
- les copules  $B$ -dimensionnelles  $\bar{C}_{j_u+p}$  traduisent la dépendance entre les  $B$  composantes  $w_{j_u+p}^{(b)}$  à l'étage de résolution  $j_u + p$ ,
- la copule  $d_{j_u}$ -dimensionnelle  $\tilde{C}_{j_u}$  traduit la lien entre le vecteur  $\mathbf{w}_{j_u}(m, n)$  et ses parents  $\mathbf{w}_{j_u+p}(m, n)$  aux étages  $j_u + p$  pour  $p = 1, \dots, J - j_u$ .

Ainsi, comme le montre la figure 3.11, trouver un modèle multivarié pour la distribution  $F_{\tilde{\mathbf{w}}_{j_u}}$  du vecteur  $\tilde{\mathbf{w}}_{j_u}(m, n)$  se résume à trouver des modèles adéquats pour ces 3 entités.

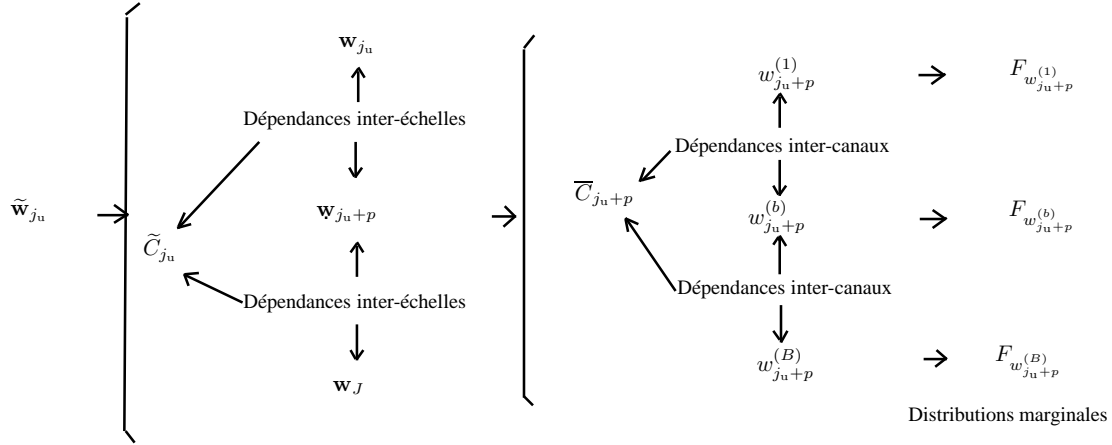


FIG. 3.11 – Modélisation des coefficients d'ondelettes prenant en compte les distributions marginales, les dépendances inter-spectrales et inter-échelles

### 3.3.3 Lois marginales

En ce qui concerne les distributions marginales  $F_{w_{j_u+p}}^{(b)}$ , elles sont modélisées comme précédemment dans la section 3.2.

### 3.3.4 Copules inter-spectrale et copules inter-échelle

Les copules  $\bar{C}_{j_u+p}$  aux étages  $j_u + p$ , ( $p = 0, \dots, J - j_u$ ) ainsi que la copule  $\tilde{C}_{j_u}$  sont choisies parmi un ensemble prédéfini de  $L$  familles de copules  $\{C_l; \theta_l\}_{l=1}^L$  par l'une des méthodes de sélection de copule évoquées précédemment.

En particulier, il est important de signaler que différentes familles de copules  $\bar{C}_{l_{j_u+p}^*, i_{db}; \bar{\theta}_{l_{j_u+p}^*, i_{db}}}$  pour traduire la dépendance inter-canaux pour chaque étage  $j_u + p$  et pour chaque image  $i_{db}$  de la base peuvent être envisagées. De même, la famille de copules  $\tilde{C}_{l_{j_u}^*, i_{db}; \tilde{\theta}_{l_{j_u}^*, i_{db}}}$  reflétant la dépendance inter-échelle peut différer selon l'image  $i_{db}$  de la base. Une fois les familles de copules  $\bar{C}_{l_{j_u+p}^*, i_{db}; \bar{\theta}_{l_{j_u+p}^*, i_{db}}}$  et  $\tilde{C}_{l_{j_u}^*, i_{db}; \tilde{\theta}_{l_{j_u}^*, i_{db}}}$  retenues, les paramètres  $\bar{\theta}_{l_{j_u+p}^*, i_{db}}$  et  $\tilde{\theta}_{l_{j_u}^*, i_{db}}$  peuvent être estimés grâce à une estimation par maximum de vraisemblance ou par une méthode des moments.

### 3.3.5 Signatures proposées

D'après ce qui précède, les coefficients d'ondelettes résultants de la décomposition d'une image  $i_{db}$  aux étages  $j_u, j_u + 1, \dots, J$  sont caractérisés par le vecteur des descripteurs  $\mathbf{f}_{j_u}^{i_{db}}$  formé des éléments suivants.

- Les paramètres des lois marginales correspondantes à chaque composante  $b$  et chaque étage  $j_u + p$ . Il s'agit des paramètres  $\alpha_{j_u+p}^{(b), i_{db}}$  et  $\beta_{j_u+p}^{(b), i_{db}}$  dans le cas d'une modélisation par des DGG soit des paramètres  $\gamma_{j_u+p}^{(b), i_{db}}$  et  $\delta_{j_u+p}^{(b), i_{db}}$  quand la DG est choisie.
- Les paramètres vecteur  $\bar{\theta}_{l_{j_u+p}^*, i_{db}}$  des copules  $\bar{C}_{l_{j_u+p}^*, i_{db}; \bar{\theta}_{l_{j_u+p}^*, i_{db}}}$  traduisant les dépendances inter-composante à chaque étage  $j_u + p$ .
- Le paramètre vecteur  $\tilde{\theta}_{l_{j_u}^*, i_{db}}$  de la copule  $\tilde{C}_{l_{j_u}^*, i_{db}; \tilde{\theta}_{l_{j_u}^*, i_{db}}}$  traduisant les dépendances inter-échelles.

Certains cas particuliers sont à noter :

- Si les copules  $\bar{C}_{l_{j_u+p}^*, i_{db}; \bar{\theta}_{l_{j_u+p}^*, i_{db}}}$  sont choisies a priori égales 1, les paramètres  $\bar{\theta}_{l_{j_u+p}^*, i_{db}}$  n'ont plus lieu d'être et les coefficients d'ondelettes relatifs aux différentes composantes  $b = 1, \dots, B$  aux étages  $j_u + p$  sont considérés comme indépendants (dépendance inter-canaux ignorée)
- Si la copule  $\tilde{C}_{l_{j_u}^*, i_{db}; \tilde{\theta}_{l_{j_u}^*, i_{db}}}$  est prise égale à 1, le paramètre  $\tilde{\theta}_{l_{j_u}^*, i_{db}}$  n'a pas lieu d'être et les dépendances inter-échelle ne sont pas prises en compte.

### 3.3.6 Récapitulons !

Nos deux premières contributions présentent les avantages suivants.

- Considérer en même temps les propriétés marginales des coefficients d’ondelettes relatifs à chaque composante  $b$ , la dépendance inter-canaux et la dépendance inter-échelle.
- Permettre de modéliser et d’estimer indépendamment les distributions des coefficients d’ondelettes marginaux relatifs à chaque composante  $b$ , la dépendance inter-canaux et la dépendance inter-échelle.
- Construire un modèle multivarié des coefficients d’ondelettes qui, contrairement à la DGGM, permet de générer une grande variété de densités conjointes couvrant la diversité des histogrammes conjoints de coefficients d’ondelettes et ceci en jouant sur la modélisation des densités marginales et du choix de la copule.

Cependant, ces deux premières méthodes d’extraction de signatures ne font que prendre en compte les dépendances des coefficients uniquement par la distribution sans exploiter au préalable ces similarités au niveau de la décomposition multirésolution. Par ailleurs, les descripteurs proposés ne sont pas forcément invariants par rapport à certaines transformations de l’image comme la rotation. Pour pallier de telles insuffisances, nous proposons une nouvelle piste à savoir agir au niveau de la décomposition multi-échelle.

## 3.4 Contribution 3 : agir sur la décomposition multi-échelle

### 3.4.1 Motivation

L’extraction des meilleurs descripteurs ne se limite pas au choix du modèle adéquat pour les coefficients d’ondelettes, il faut également choisir une décomposition multirésolution qui reflète le plus “fidèlement” possible les caractéristiques statistiques de l’image. Jusqu’à présent, la décomposition multirésolution envisagée est la S-TOD appliquée séparément à chaque composante de l’image multispectrale. Néanmoins, malgré sa simplicité, cette décomposition présente les inconvénients suivants.

- Seules les redondances spatiales sont exploitées tandis que les redondances spectrales (dépendances inter-canaux) ne sont pas prises en considération au niveau des calculs des coefficients.
- Une faible translation de l’image peut induire une modification significative des coefficients d’ondelettes (absence d’invariance par translation).
- Une décomposition dyadique souffre d’être pauvre en sélectivité directionnelle. En effet, les coefficients d’ondelettes sont orientés uniquement selon 3 orientations différentes (horizontale, verticale et diagonale). Ainsi, les contours orientés à  $+45$  et  $-45$  degrés vont contribuer aux coefficients de la même sous-bande diagonale.

Pour contourner ces problèmes, nous proposons de remplacer la décomposition en ondelettes conventionnelle par d'autres décompositions multirésolutions pour satisfaire les deux objectifs suivants :

- exploiter les dépendances inter-canaux au niveau du choix d'une transformation multirésolution qui tient compte, en plus des redondances spatiales, des redondances spectrales entre différents spectres ou canaux de l'image [144] ;
- choisir une décomposition multirésolution qui réalise un compromis entre le facteur de redondance et la richesse en information directionnelle afin de permettre d'extraire des signatures invariantes par rotation tout en respectant les contraintes de compression [145].

### 3.4.2 Recours au schéma de lifting vectoriel

Jusqu'alors, nous avons utilisé une décomposition en ondelettes se faisant de manière indépendante pour chaque composante de l'image. Or, cette décomposition exploite seulement les redondances intra-composante tout en négligeant les redondances inter-composante qui peuvent jouer un rôle important dans la description du contenu de l'image. Le schéma du lifting vectoriel proposé en [118] a été conçu pour résoudre ce problème.

Comme expliqué dans l'annexe C, la décomposition en  $J$  étages d'un signal 1D à  $B$  composantes résulte, comme pour le cas du schéma de lifting conventionnel, en  $B$  signaux approximation à la résolution la plus grossière  $J$  et  $B$  signaux détails à chaque résolution  $1, \dots, J$ . La différence avec le schéma de lifting conventionnel est qu'un codage hybride (intra et inter-canaux) est utilisé. Pour une permutation donnée  $(b_1, \dots, b_B)$  des  $B$  composantes, il a pour principe de coder la composante  $b_1$  en mode intra c'est-à-dire selon le schéma de lifting conventionnel tandis que les autres composantes  $b_i$  ( $i = 2, \dots, B$ ) sont codées en utilisant à la fois l'information spatiale de la même composante  $b_i$  et les informations relatives aux composantes précédentes  $b_k$  ( $k < i$ ).

La décomposition d'une image multispectrale résulte comme dans le cas du schéma de lifting conventionnel en une image approximation multispectrale  $(a_J^{(b)})_{b=1}^B$  de taille  $N_J \times N_J$  avec  $N_J = \frac{N}{2^J}$  à l'étage le plus grossier  $J$  et 3 sous-bandes détails multispectrales  $(w_{j,o}^{(b)})_{b=1}^B$  de taille  $N_j \times N_j \times B$  avec  $N_j = N/2^j$  orientées horizontalement ( $o = 1$ ), verticalement ( $o = 2$ ) et diagonalement ( $o = 3$ ) à chaque étage  $j = 1, \dots, J$ . Grâce au mode de codage hybride, les redondances spatiales et spectrales présentes au niveau de l'image multispectrale sont simultanément exploitées dans le calcul des coefficients détails  $(w_{j,o}^{(b)})_{b=1}^B$ . Ainsi, la dépendance statistique entre une composante  $w_{j,o}^{(b)}$  et une autre composante  $w_{j,o}^{(b')}$  de la sous-bande  $(w_{j,o}^{(b)})_{b=1}^B$  est-elle moins importante que dans le cas du schéma de lifting conventionnel.

### 3.4.3 Recours à la transformation bidimensionnelle $M$ bande en arbre dual

Comme on a l'a mentionné dans la section. 3.1, une alternative largement utilisée dans la littérature pour extraire des descripteurs invariants par rotation consiste à calculer des signatures sur des coefficients résultants de la décomposition de l'image par des transformations multirésolutions sélectives en direction telles que les pyramides orientables [70] et la transformée en ondelette en arbre dual [113, 77]. Comme notre travail est situé dans le contexte de recherche d'images stockées sous forme codée, ces transformations présentent l'inconvénient d'être lourdement redondantes. Le facteur de redondance pourrait s'élever, par exemple, à  $4O/3 = 8$  dans le cas d'une décomposition en pyramides orientables à  $O = 6$  orientations. Ainsi, un défi consiste à choisir une transformation multirésolution qui assure un compromis entre la sélectivité directionnelle et la compacité de la représentation. Dans notre travail, on a proposé de tester la transformée en ondelettes  $M$  bandes en arbre dual [80] (2D TOM-AD) décrite dans le chapitre 2. Celle-ci présente l'avantage d'un faible facteur de redondance (égal à 2) et ceci indépendamment du nombre d'orientations  $2(M^2 - 1)$ .

Dans le cas d'une image multispectrale, la décomposition TOM-AD sur  $J$  étages se fait indépendamment sur chaque canal de l'image ce qui résulte pour chaque canal  $b = 1, \dots, B$  en 2 images approximations  $a_j^{(b,h)}(p, q)$ ,  $h \in \{0, 1\}$ , ( $h = 0$  : arbre primaire et  $h = 1$  : arbre dual),  $p, q \in \{1, \dots, \frac{N}{M^j}\}$  à l'étage le plus grossier  $J$  et  $2 \times (M^2 - 1)$  sous-bandes détail à chaque étage  $j$  orientées dans des directions différentes  $w_{j,m,m'}^{(h,b)}(p, q)$ , avec  $h \in \{0, 1\}$ ,  $(m, m') \in \{1, \dots, M\}$  et  $(m, m') \neq (0, 0)$ .

**Descripteurs relatifs à chaque sous-bande** Pour chaque arbre  $h \in \{0, 1\}$ ,  $(m, m') \in \{1, \dots, M\}$  et chaque sous-bande  $(m, m') \neq (0, 0)$  nous définissons à un étage  $j$  le vecteur de coefficients :

$$\mathbf{w}_{j,m,m'}^{(h)}(p, q) \triangleq (w_{j,m,m'}^{(h,1)}(p, q), \dots, w_{j,m,m'}^{(h,B)}(p, q))^T. \quad (3.28)$$

Soit  $j_u$  un étage choisi par l'utilisateur, on considère alors, le vecteur  $\tilde{\mathbf{w}}_{j_u,m,m'}^{(h)}(p, q)$  défini par :

$$\tilde{\mathbf{w}}_{j_u,m,m'}^{(h)}(p, q) = (\dots, \mathbf{w}_{j_u+k,m,m'}^{(h)T}(\lfloor \frac{p}{2^{j_u+k}} \rfloor, \lfloor \frac{q}{2^{j_u+k}} \rfloor), \dots)^T \quad (3.29)$$

où  $k = 1, \dots, J - j_u$ ,  $w_{j_u+k,m,m'}^{(h,b)}(\lfloor \frac{p}{2^{j_u+k}} \rfloor, \lfloor \frac{q}{2^{j_u+k}} \rfloor)$  est l'ascendant à l'étage  $j_u + k$  de la  $b$ -ième composante  $w_{j_u,m,m'}^{(h,b)}(p, q)$  du vecteur  $\mathbf{w}_{j_u,m,m'}^{(h)T}(p, q)$ .

La fonction de répartition conjointe  $F_{\tilde{\mathbf{w}}_{j_u,m,m'}^{(h)}}$  est modélisée grâce aux théorème de Sklar généralisé. Cette distribution est alors exprimée en fonction des distributions marginales  $F_{w_{j_u+k,m,m'}^{(h,b)}}$  des coefficients d'ondelettes relatives à chaque composante  $b$  et chaque étage  $j_u + k$  et à l'orientation choisie, des copules  $\overline{C}_{j_u+k,m,m'}^{(h)}$  traduisant la dépendance inter-canaux à chaque étage  $j_u + k$  et la copule  $\tilde{C}_{j_u,m,m'}^{(h)}$  reflétant la dépendance inter-échelles.

Ainsi, les coefficients résultant de la décomposition 2D TOM-AD en  $J - j_u$  étages d'une

image  $i_{db}$  de la base sont caractérisés à une orientation donnée  $(h, m, m')$  par le vecteur des signature  $\mathbf{f}_{j_u, m, m'}^{i_{db}, (h)}$  formé par les éléments suivants :

- les paramètres  $\alpha_{j_u+k, m, m'}^{(b, h), i_{db}}$  et  $\beta_{j_u+k, m, m'}^{(b, h), i_{db}}$  si les distributions marginales  $F_{j_u+k, m, m'}^{(b, h), i_{db}}$  sont modélisées par des DGG ou les paramètres  $\gamma_{j_u+k, m, m'}^{(b, h), i_{db}}$  et  $\delta_{j_u+k, m, m'}^{(b, h), i_{db}}$  quand elles sont modélisées par des DG ;
- les paramètres dans le vecteur  $\bar{\boldsymbol{\theta}}_{j_u+k, m, m'}^{l^*, (h), i_{db}}$  des copules  $\bar{C}_{j_u+k, m, m'}^{l^*, (h), i_{db}}$  traduisant les dépendances inter-composante à chaque étage  $j_u + k$  ;
- les paramètres dans le vecteur  $\tilde{\boldsymbol{\theta}}_{j_u, m, m'}^{l^*, (h), i_{db}}$  de la copule  $\tilde{C}_{j_u, m, m'}^{l^*, (h), i_{db}}$  traduisant les dépendances inter-échelles [145].

Rappelons que les copules  $\bar{C}_{j_u+k, m, m'}^{l^*, (h), i_{db}}$  et  $\tilde{C}_{j_u, m, m'}^{l^*, (h), i_{db}}$  sont choisies parmi les listes prédéfinies de familles de copules décrite précédemment  $\{\bar{C}_{j_u+k, m, m'}^{l(h)}, \bar{\boldsymbol{\theta}}_{j_u+k, m, m'}^{l(h)}\}_{l=1}^L$  et  $\{\tilde{C}_{j_u, m, m'}^{l(h)}, \tilde{\boldsymbol{\theta}}_{j_u, m, m'}^{l(h)}\}_{l=1}^{L'}$ .

**Signatures invariantes par rotation** D'après ce qui précède, on remarque que chaque sous-bande  $\tilde{\mathbf{w}}_{j_u, m, m'}^{i_{db}, (h)}$  d'une image  $i_{db}$  de la base contient de l'information relative à une orientation donnée caractérisée par les paramètres  $m, m', h$ . Ainsi une version tournée  $\underline{i}_{db}$  de l'image  $i_{db}$  aurait probablement les mêmes coefficients d'ondelettes que ceux de l'image  $i_{db}$  à une orientation  $m, m', h$  mais dans une autre orientation  $\underline{m}, \underline{m}', \underline{h}$ . Ainsi la moyenne de l'information donnée par les  $2(M^2 - 1)$  sous-bandes peut être considérée comme inchangée pour une image  $i_{db}$  et sa copie tournée  $\underline{i}_{db}$ . On en déduit facilement que pour calculer des signatures invariantes par rotation pour une image  $i_{db}$  de la base, il suffit de calculer le vecteur moyenne des vecteurs  $\mathbf{f}_{j_u, m, m'}^{i_{db}, (h)}$  sur toutes les  $2(M^2 - 1)$  orientations [145] dans le but d'obtenir le vecteur des signatures invariant par rotation  $\mathbf{f}_{RI_{j_u}}^{i_{db}}$  donné par :

$$\mathbf{f}_{RI_{j_u}}^{i_{db}} = \frac{1}{2(M^2 - 1)} \sum_{(m, m') \neq (0, 0)}^{M-1} \sum_{h=0}^1 \mathbf{f}_{j_u, m, m'}^{i_{db}, (h)}. \quad (3.30)$$

### 3.5 Conclusion

Dans ce chapitre, nous avons présenté la première partie de nos contributions qui se rapporte à l'étape d'extraction des signatures à partir d'images multispectrales dans le domaine des ondelettes. Dans une première partie, nous avons considéré des décompositions multi-échelle appliquée indépendamment à chaque canal de l'image multispectrale. Nous avons, alors, eu recours à la théorie des copules pour établir des modèles conjoints multivariés adéquats qui prennent en considération en plus des distributions marginales des coefficients



d'ondelettes de chaque composante et à chaque étage, les dépendances inter-canaux et inter-échelles. Nous, avons essayé par la suite de généraliser ce travail en appliquant des décompositions multirésolutions ayant des propriétés supplémentaires par rapport aux décompositions conventionnelles pour traduire des caractéristiques statistiques de l'image multispectrale. En particulier, nous avons adopté le schéma de lifting vectoriel pour tenir compte des dépendances spectrales ainsi que la 2D TOM-AD pour assurer l'invariance par rotation. Dans le prochain chapitre, nous aborderons le second volet de nos contributions qui consiste à optimiser l'étape de la procédure de recherche.

## Chapitre 4

# Accélération de la procédure de recherche

Une fois les descripteurs des images de la base calculés, l'étape suivante est la procédure de recherche qui permet à l'utilisateur de retrouver les images de la base semblables à une image requête. En effet, le système doit comparer le vecteur des signatures de la requête aux vecteurs des signatures stockés dans la base des index. Dans ce cadre, deux problèmes se posent. Le premier est celui du choix d'une fonction de similarité nécessaire pour effectuer cette comparaison. Le second problème touche à la conception de l'algorithme de recherche qui définit la manière dont la base des index est parcourue afin de comparer les vecteurs des signatures au vecteur requête. Dans ce chapitre, nous aborderons ces deux problèmes afin d'adapter la procédure de recherche aux signatures calculées dans le chapitre précédent. Dans la section 4.1, nous commençons par présenter un bref état de l'art sur les fonctions de similarité les plus utilisées en indexation y compris la distance euclidienne normalisée. Par la suite, nous décrivons une autre distance considérée comme plus adaptée à notre cas à savoir la distance de Kullback-Leibler symétrique. Dans la section 4.2, nous proposons une recherche arborescente exploitant la propriété d'échelonnabilité des décompositions multirésolutions. Enfin, certaines conclusions seront établies dans la section 4.3.

### 4.1 Choix de la mesure de similarité

#### 4.1.1 Intérêt d'une fonction de similarité

Avant de concevoir un algorithme de recherche, une fonction de similarité doit être définie pour comparer un vecteur des signatures  $\mathbf{f}_j^{i_q}$  d'une image requête  $i_q$  à un vecteur des signatures  $\mathbf{f}_j^{i_{db}}$  d'une image  $i_{db}$  de la base d'images. Dans la littérature, plusieurs définitions d'une fonction de similarité, sont employées dans plusieurs applications de traitement d'images telles que la classification, la reconnaissance et l'indexation [147]. Les plus connues sont celles

qui correspondent à une distance  $d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_{db}})$  satisfaisant les axiomes suivants [146, 148].

- l’auto-similarité  $d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_q}) = d(\mathbf{f}_j^{i_{db}}, \mathbf{f}_j^{i_{db}}) = 0$  ;
- la minimalité  $d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_{db}}) \geq d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_q}) \geq 0$  ;
- la symétrie  $d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_{db}}) = d(\mathbf{f}_j^{i_{db}}, \mathbf{f}_j^{i_q})$  ;
- l’inégalité triangulaire  $d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i_{db}}) + d(\mathbf{f}_j^{i_{db}}, \mathbf{f}_j^{i'_{db}}) \geq d(\mathbf{f}_j^{i_q}, \mathbf{f}_j^{i'_{db}})$ .

Notons que le dernier axiome a l’intérêt de réduire le complexité de la recherche.

#### 4.1.2 Mesures de similarité de type distances de Minkowski

Plusieurs distances ont été adoptées en indexation d’images pour la comparaison de vecteurs des signatures  $\mathbf{s}_1$  et  $\mathbf{s}_2$  dans le contexte de recherche d’images. Parmi elles, on peut citer les distances conventionnelles de Minkowski qui peuvent s’écrire sous la forme :

$$\forall \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^L \quad d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt[p]{\sum_{i=1}^L |\mathbf{s}_1(i) - \mathbf{s}_2(i)|^p} \quad (4.1)$$

où  $p \in \mathbb{N}^*$  et  $\mathbf{s}_1(i)$ , (resp.  $\mathbf{s}_2(i)$ ) représente le  $i$ ’ième élément du vecteur  $\mathbf{s}_1$ , (resp.  $\mathbf{s}_2$ ).

Des exemples de distance de Minkowski les plus utilisées dans le contexte de recherche d’images sont les suivants :

- la distance euclidienne correspond à  $p = 2$  ;
- la distance de Manhattan (ou distance  $L^1$ ) est définie pour  $p = 1$ .

Dans le cas de la distance euclidienne conventionnelle, certaines composantes des vecteurs des signatures  $\mathbf{s}_1$  et  $\mathbf{s}_2$  ont des variances nettement plus grandes que les autres composantes. Elles ont tendance à dominer la mesure de la distance euclidienne. D’autre part, certaines composantes sont bruitées ce qui affecte la mesure de la distance. Deux distances sont conçues pour remédier à ces insuffisances :

- La distance de Mahalanobis qui s’exprime par :

$$d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{(\mathbf{s}_1 - \mathbf{s}_2)^T \Sigma^{-1} (\mathbf{s}_1 - \mathbf{s}_2)} \quad (4.2)$$

où  $\Sigma$  est la matrice de covariance du vecteur des signatures  $\mathbf{s}$  estimé sur toutes les images de la base.

- La distance euclidienne normalisée (DEN) qui est définie par :

$$d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{(\tilde{\mathbf{s}}_1 - \tilde{\mathbf{s}}_2)^T (\tilde{\mathbf{s}}_1 - \tilde{\mathbf{s}}_2)} \quad (4.3)$$

où  $\tilde{\mathbf{s}}_1$  et  $\tilde{\mathbf{s}}_2$  sont les vecteurs des signatures normalisés obtenus par :

$$\begin{aligned} \tilde{\mathbf{s}}_1(i) &= \frac{1}{md(i)} (\mathbf{s}_1(i) - m(i)) \\ \tilde{\mathbf{s}}_2(i) &= \frac{1}{md(i)} (\mathbf{s}_2(i) - m(i)) \end{aligned} \quad (4.4)$$

avec  $m(i)$  et  $md(i)$  correspondent respectivement à la valeur médiane et à la déviation moyenne de la  $i$ -ième composante du vecteur des signatures calculées sur toutes les images de la base :

$$md(i) = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} |s_n(i) - moy(i)| \quad (4.5)$$

où  $\mathcal{N}$  est le nombre d'images de la base et  $moy(i)$  est la moyenne de la composante  $s_n(i)$  sur toutes les images de la base.

Ainsi, toutes les composantes  $\tilde{s}_n(i)$  d'un vecteur  $\tilde{s}_n$  d'une image  $n$  de la base auront la même déviation moyenne qui est égale à 1. Par suite, tous les termes  $(\tilde{s}_1(i) - \tilde{s}_2(i))^2$  de toutes les composantes dans l'expression de la DEN auront le même ordre de grandeur.

Il n'y aura donc pas de composantes qui tendent à dominer cette distance. De plus, la DEN est plus robuste au bruit et aux données aberrantes que la distance euclidienne.

Ces métriques sont largement utilisées dans les systèmes d'indexation [149, 130] vu leur simplicité ainsi que leur rapidité de mise en oeuvre. Néanmoins, plusieurs travaux ont mis en évidence les limitations de ces métriques. En effet, ces distances considèrent que deux images sont similaires si leurs vecteurs des signatures sont proches. Ceci pourrait être insuffisant pour traduire adéquatement les caractéristiques du système visuel humain dans la distinction perceptuelle des images [150]. D'autre part, elles nécessitent souvent une normalisation des données de l'espace des signatures [46]. De plus, la distance euclidienne se justifie pleinement quand les données à comparer (dans notre cas les vecteurs des signatures) suivent une loi gaussienne, tandis que la distance  $L^1$  convient pour une loi exponentielle. En l'absence de connaissance de la distribution statistique des signatures, il s'avère difficile de choisir une distance même si des travaux émergents ont été entrepris dans ce sens [151]. Ainsi, ces distances ne sont pas spécifiquement adaptées à la comparaison de modèles statistiques dans le cadre d'une approche statistique de modélisation des coefficients d'ondelettes. Pour remédier à ces limites, il est recommandé dans [46] de recourir à la distance symétrique de Kullback-Leibler (DSKL).

### 4.1.3 Distance symétrique de Kullback-Leibler

**Définition** Dans le contexte d'indexation d'images de texture monochromes, Do et Vetterli [46] ont proposé d'utiliser la DSKL pour trouver les images  $i_{db}$  de la base similaires à une image requête  $i_q$  et ceci dans le cadre d'une modélisation statistique de leurs coefficients d'ondelettes. Le but consiste alors à trouver les images  $i_{db}$  caractérisées par des densités des coefficients d'ondelettes  $p_{i_{db}}(x)$  "proches" de celle de l'image requête  $p_{i_q}(x)$ . En théorie de l'information, il est d'usage de définir des divergences pour effectuer cette comparaison. Leur principe est de mesurer l'inefficacité moyenne de coder la loi  $p_{i_q}$  en utilisant le code de  $p_{i_{db}}$ . Parmi les multiples définitions de divergence, la divergence de Kullback-Leibler est

fréquentement retenue. En effet, elle présente plusieurs avantages.

- Elle s'avère robuste par rapport au bruit entachant éventuellement les données.
- Elle est numériquement stable (notamment en fonction du nombre de données).

L'expression de la divergence de Kullback-Leibler est la suivante :

$$d(p_{i_q} \parallel p_{i_{db}}) = \int p_{i_q}(x) \log\left(\frac{p_{i_q}(x)}{p_{i_{db}}(x)}\right) dx. \quad (4.6)$$

Comme la divergence de Kullback-Leibler, n'est pas symétrique, elle ne peut pas être une distance. Or, comme on l'a mentionné précédemment, l'inégalité triangulaire est une propriété intéressante pour réduire la complexité opératoire de la recherche. On définit alors la distance de Kullback-Leibler symétrique en moyennant entre la divergence  $d(p_{i_q} \parallel p_{i_{db}})$  et  $d(p_{i_{db}} \parallel p_{i_q})$  :

$$D(p_{i_q}, p_{i_{db}}) = \frac{1}{2} (d(p_{i_q} \parallel p_{i_{db}}) + d(p_{i_{db}} \parallel p_{i_q})). \quad (4.7)$$

### Cas où seules les densités marginales des coefficients d'ondelettes sont considérées

Dans le cas conventionnel où les dépendances inter-canaux et inter-échelle sont négligées, la densité  $f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}$  d'une image de la base s'écrit comme produit des densités marginales des coefficients d'ondelettes  $w_{j_u+p}^{(b),i_{db}}(m, n)$  des  $B$  différentes composantes et aux différents étages  $j_u + p = 1, \dots, J$ , considérés comme indépendants :

$$\forall (x_1, \dots, x_{d_{j_u}})^T \in \mathbb{R}^{d_{j_u}} \quad f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}(x_1, \dots, x_{d_{j_u}}) = \prod_{b=1}^B \prod_{p=0}^{J-j_u} f_{w_{j_u+p}^{(b),i_{db}}}(x_{b+pB}) \quad (4.8)$$

où  $d_{j_u} = B(J - j_u)$ . Il en va de même pour une image requête  $i_q$ . Une propriété importante de la DSKL [152] est que  $D(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}, f_{\tilde{\mathbf{w}}_{j_u}^{i_q}})$  peut s'écrire comme la somme des DSKL entre les densités marginales  $f_{w_{j_u+p}^{(b),i_q}}$  et  $f_{w_{j_u+p}^{(b),i_{db}}}$  correspondantes :

$$D(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}, f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}) = \sum_{b=1}^B \sum_{p=0}^{J-j_u} D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}}). \quad (4.9)$$

Comme on l'a vu dans le chapitre précédent, les densités marginales  $f_{w_{j_u+p}^{(b),i_q}}$  et  $f_{w_{j_u+p}^{(b),i_{db}}}$  sont modélisées par des distributions gaussiennes généralisées (DGG) ou par des distributions gamma (DG). Dans les deux cas, la DSKL  $D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}})$  présente l'avantage d'admettre une forme explicite.

– Cas de la DGG :

Il peut être facilement prouvé que la DSKL  $D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}})$  est donnée par :

$$D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}}) = \frac{1}{2} \left( \left( \frac{\alpha_{j_u+p}^{(b),i_q}}{\alpha_{j_u+p}^{(b),i_{db}}} \right)^{\beta_{j_u+p}^{(b),i_{db}}} \frac{\Gamma(\beta_{j_u+p}^{(b),i_{db}} + 1) / \beta_{j_u+p}^{(b),i_q}}{\Gamma(1/\beta_{j_u+p}^{(b),i_q})} - 1 / \beta_{j_u+p}^{(b),i_q} \right) \\ + \frac{1}{2} \left( \left( \frac{\alpha_{j_u+p}^{(b),i_{db}}}{\alpha_{j_u+p}^{(b),i_q}} \right)^{\beta_{j_u+p}^{(b),i_q}} \times \frac{\Gamma(\beta_{j_u+p}^{(b),i_q} + 1) / \beta_{j_u+p}^{(b),i_{db}}}{\Gamma(1/\beta_{j_u+p}^{(b),i_{db}})} - 1 / \beta_{j_u+p}^{(b),i_{db}} \right). \quad (4.10)$$

où  $\beta_{j_u+p}^{(b),i_q}$ ,  $\alpha_{j_u+p}^{(b),i_q}$ , (resp.  $\beta_{j_u+p}^{(b),i_{db}}$ ,  $\alpha_{j_u+p}^{(b),i_{db}}$ ) sont les paramètres de forme et d'échelle des densités  $f_{w_{j_u+p}^{(b),i_q}}$ , (resp.  $f_{w_{j_u+p}^{(b),i_{db}}}$ ).

– Cas de la DG :

Dans le cas de la DG, une formule explicite de la DSKL  $D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}})$  peut être établie [153] :

$$D(f_{w_{j_u+p}^{(b),i_q}}, f_{w_{j_u+p}^{(b),i_{db}}}) = (\gamma_{j_u+p}^{(b),i_q} - \gamma_{j_u+p}^{(b),i_{db}}) \left( \log \left( \frac{\delta_{j_u+p}^{(b),i_q}}{\delta_{j_u+p}^{(b),i_{db}}} \right) + \psi(\gamma_{j_u+p}^{(b),i_q}) - \psi(\gamma_{j_u+p}^{(b),i_{db}}) \right) \\ + \frac{\delta_{j_u+p}^{(b),i_q}}{\delta_{j_u+p}^{(b),i_{db}}} + \frac{\delta_{j_u+p}^{(b),i_{db}}}{\delta_{j_u+p}^{(b),i_q}} - \gamma_{j_u+p}^{(b),i_q} - \gamma_{j_u+p}^{(b),i_{db}} \quad (4.11)$$

où  $\gamma_{j_u+p}^{(b),i_q}$ ,  $\delta_{j_u+p}^{(b),i_q}$ , et (resp.  $\gamma_{j_u+p}^{(b),i_{db}}$ ,  $\delta_{j_u+p}^{(b),i_{db}}$ ) sont les paramètres de forme et d'échelle des densités  $f_{w_{j_u+p}^{(b),i_q}}$ , (resp.  $f_{w_{j_u+p}^{(b),i_{db}}}$ ) et  $\psi$  est la fonction digamma.

### Cas où les dépendances entre coefficients d'ondelettes sont considérées

Si les dépendances inter-spectrales et/ou inter-échelles sont considérées, la DSKL  $D(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}, f_{\tilde{\mathbf{w}}_{j_u}^{i_q}})$  entre les densités jointes  $f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}$  et  $f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}$  n'admet pas de forme explicite. Dans ce cas, nous proposons [155] de l'approximer empiriquement grâce une approche de Monte Carlo [154, 155]. Plus précisément, considérons une séquence aléatoire  $\mathbf{s}_1, \dots, \mathbf{s}_{L_{j_u}^{i_{db}}}$  de longueur  $L_{j_u}^{i_{db}}$  générée à partir du modèle de densité  $f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}$ , une approximation  $\hat{d}(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}} \| f_{\tilde{\mathbf{w}}_{j_u}^{i_q}})$  de la divergence de Kullback-Leibler entre les densités  $f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}$  et  $f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}$  est donnée par :

$$\hat{d}(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}} \| f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}) = \frac{1}{L_{j_u}^{i_{db}}} \sum_{\ell=1}^{L_{j_u}^{i_{db}}} \log(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}(\mathbf{s}_\ell)) - \log(f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}(\mathbf{s}_\ell)). \quad (4.12)$$

La divergence  $\hat{d}(f_{\tilde{\mathbf{w}}_{j_u}^{i_q}} \| f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}})$  est également estimée de la même manière en s'appuyant sur une seconde séquence aléatoire  $\mathbf{s}'_1, \dots, \mathbf{s}'_{L_{j_u}^{i_q}}$  de longueur  $L_{j_u}^{i_q}$  générée à partir du modèle de densité  $f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}$ .

Enfin, la DSKL entre les densités  $f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}$  et  $f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}$  est estimée par :

$$\hat{D}(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}, f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}) = \frac{1}{2} (\hat{d}(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}} \| f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}) + \hat{d}(f_{\tilde{\mathbf{w}}_{j_u}^{i_q}} \| f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}})) \quad (4.13)$$

Il est à noter que la précision de cette approximation dépend des longueurs  $L_{j_u}^{i_q}$  et  $L_{j_u}^{i_{db}}$  des séquences aléatoires  $\mathbf{s}_1, \dots, \mathbf{s}_{L_{j_u}^{i_{db}}}$  et  $\mathbf{s}'_1, \dots, \mathbf{s}'_{L_{j_u}^{i_q}}$ . En effet, plus le nombre de réalisations aléatoires est grand, meilleure est la précision, au prix d'une augmentation de la complexité opératoire. Un compromis doit alors être recherché entre précision et temps d'exécution de cette approximation. Il a été empiriquement établi que les longueurs  $L_{j_u}^{i_q}$  et  $L_{j_u}^{i_{db}}$  considérées sont de l'ordre de  $10^3$  [154].

Le choix de la mesure de similarité est couplé à un algorithme de recherche qui définit la manière dont la base des signatures est parcourue pour comparer les signatures qui y sont stockées à la celles de la requête. Cet algorithme doit assurer le maximum de précision selon un temps d'exécution raisonnable. Le coût élevé du calcul de la DSKL (proportionnel à  $10^3$ ) ainsi que la propriété d'échelonnabilité de la décomposition en ondelettes nous ont poussé à concevoir un algorithme rapide de recherche.

## 4.2 Conception d'un algorithme de recherche arborescente

Une mesure de similarité étant choisie, l'algorithme de recherche permet de parcourir la base des index afin de comparer les signatures des images de la base à celle de la requête. Etant donné un paramètre  $R$  choisi par l'utilisateur, les  $R$  images jugées comme étant les plus semblables à la requête sont alors choisies. La recherche exhaustive est la plus intuitive. Son principe consiste à comparer systématiquement la signature d'une image requête donnée à toutes les signatures de la base des index. Comme toutes les images de la bases sont considérées comme candidates, cette approche est clairement la plus précise. Néanmoins, dans le cas de bases d'images de grande taille, la recherche exhaustive peut s'avérer inadéquate en raison de la complexité opératoire rédhibitoire. En effet, pour une image requête donnée, la recherche des  $R$  meilleures images de la base nécessite  $R(\mathcal{N} - \frac{R}{2} + \frac{1}{2})$  calculs de la mesure de similarité et comparaisons où  $\mathcal{N}$  est la taille de la base d'images. Si on prend l'exemple d'une large base de 10000 images et  $R = 10$ , le nombre d'opérations de comparaison et de calculs de distance est de l'ordre de  $10 \times 10000 = 10^5$  opérations. Si en plus, on utilise la distance DSKL qui a une coût élevé en terme de calculs, cette recherche devient difficile à envisager sur le plan pratique d'autant plus que la recherche se fait en ligne en temps réel. Pour remédier à ce problème, plusieurs travaux essayent de concevoir des algorithmes rapides de recherche. Leur principe est de parcourir judicieusement la base des signatures pour ne calculer la similarité qu'avec certaines images de la base ce qui permet de réduire drastiquement le nombre de distances calculées. La sélection de ces images de la base se fait en général grâce à une structuration préalable de la base des index qui se fait hors ligne. Parmi les méthodes classiques d'organisation de la base des index, certaines procèdent à un découpage rectangulaire de l'espace comme le R-tree [156, 157] et le X-tree [158], d'autres à un découpage sphérique tel que le SS-tree [159] et le SR-tree [160]. Ainsi, l'algorithme de recherche est accéléré par

le parcours d'un arbre formé par différentes régions de l'espace des index au lieu de l'espace tout entier. Majdoub *et. al* [161] ont également conçu un algorithme d'indexation rapide qui repose sur une structuration hiérarchique de la base des index. Cette structuration est basée sur la classification des vecteurs résultant de la quantification des vecteurs des signatures par un réseau en treillis.

Dans cette thèse, nous proposons une autre alternative qui exploite la propriété d'analyse multirésolution des transformée en ondelette considérées pour concevoir un algorithme de recherche arborescente [162] reposant sur une organisation préalable de la base des index . Celle-ci correspond dans notre cas aux vecteurs descripteurs  $\mathbf{f}_j^{i_{db}}$  des images  $i_{db}$  de la base aux différents étages  $j = 1, \dots, J$ , traduisant les densités  $f_{\mathbf{w}_j}^{i_{db}}$  des vecteurs  $\mathbf{w}_j^{i_{db}}$ ,  $j = 1, \dots, \dots, J$ . Dans ce qui suit, nous décrivons les deux étapes de notre approche à savoir l'étape de l'organisation des signatures qui se fait hors-ligne et l'étape de recherche qui se fait en ligne.

### 4.2.1 Etape d'organisation des signatures

L'étape d'organisation de la base des index a pour principe de grouper les différentes signatures à chaque étage  $j$  en ayant recours à un algorithme donné de classification et ceci d'une manière emboîtée qui exploite la propriété d'analyse en multirésolution assurée par les transformations en ondelettes. Nous commençons par décrire le principe général de l'étape d'organisation des signatures. Ensuite, nous nous focalisons sur le choix de l'algorithme de classification.

#### Principe de l'organisation des signatures

Comme illustrée dans la figure 4.1, l'organisation des signatures se fait d'une manière récursive, en partant des échelles les plus grossières vers les échelles les plus fines et ceci de la manière suivante.

- A l'échelle la plus grossière,  $j = J$ , les vecteurs  $\mathbf{f}_J^{i_{db}}$  sont classés en un nombre prédéfini  $K_J$  de classes disjointes en utilisant un algorithme donné de classification relativement à la distance  $d$  choisie. A chaque classe  $k_J = 1, \dots, K_J$  correspondra un prototype (ou centroïde)  $\mathbf{p}_J[k_J]$ . Ces prototypes sont généralement virtuels car ils ne correspondent pas forcément à des images effectives de la base des index. Un vecteur d'entiers  $\mathbf{k}_{J-1}[k_J]$  est également choisi pour définir le nombre de sous-classes à l'échelle la plus fine.
- La classification des index à un étage  $1 \leq j < J$  se fait de la manière suivante. Pour chaque classe  $k_{j+1}$  des  $K_{j+1}$  classes de l'échelle la plus grossière  $j + 1$ , seules les images dont les index à l'échelle  $j + 1$  appartiennent à la classe  $k_{j+1}$  sont considérées dans la procédure de classification. Les vecteurs  $\mathbf{f}_j^{i_{db}}$  qui correspondent aux images concernées sont alors classifiés en utilisant le même algorithme de classification en  $\mathbf{k}_j[k_{j+1}]$  sous-classes disjointes relatives à la classe mère  $k_{j+1}$ . Notons que le vecteur  $\mathbf{k}_j$  est un vecteur



de dimension  $K_{j+1}$  dont les éléments correspondent aux nombres de sous-classes pour chaque classe-mère  $k_{j+1} = 1, \dots, K_{j+1}$ . Ainsi le nombre total de sous-classes à l'étage  $j$  est donné par

$$K_j = \sum_{k_{j+1}=1}^{K_{j+1}} \mathbf{k}_j[k_{j+1}]. \quad (4.14)$$

De nouveau, chaque classe  $k_j$  de ces  $K_j$  classes est associée à un centroïde virtuel  $\mathbf{p}_j[k_j]$  et la valeur de  $j$  est décrémentée d'une unité.

- La procédure continue jusqu'à  $j = 1$ .

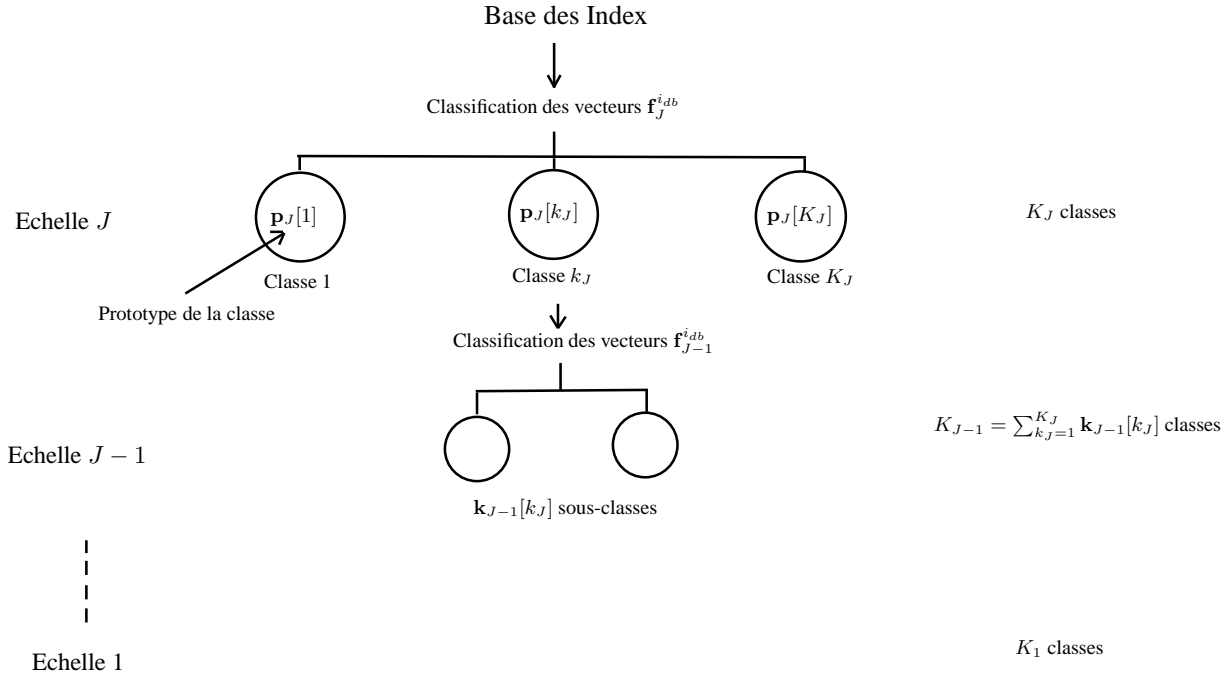


FIG. 4.1 – Organisation de la base des signatures.

### Choix de l'algorithme de classification

Il est évident que la rapidité et la pertinence de cette approche sont fortement liées au choix de l'algorithme de classification. Dans notre travail, nous avons choisi d'adopter deux algorithmes de classification différents : l'algorithme K-moyennes et l'Algorithme d'Agglomération Compétitive (AAC) [163]. Les principes de ces deux algorithmes sont décrits dans notre contexte (base des index) comme suit :

- *L'algorithme K-moyennes :*
- A l'échelle la plus grossière  $J$ , le nombre de classes  $K_J$  est prédéfini et  $K_J$  prototypes initiaux sont choisis comme étant des points arbitraires de la base des index, de

préférence dans les régions de la base des index les plus peuplées.

- $K_J$  classes initiales sont créées en attribuant chaque vecteur  $\mathbf{f}_j^{i,db}$  de la base des index à l'étage  $J$  au prototype le plus proche relativement à la distance choisie  $d$  (DEN ou DSKL). Chaque classe est alors formée par les points de la base des index qui sont plus proches d'un prototype donné qu'aux autres (règle du plus proche voisin). On a besoin de recalculer  $K_J$  nouveaux centroïdes, cette fois virtuels. Il s'agit de calculer le barycentre de chacune des classes obtenues relativement à la distance  $d$  choisie. Dans le cas de la DEN, le nouveau centroïde de chaque classe correspond à la moyenne arithmétique des points de la classes, dans le cas de la DSKL, le centroïde n'admet pas une forme explicite. Il peut être approximé [164] comme décrit dans l'annexe D. Les points de la base des index à l'étage  $J$  sont alors re-classifiés par rapport aux nouveaux centroïdes.
- Cette procédure se poursuit à l'étage  $J$  de manière itérative en changeant les centroïdes à chaque itération. La procédure s'arrête quand le changement des centroïdes cesse ou devient négligeable. On aura alors  $K_J$  classes finales caractérisées par leurs centroïdes  $\mathbf{p}_J[k_J]$ ,  $k_J = 1, \dots, K_J$ .
- La classification procède de la même manière pour les échelles suivantes à la différence qu'à un étage  $1 \leq j < J$ , seules les images dont les index à l'étage  $j+1$  appartiennent à une classe  $k_{j+1}$ ,  $k_{j+1} = 1, \dots, K_{j+1}$ , sont concernées par l'algorithme de classification. Ainsi, pour chaque classe  $k_{j+1} = 1, \dots, K_{j+1}$ , le nombre de sous-classes  $\mathbf{k}_j[k_{j+1}]$  est choisi, ensuite les vecteurs  $\mathbf{f}_j^{i,db}$  concernés sont classifiés en  $\mathbf{k}_j[k_{j+1}]$  sous-classes, en utilisant la même procédure décrite précédemment pour l'étage  $J$ .

La simplicité de l'algorithme K-moyennes ainsi que sa popularité nous ont poussé à l'appliquer dans notre procédure de recherche. Néanmoins pour cet algorithme, les nombres de classes  $K_J, \dots, K_1$  à chaque étage  $j$  ainsi que les nombres de sous-classes  $\mathbf{k}_j[k_{j+1}]$  sont choisis dès le départ d'une manière arbitraire. De plus ces nombres choisis au départ ne peuvent pas être mis à jour au cours de l'algorithme, cela pourrait aboutir à l'obtention de classes insignifiantes de faible effectif ce qui pourrait affecter la pertinence de l'étape de recherche décrite par la suite. Pour éviter ce problème, on a eu recours à l'Algorithme d'Agglomération Compétitive (AAC) [163] décrit ci-dessous.

- *L' Algorithme d'Agglomération Compétitive :*

Il s'agit d'un algorithme de classification non-supervisée *hiérarchique*, dont le principe est la création de classes successives à partir de classes pré-établies précédemment. C'est un algorithme *agglomératif*. En effet, il commence par attribuer les données à classer à un grand nombre initial de classes et enchaîne par des fusions successives de classes de faibles effectifs en des classes de plus en plus grandes. D'autre part, le AAC n'est pas un

algorithme *statique*. En effet, les points de l'espace appartiennent aux différentes classes d'une manière floue (c'est-à-dire suivant un degré d'appartenance). Le AAC a été utilisé pour la classification de vecteurs des descripteurs extraits à partir de la représentation spatiale d'images d'une base [163, 165]. Dans cette thèse, on propose de l'appliquer pour l'organisation des vecteurs extraits dans le domaine des ondelettes.

- A l'étage  $J$  le plus grossier, on prédéfinit le nombre maximal de classes  $K_{J,max}$  ainsi que  $K_{J,max}$  centroïdes initiaux  $\mathbf{p}_{J,init}[k_J]$  choisis d'une manière arbitraire. Ensuite, le AAC est appliqué aux vecteurs signatures  $\mathbf{f}_J^{i_{db}}$  dans le but d'obtenir un nombre final de classes  $K_J$  de manière à minimiser la fonction objective suivante :

$$G_J = \sum_{k_J=1}^{K_J} \sum_{i_{db}=1}^{\mathcal{N}} u_{k_J,i_{db}}^2 d^2(\mathbf{p}_J[k_J], \mathbf{f}_J^{i_{db}}) - \lambda \sum_{k_J=1}^{K_J} \left[ \sum_{i_{db}=1}^{\mathcal{N}} u_{k_J,i_{db}} \right]^2 \quad (4.15)$$

où  $\mathbf{p}_J[k_J]$  est le centroïde de la classe  $k_J$ ,  $d$  la fonction de similarité choisie,  $\mathcal{N}$  l'effectif de la base d'images et  $u_{k_J,i_{db}}$  est le degré d'appartenance de l'image  $i_{db}$  à la classe  $k_J$ . La fonction  $G_J$  est minimisée sous les conditions suivantes pour toute image  $i_{db}$  de la base :

$$\sum_{k_J=1}^{K_J} u_{k_J,i_{db}} = 1, \quad u_{k_J,i_{db}} \in [0, 1]. \quad (4.16)$$

Le degré d'appartenance  $u_{k_J,i_{db}}$  peut en fait s'écrire sous la forme de 2 termes  $u_{k_J,i_{db}} = u_{k_J,i_{db}}^{FCM} + u_{k_J,i_{db}}^{BIAIS}$ . Le premier terme prend en considération les distances des points par rapport à une classe  $k_J$  :

$$u_{k_J,i_{db}}^{FCM} = \frac{1/d^2(\mathbf{p}_J[k_J], \mathbf{f}_J^{i_{db}})}{\sum_{c_J=1}^{K_J} 1/d^2(\mathbf{p}_J[c_J], \mathbf{f}_J^{i_{db}})}. \quad (4.17)$$

Le second terme permet de réduire les cardinalités des grandes classes afin de ne garder que des classes significatives, il est donné par :

$$u_{k_J,i_{db}}^{BIAIS} = \frac{\lambda}{d^2(\mathbf{p}_J[k_J], \mathbf{f}_J^{i_{db}})} \left[ \mathcal{N}_{k_J} - \frac{\sum_{c_J=1}^{K_J} 1/d^2(\mathbf{p}_J[c_J], \mathbf{f}_J^{i_{db}}) \mathcal{N}_{c_J}}{\sum_{c_J=1}^{K_J} 1/d^2(\mathbf{p}_J[c_J], \mathbf{f}_J^{i_{db}})} \right] \quad (4.18)$$

où  $\mathcal{N}_{k_J}$  est l'effectif de la classe  $k_J$ .

Il est également intéressant de noter que le critère  $G_J$  combine deux termes qui se complètent. Le premier terme correspond à la fonction objective de l'algorithme de classification Fuzzy C-means (FCM) [166] et atteint son minimum quand chaque point appartient à une classe séparée. A l'opposé, le second terme atteint son minimum quand tous les points appartiennent à une même classe, il contrôle ainsi le nombre de classes. Le paramètre  $\lambda$  a pour rôle d'assurer un équilibre entre ces deux termes. Il décroît au cours des itérations pour faciliter l'agglomération des classes.

- Pour les étages  $1 \leq j < J$ , la classification se fait de la même manière pour la génération de sous-classes à partir des classes mères à l'étage  $j + 1$ . Pour chaque

classe  $k_{j+1} = 1, \dots, K_{j+1}$ , le nombre maximal de sous-classes  $\mathbf{k}_{j,max}[k_{j+1}]$  est choisi. Ensuite, les vecteurs  $\mathbf{f}_j^{i_{db}}$  concernés sont classifiés par le AAC en  $\mathbf{k}_j[k_{j+1}]$  sous-classes.

### 4.2.2 Etape de recherche hiérarchique

Après avoir organisé la base des index, la procédure de recherche, illustrée par la figure 4.2, est itérative et hiérarchique. Elle commence à l'échelle la plus grossière  $J$  et s'arrête à un étage  $j_u$  choisi par l'utilisateur et ceci de la manière suivante.

- A l'échelle la plus grossière  $J$ , le vecteur des signatures  $\mathbf{f}_J^{i_q}$  de l'image requête est comparé aux prototypes  $\mathbf{p}_J[k_J]$  des classe  $k_J = 1, \dots, K_J$ . Ensuite, la meilleure classe choisie  $k_J^{*,i_q}$  est celle dont le centroïde  $\mathbf{p}_J[k_J^{*,i_q}]$  est le plus proche du vecteur des signatures  $\mathbf{f}_J^{i_q}$  relativement à la distance  $d$  choisie.
- A un étage  $j_u \leq j < J$ , le vecteur signature  $\mathbf{f}_j^{i_q}$  est comparé seulement aux prototypes des sous-classes de la classe-mère  $k_{j+1}^{*,i_q}$  choisie à l'échelle la plus grossière  $j + 1$ . Parmi ces dernières, la meilleure sous-classe  $k_j^{*,i_q}$  est choisie de manière à avoir le centroïde le plus proche de l'index de la requête.
- Si  $j = j_u$ , la meilleure sous-classe  $k_{j_u}^{*,i_q}$  est sélectionnée et le système délivre en sortie les  $R$  meilleures images dont les index appartiennent à cette dernière sous-classe, et vérifiant la condition suivante :

$$i_{db}^* = \arg \min_{i_{db}, \mathbf{f}_{j_u}^{i_{db}} \in k_{j_u}^{*,i_q}} d(f_{\tilde{\mathbf{w}}_{j_u}^{i_{db}}}, f_{\tilde{\mathbf{w}}_{j_u}^{i_q}}). \quad (4.19)$$

### 4.2.3 Avantages de notre approche

- Il est évident d'après ce qui précède que notre approche permet de gagner en temps d'exécution et en complexité opératoire par rapport à la recherche exhaustive vu que la signature de la requête à chaque étage  $j$  ne sera pas comparée à toutes les signatures de la base des index mais seulement à celles appartenant à la sous-classe  $k_{j_u}^{*,i_q}$  choisie.
- Notre approche permet d'exploiter la propriété d'échelonnabilité de la décomposition en multirésolution en assurant une recherche arborescente qui permet à l'utilisateur de raffiner sa recherche en partant à l'échelle la plus grossière  $J$  et s'arrêtant à échelle  $j_u$  lorsqu'il est satisfait.
- La flexibilité dans le choix de l'algorithme de classification permet de choisir l'algorithme qui assure la meilleure répartition de la base des signatures ce qui influencerait les performances de la recherche en diminuant le nombre d'images pertinentes (semblables à le requête) et ignorées dans la procédure de comparaison.

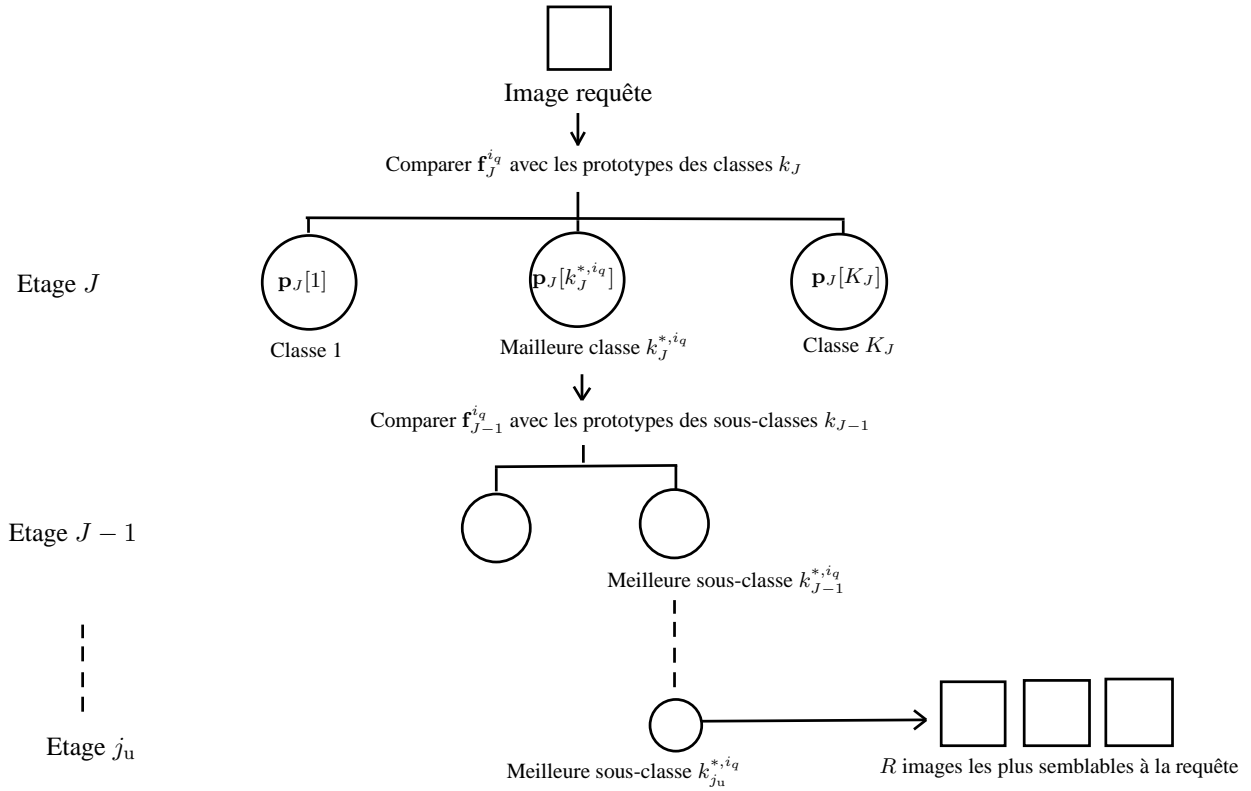


FIG. 4.2 – Procédure de recherche arborescente.

### 4.3 Conclusion

Dans ce chapitre, nous avons présenté la seconde partie de nos contributions qui se rapporte à l'étape de la procédure de recherche. Dans une première partie, nous avons présenté les mesures de similarité les plus utilisées en indexation d'images et nous avons souligné que la DSKL comme étant une distance mieux adaptée en terme de précision que la DEN au type de signature que nous avons proposé. Dans une seconde partie, nous avons expliqué le principe de l'algorithme de recherche arborescente que nous avons mis en œuvre dans le but de minimiser la complexité opératoire ainsi que d'exploiter la propriété d'échelonnabilité assurée par la décomposition en ondelettes. Nous avons également insisté sur l'importance du choix de l'algorithme de classification pour l'étape d'organisation des index.

Dans le prochain chapitre, nous essayerons de valider nos contributions par des simulations réalisés sur des bases d'images satellitaires multispectrales.

# Chapitre 5

## Résultats expérimentaux

Ce chapitre est consacré à la partie expérimentale. En effet, dans ce chapitre, nous essayons de valider nos contributions, développées sur le plan théorique dans les chapitre 3 et 4, par des simulations conduites sur des bases d’images multispectrales SPOT. Dans la section 5.1, nous définissons le cadre expérimental en présentant les bases d’apprentissage et de test, les critères d’évaluation des performances, ainsi que les travaux de références auxquels nous nous comparons le long de ce chapitre. Les sections 5.2 et 5.3 sont consacrées à la validation de nos contributions dans l’étape d’extraction des signatures. Dans la section 5.2, nous retenons comme décomposition multirésolution la transformée en ondelettes S-TOD, et nous étudions l’apport sur les performances de notre modèle multivarié reposant sur la théorie des copules. Dans la section 5.3, nous étudions l’influence du choix de la décomposition multirésolution sur les performances de la recherche.

Dans les sections 5.4 et 5.5, nos contributions concernant la partie procédure de recherche sont testées sur le plan expérimental. Dans la section 5.4, nous étudions l’influence du choix de la mesure de similarité sur les performances de la recherche, tandis que dans la section 5.5, notre algorithme de recherche arborescente est évalué en termes de performances de la recherche et de la complexité opératoire. Enfin, certaines conclusions sont faites dans la section 5.6.

### 5.1 Cadre expérimental

#### 5.1.1 Présentation des bases d’apprentissage et de test

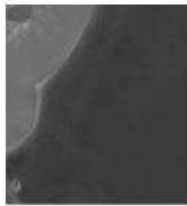
Pour nos simulations, nous utilisons la plateforme d’images satellitaires formées par les bases d’apprentissage et de test suivantes.

- La base d’apprentissage contient 2704 images satellitaires SPOT4 de tailles  $256 \times 256 \times 4$  ( $B = 4$ ) correspondant à différentes régions de la Tunisie. Pour chaque image de la base, la base contient aussi les versions issues de rotation d’un angle prenant 12 valeurs différentes ( $\pm 15^\circ, \pm 30^\circ, \pm 45^\circ, \pm 60^\circ, \pm 75^\circ, \pm 90^\circ$ ). Les images de la base appartiennent

visuellement aux 9 catégories suivantes :

- eau,
- zone urbaine,
- montagne,
- zone rurale,
- zone urbaine et plan d'eau,
- zone mixte urbaine et rurale,
- zone rurale et montagne,
- zone rurale et plan d'eau,
- montagne et plan d'eau.

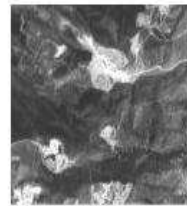
Un exemple d'image de chaque catégorie est donné dans la figure 5.1 où seule la première composante de chaque image est affichée.



eau



zone urbaine



montagne



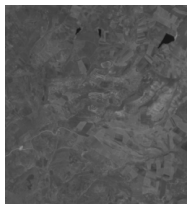
zone rurale



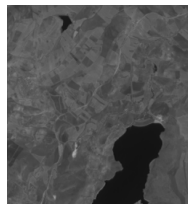
zone urbaine et eau



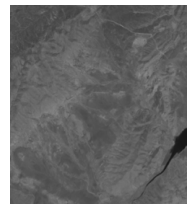
zone urbaine et rurale



zone rurale et montagne



zone rurale et eau



Montagne et eau

FIG. 5.1 – Différentes catégories d'images dans la base d'apprentissage.

- La base de test contient 208 images appartenant à la base d’apprentissage et aux différentes catégories. Il est à noter que la base de test ne contient pas des versions d’une même image à des angles différents.

Dans ce contexte, nous soulignons que nous sommes bien conscients que les tailles des bases d’images utilisées sont bien faibles au regard des tailles que peuvent atteindre certaines bases d’images (plusieurs millions). En effet, nous avons rencontré une difficulté handicapante : l’obtention des bases d’images multispectrales se fait moyennant une contrepartie financière conséquente auprès des fournisseurs d’images satellitaires qui dépasse les capacités financières de notre équipe de recherche.

### 5.1.2 Critères d’évaluation de performances

Pour évaluer les performances de recherche, plusieurs critères sont utilisés dans les travaux d’indexation d’images tels que la précision, le rappel et le rang de recherche modifié normalisé moyen (RRMNM) pour *Average Normalized Modified Retrieval Rank* retenu dans la norme d’indexation d’images MPEG7.

Dans ce travail, pour une image requête  $i_q$ , nous utilisons la précision ( $PR_{i_q}$ ) et le rappel ( $RC_{i_q}$ ) définis par :

$$PR_{i_q} = \frac{R_{i_q}}{R}, \quad RC_{i_q} = \frac{R_{i_q}}{R_{t,i_q}} \quad (5.1)$$

où

- $R$  est le nombre total d’images retournées par le système de recherche.
- $R_{i_q}$  le nombre d’images retournées considérées comme pertinentes par rapport à l’image requête  $i_q$  selon la vérité-terrain.
- $R_{t,i_q}$  le nombre total d’images pertinentes par rapport à l’image requête  $i_q$  qui se trouvent dans la base d’apprentissage.

Dans notre cas, une image est considérée comme pertinente si elle appartient à la même catégorie que la requête  $i_q$ . Dans nos simulations, nous calculons les valeurs moyennes  $PR$  et  $RC$  de précision et rappel calculées sur toutes les images  $Q$  de la base de test pour avoir une idée plus significative des performances de la recherche :

$$PR = \frac{1}{Q} \sum_{i_q=1}^Q PR_{i_q}, \quad RC = \frac{1}{Q} \sum_{i_q=1}^Q RC_{i_q}. \quad (5.2)$$

### 5.1.3 Travaux de références

Pour mieux mettre en évidence nos contributions, nous avons comparé celles-ci aux travaux de recherche rapportés dans la littérature s’inscrivant dans le même contexte applicatif à savoir la recherche par le contenu dans le domaine transformé en ondelettes. Les méthodes en question sont les suivantes.



- L’approche univariée qui repose sur l’utilisation de la distribution gaussienne généralisée (DGG) pour les lois marginales.
- L’approche multivariée basée sur l’utilisation de la distribution gaussienne généralisée multivariée (DGGM) pour modéliser la distribution conjointe des coefficients d’ondelettes.
- L’approche basée sur la réduction des dépendances inter-spectrales et/ou inter-échelles par une Analyse en Composantes Indépendantes (ACI).

Tous ces travaux de référence seront implémentés en utilisant la transformée 5/3 appliquée séparément à chaque composante de l’image multispectrale.

Dans la section suivante, nous nous limitons à l’étude de l’influence des modèles conçus pour la caractérisation des coefficients d’ondelettes multivariés. Ainsi,  $J$  étages de la transformation 5/3 sont appliqués séparément aux  $B = 4$  composantes des images des bases d’apprentissage et de test. On prend  $J = 4$  pour l’étude de l’influence des dépendances inter-échelle et  $J = 3$  pour le reste des simulations. Dans cette première phase, la recherche exhaustive est choisie comme algorithme de recherche et la distance euclidienne normalisée (DEN) comme mesure de similarité.

## 5.2 Performances du modèle multivarié reposant sur la théorie des copules

### 5.2.1 Choix des densités marginales

Nous voulons comparer les performances de la recherche dans les deux cas où la DGG (distribution gaussienne généralisée) ou la DG (distribution gamma) est choisie comme modèle pour les distributions marginales des composantes  $w_j^b(m, n)$ . La méthode univariée est appliquée c’est-à-dire que les dépendances inter-spectrales et inter-échelles entre coefficients d’ondelettes sont ignorées. La totalité des 4 composantes ( $B = 4$ ) ainsi que trois étages de coefficients d’ondelettes ( $J = 3, j_u = 1$ ) sont considérés où  $j_u$  désigne l’échelle de résolution à laquelle s’arrête la recherche. La figure 5.2 montre les performances en termes de précision-rappel obtenues avec l’approche univariée pour des distributions marginales modélisées par des DGG ou des DG. La figure 5.2 prouve que choisir une DG pour les lois marginales assure de meilleures performances que la DGG. Ceci justifie le choix de la DG comme modèle des densités marginales dans la plupart des simulations qui suivront.

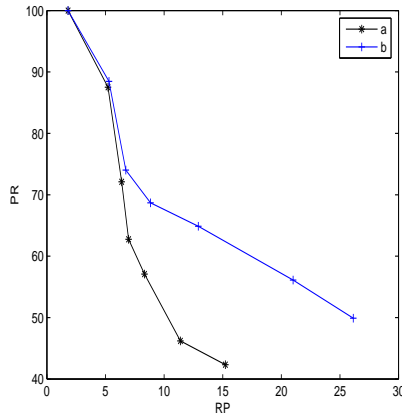


FIG. 5.2 – Courbes précision-rappel obtenues avec l’approche univariée **a** : densités marginales modélisées par des DGG **b** : densités marginales modélisées par des DG.

### 5.2.2 Prise en compte des dépendances inter-spectrales

**Mérite de l’approche multivariée :** Comme cela a été expliqué dans le chapitre 3, il s’avère judicieux d’exploiter les dépendances inter-spectrales qui existent entre les différentes composantes  $w_j^b(m, n)$  des coefficients d’ondelettes. La figure 5.3 montre, en effet, qu’une approche multivariée qui prend en considération les dépendances inter-spectrales assure de meilleures résultats que l’approche univariée. Pour le prouver, on a considéré l’exemple de la modélisation de la densité conjointe  $f_{\mathbf{w}_j}$  du vecteur  $\mathbf{w}_j(m, n)$ , défini par l’équation (3.2), par une DGGM [120].

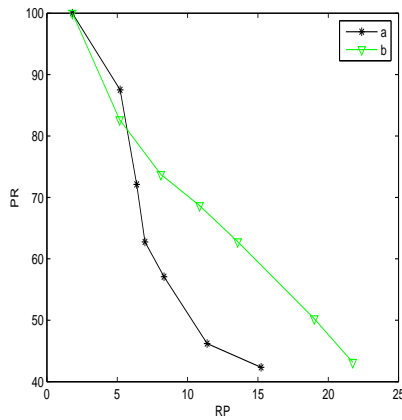


FIG. 5.3 – Gain en performances dû à l’utilisation d’une approche multivariée **a** : approche univariée avec des densités marginales modélisées par des DGG, **b** : densité conjointe modélisée par une DGGM.

**Apport de notre première contribution :** Dans la figure 5.4, nous illustrons l'apport de notre première contribution qui consiste à établir un modèle multivariée de la densité conjointe  $f_{\mathbf{w}_j}$  du vecteur  $\mathbf{w}_j(m, n)$  associé à une copule gaussienne et à des densités marginales DG. Notre approche est comparée aux trois travaux de référence suivants :

- l'approche univariée avec l'utilisation de la DG pour la modélisation des distributions marginales ;
- la décorrélation des  $B = 4$  composantes par une ACI, suivie de l'application de l'approche univariée ;
- la modélisation de la densité conjointe  $f_{\mathbf{w}_j}$  par une DGGM.

La figure 5.4 montre que notre approche fournit nettement les meilleures performances en termes de précision et rappel.

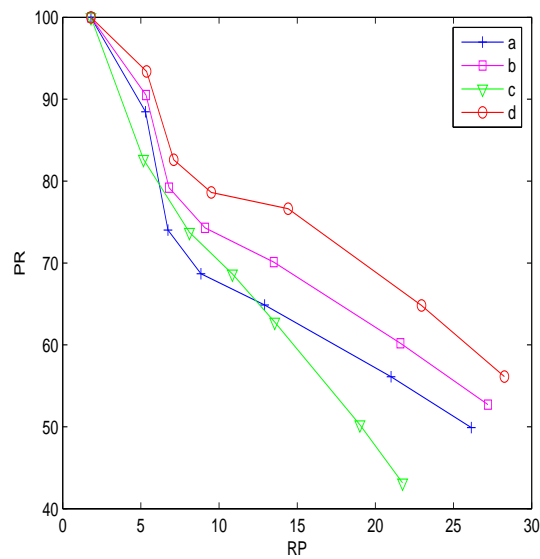


FIG. 5.4 – Apport en terme de rappel-précision de notre première contribution : **a** : approche univariée avec des densités marginales modélisées par des DG, **b** : approche basée sur l'utilisation d'une ACI, **c** : approche multivariée utilisant une DGGM, **d** : dépendances inter-spectrales modélisées par une copule gaussienne.

### 5.2.3 Prise en compte simultanée des dépendances inter-spectrales et inter-échelles

Dans le chapitre 3, nous avons montré que la théorie de copule nous a permis de modéliser simultanément les distributions marginales, les dépendances inter-spectrales et les dépendances inter-échelles. Pour valider la gain en précision-rappel, nous avons comparé les performances de recherche obtenues avec les 4 cas suivants où les distributions marginales sont toujours modélisées par des DG.

- Approche univariée : Seules les distributions marginales sont considérées. Les dépendances inter-spectrales et inter-échelles sont ignorées.
- Les dépendances inter-spectrales sont modélisées par des copules gaussiennes et les dépendances inter-échelles sont ignorées.
- Les dépendances inter-échelles sont modélisées par une copule gaussienne et les dépendances inter-spectrales sont ignorées.
- Les dépendances inter-échelles entre toutes les échelles de décomposition considérées, ainsi que les dépendances inter-spectrales à chaque échelle de résolution sont modélisées par des copules gaussiennes.

Dans cette partie des simulations, nous considérons les trois premières composantes spectrales ( $B = 3$ ). Ainsi que  $J = 4$  étages de décomposition en ondelettes. Le choix d'un grand nombre d'échelles de résolution est justifié par une meilleure mise en évidence de l'importance des dépendances inter-échelles. La figure 5.5 illustre les courbe précision-rappel obtenues avec les 4 cas cités précédemment pour deux valeurs différentes de  $j_u$  :  $j_u = 2$  et  $j_u = 1$  qui correspondent respectivement à un nombre d'échelles considérées  $J - j_u + 1 = 3$  et  $J - j_u + 1 = 4$ .

La figure 5.5, montre que la prise en compte simultanée des dépendances inter-canaux et inter-échelles assure les meilleures performances de cette recherche. Néanmoins, en général, l'amélioration due à la prise en compte des dépendances inte-échelles est moins importante que celle correspondant aux dépendances inter-canaux. Ceci s'explique par l'importance des similarités entre les différentes composantes d'une image multispectrale. l'apport des dépendances inter-échelles, est toutefois plus évident quand le nombre d'échelles considérées est plus grand.

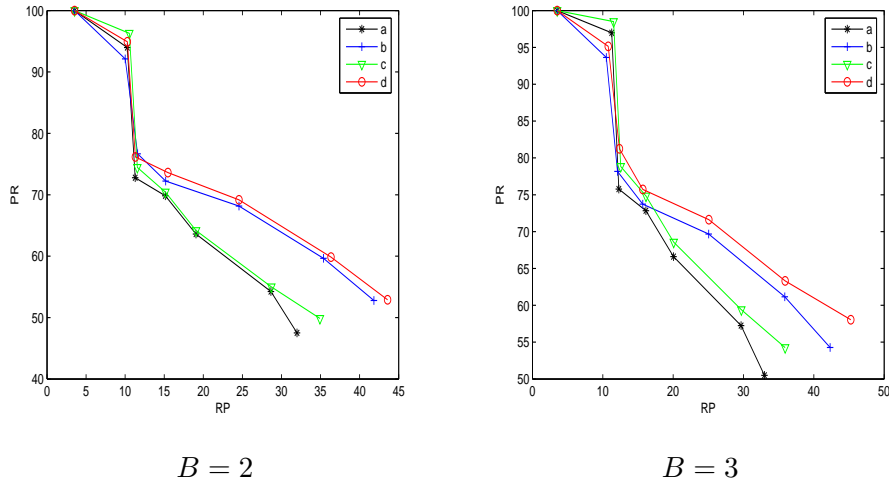


FIG. 5.5 – Apport en rappel-précision de la prise en compte simultanée des dépendances inter-canaux et inter-échelles : **a** : dépendances inter-spectrales et inter-échelles ignorées, **b** : dépendances inter-spectrales modélisées par une copule gaussienne et dépendances inter-échelles ignorées, **c** : dépendances inter-échelles modélisées par une copule gaussienne et dépendances inter-spectrales ignorées, **d** : prise en compte simultanée des dépendances inter-spectrales et inter-échelles.

## 5.2.4 Influence du choix de la copule sur les performances

### Comparaison de plusieurs familles de copules

Dans ce qui précède, nous nous sommes limité à l'utilisation de la copule gaussienne pour modéliser les dépendances inter-spectrales et inter-échelles. Or, celle-ci n'est pas forcément la copule idéale pour traduire les propriétés statistiques des dépendances entre canaux et/ ou entre échelles. Ainsi, nous avons proposé de comparer les performances de la recherche données avec 25 combinaisons de choix de familles de copules pour modéliser les dépendances inter-spectrales (modélisées par les copules  $\overline{C}_{j_u+p}$ ,  $p = 1, \dots, J - j_u + 1$ ) et inter-échelles (modélisées par la copule  $\tilde{C}_{j_u}$ ) et ceci pour  $B = 3$  (les 3 premières composantes sont considérées),  $J = 3$ ,  $j_u = 1$  (3 échelles de décompositions sont considérées) et  $R = 20$ . Pour les distributions marginales, on a retenu la DG. Les familles de copules testées sont :

- la copule égale à 1 (les dépendances inter-canaux/inter-échelles sont ignorées) ;
- la copule gaussienne ;
- la copule de Student ;
- la copule de Gumbel ;
- la copule de Clayton.

Les tableaux 5.1 et 5.2 donnent les valeurs de précision et de rappel moyens (en pourcentage)

pour les 25 combinaisons de familles de copules prédéfinies.

TAB. 5.1 – Précision moyenne en pourcentage obtenue avec différents couples de familles de copules traduisant les dépendances inter-canaux (en colonnes) et inter-échelles (en lignes).

Couple	1	gaussienne	Student	Clayton	Gumbel
1	57.40	88.87	89.03	58.93	59.11
gaussienne	57.61	88.87	89.03	58.95	59.12
Student	59.79	89.06	<b>89.22</b>	60.61	60.60
Clayton	57.66	88.87	89.03	58.95	59.11
Gumbel	57.57	88.87	89.03	58.99	59.19

TAB. 5.2 – Rappel moyen en pourcentage obtenu avec différents couples de familles de copules traduisant les dépendances inter-canaux (en colonnes) et inter-échelles (en lignes).

Couple	1	gaussienne	Student	Clayton	Gumbel
1	24.41	42.86	43.03	24.77	24.88
gaussienne	24.48	42.86	43.03	24.78	24.89
Student	25.79	43.08	<b>43.20</b>	26.11	26.11
Clayton	24.50	42.86	43.03	24.78	24.88
Gumbel	24.47	42.86	43.03	24.80	24.93

D’après ces deux tableaux, on remarque que le choix des familles des copules influe considérablement sur les performances de la recherche. Dans notre cas, la combinaison Student/Student assure les meilleures performances. Cependant, ce résultat peut varier selon les bases d’images d’apprentissage et/ou de test ou même selon le nombre d’étages considérés. D’autre part, pour ces résultats, on s’est contenté de prédéfinir à chaque fois une combinaison de familles de copules pour toutes les images de la base et à tous les étages et ensuite de tester les performances obtenues avec la combinaison choisie sans se poser la question de la sélection des meilleures familles de copules inter-spectrale et inter-échelle pour chaque image des bases d’apprentissage et de test.

### Approche bayésienne

Pour les raisons évoquées précédemment, nous nous sommes placés dans cette partie des simulations dans le cadre de l’estimation bayésienne présentée dans le chapitre 3 pour déterminer les “meilleures” familles de copules pour chaque image de la base et à chaque échelle. Pour simplifier, nous nous sommes contenté de tester l’approche bayésienne dans le cas de

la seule prise en compte des dépendances inter-canaux (les dépendances inter-échelle étant ignorées). Nous nous sommes également limité au cas de familles de copules ayant un paramètre  $\theta$  scalaire et ceci pour les 2 cas où  $B = 2$  (cas bivarié décrit dans [140], dans ce cas on considère les deux premières composantes d'une image) et  $B = 3$  (généralisation décrite dans l'annexe A). Dans le cas bivarié, l'approche bayésienne choisit la meilleure copule parmi la liste des familles de copules  $\{C_{l,\theta_l}\}_{l=L=5}$  formée par les 5 copules suivantes : Gumbel, Clayton, Frank, Ali-Mikhail-Haq (AHM) et la copule gaussienne, où les 4 premières sont des copules archimédiennes. Dans le cas  $B = 3$ , la copule gaussienne est écartée de la sélection car elle est caractérisée par 3 coefficients de corrélation. Dans les 2 cas ( $B = 2$  et  $B = 3$ ), 2 étages de décomposition sont considérés ( $J = 3, j_u = 2$ ).

Il est important de souligner que les familles de copules sélectionnées par cette approche peuvent différer d'une image de la base à une autre et également d'une échelle à une autre pour une même image. Ainsi les signatures d'une image de la base d'apprentissage ne sont comparées à celles d'une image requête que si les deux images ont les mêmes familles de copules inter-spectrales pour les 2 étages de décomposition  $j = J = 3$  et  $j = j_u = 2$ .

La figure 5.6 illustre l'amélioration en performances de la recherche obtenue grâce à l'utilisation de l'approche bayésienne par rapport à différents cas où la famille de copules inter-spectrales est prédéfinie.

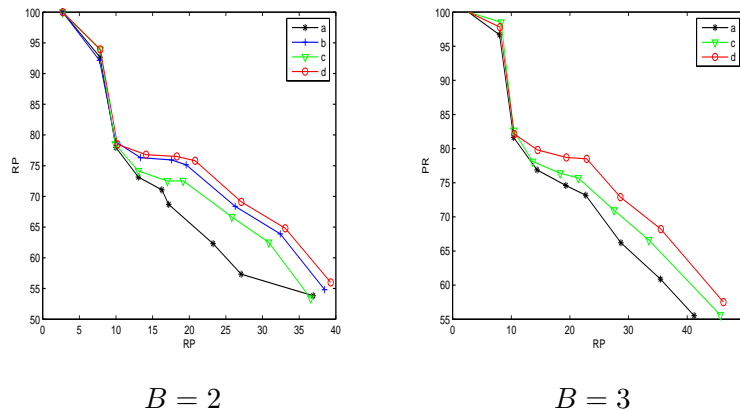


FIG. 5.6 – Apport en précision-rappel de l'approche bayésienne : **a** : dépendances inter-spectrales ignorées, **b** : dépendances inter-spectrales modélisées par une copule gaussienne, **c** : dépendance inter-spectrales modélisées par une copule de Gumbel, **d** : approche bayésienne.

## 5.3 Influence du choix de la décomposition

Si dans les simulations précédentes nous nous sommes limité à la décomposition 5/3 appliquée séparément à chaque composante des images des bases d'apprentissage et de test, nous essayons de cette partie de montrer l'influence de l'utilisation d'autres décompositions multirésolutions sur les performances de recherche. Soulignons que la recherche dans la base se fait toujours de manière exhaustive et la DEN reste le critère de similarité adopté.

### 5.3.1 Décomposition en schéma de lifting vectoriel

Les propriétés intéressantes du schéma de lifting vectoriel, en particulier, l'habilité à exploiter, simultanément, les redondances spectrales et spatiales, nous ont motivé à étudier l'apport de cette décomposition sur les performances de la recherche. Pour cela, on a choisi d'appliquer sur 3 étages ( $J = 3$ ), un schéma de lifting vectoriel (transformation DEC3 décrite dans l'annexe C), aux trois premières composantes des images d'apprentissage ( $B = 3$ ) et de test.

La figure 5.7 compare les performances obtenues avec le schéma de lifting conventionnel (transformée 5/3) et le schéma de lifting vectoriel (transformée DEC3) dans le deux cas suivants.

- Seules des distributions marginales sont modélisées par des DG (approche univariée).
- Les dépendances inter-spectrale sont modélisées par une copule gaussienne (notre approche multivariée).

Nous en déduisons les conclusions suivantes.

- Une amélioration considérable des performances est apportée par le schéma de lifting vectoriel dans les deux cas de l'approche univariée et multivariée. Ceci est attendu vu que le lifting vectoriel exploite les dépendances inter-spectrales dès l'étape de la décomposition.
- Dans la cas du schéma de lifting vectoriel, l'apport de l'utilisation de la copule pour modéliser les dépendances inter-canaux est minime. Ceci s'explique par le fait que les sous-bandes de coefficients d'ondelettes relatives à différents canaux spectraux présentent des redondances spectrales résiduelles faibles en raison du schéma de lifting vectoriel.



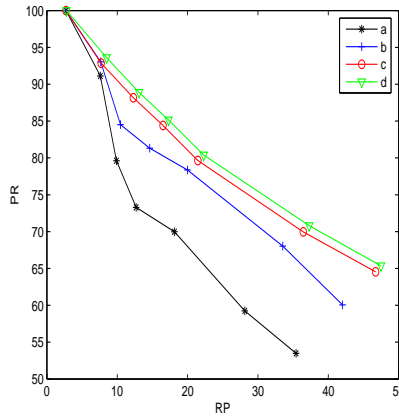


FIG. 5.7 – Comparaison des performances en terme de précision-rappel obtenues avec la décomposition 5/3 et la décompositions DEC3 : **a** : décomposition 5/3 suivie de l’approche univariée, **b** : décomposition 5/3 et dépendances inter-spectrales modélisées par une copule gaussienne, **c** : décomposition DEC3 suivie de l’approche univariée, **d** : décomposition DEC3 et dépendances inter-spectrales modélisées par une copule gaussienne.

### 5.3.2 Apport de la décomposition 2D TOM-AD et invariance par rotation

Dans tout ce qui précède, les signatures représentant les paramètres des modèles adoptés ne sont pas forcément invariantes par rapport à la rotation de l’image. Comme nous l’avons expliqué dans le chapitre 3, nous avons défini des signatures invariantes par rotation extraites à partir d’une décomposition de l’image par une transformation  $M$ -bandes en arbre dual 2D TOM-AD. Dans cette partie des simulations nous essayons de valider la pertinence de ce choix. Pour cela, nous avons proposé de comparer les performances de la recherche obtenues avec 3 étages ( $J = 3, j_u = 1$ ) de 4 décompositions multirésolutions différentes, décrites dans le chapitre 2 : (la S-TOD, la TOC-AD, les pyramides orientables et la TOM-AD). Dans le cas de la décomposition en pyramide orientable, le nombre d’orientations est choisi comme étant  $O = 6$ . Dans ce cas, le facteur de redondance est égal à  $4O/3 = 8$ . La TOM-AD est testée pour deux valeurs différentes du paramètre  $M$  :  $M = 2$  et  $M = 3$ . Dans les 4 cas, les vecteurs des signatures invariantes par rotation sont calculés comme la moyenne des vecteurs des signatures relatifs aux différentes orientations. L’approche d’extraction des signatures à partir des sous-bandes est l’approche multivariée qui repose sur la modélisation des dépendances inter-spectrales par une copule gaussienne. Les dépendances inter-échelles sont ignorées.

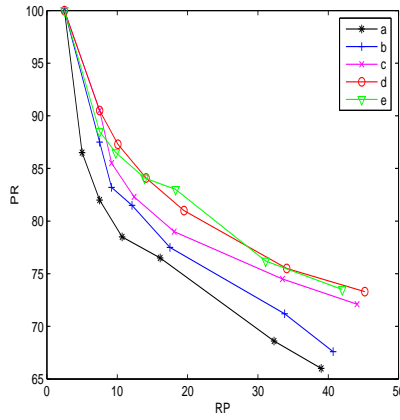


FIG. 5.8 – Comparaison des performances en terme de précision-rappel obtenues par différentes décompositions multirésolutions : **a** : S-TOD, **b** : TOC-AD, **c** : TOM-AD,  $M = 2$ , **d** : TOM-AD,  $M = 3$ , **e** : Décomposition en pyramides orientables,  $O = 6$ .

La figure 5.8 illustre les performances obtenues pour des signatures invariantes par rotation calculées à partir des différentes décompositions multirésolutions que nous venons de citer.

L'examen des courbes de performances nous conduit à faire les remarques suivantes.

- Toutes les transformations multirésolution directionnelles aboutissent à de meilleures performances que celles obtenues par la S-TOD. Ceci est attendu vu la faible directionnalité de la S-TOD qui devient une limitation réelle dans le cas de l'invariance par rotation.
- Bien que les pyramides orientables aboutissent généralement aux meilleures performances, elles présentent l'inconvénient d'avoir un facteur de redondance élevé ce qui est contraignant dans le contexte de la compression.
- La 2D TOM-AD réalise le meilleur compromis entre performances (proches de celles du pyramide orientables dans le cas où  $M = 3$ ) et le facteur de redondance qui reste toujours acceptable (égal à 2) quel que soit le choix de  $M$ .
- Les performances s'améliorent quand la valeur de  $M$  augmente et toujours à facteur de redondance constant.

## 5.4 Influence du choix de la distance

Dans tout ce qui précède, nous avons utilisé la distance euclidienne normalisée (DEN) pour comparer les signatures (paramètres des modèles statistiques mis en œuvre). Néanmoins, comme l'on a vu dans le chapitre 4, la distance symétrique de Kullback-Leibler (DSKL) est plus adaptée à comparer des distributions statistiques. Pour cette raison, nous avons comparé les performances de la recherche obtenues avec la DEN et la DSKL et ceci dans les différents cas où les distributions marginales sont modélisées par des DG ou des DGG et où

les dépendances inter-spectrales sont ignorées ou modélisées par une copule gaussienne. Pour simplifier, nous avons ignoré dans cette partie les dépendances inter-échelles. Il est à rappeler que le calcul de la DSKL se fait en utilisant sa forme explicite dans le cas univarié et en ayant recours à la méthode de Monte Carlo dans le cas où la copule gaussienne est considérée. Ainsi, vu la complexité opératoire accrue de la DSKL, nous avons utilisé seulement 1000 images de la base d'apprentissage qui appartiennent toujours aux différentes catégories présentées dans la section 5.1.

La figure 5.9 montre les courbes précision-rappel pour les différents cas suivants.

- **a.** Les lois marginales sont modélisées par des DGG, les dépendances inter-spectrales sont ignorées et la DEN est utilisée.
- **b.** Les lois marginales sont modélisées par des DGG, les dépendances inter-spectrales sont ignorées et la DSKL est utilisée. Cette méthode correspond en fait à la méthode de Do et Vetterli [46].
- **c.** Les lois marginales sont modélisées par des DG, les dépendances inter-spectrales sont ignorées et la DEN est utilisée.
- **d.** Les lois marginales sont modélisées par des DG, les dépendances inter-spectrales sont ignorées et la DSKL est utilisée.
- **e.** Les lois marginales sont modélisées par des DG, les dépendances inter-spectrales sont modélisées par une copule gaussienne et la DEN est utilisée.
- **f.** Les lois marginales sont modélisées par des DG, les dépendances inter-spectrales sont modélisées par une copule gaussienne et la DSKL est utilisée.

On remarque alors aussi que quels que soient les modèles utilisés pour les distributions marginales et la copule, la comparaison des signatures correspondantes par la DSKL assure de meilleures performances que la comparaison par la DEN. Ceci a été déjà établi par Do et Vetterli [46] dans le cas de l'approche univariée et quand la DGG est utilisée pour modéliser les distributions marginales. Pour notre part, nous avons généralisé ce résultat dans le cas multivarié s'appuyant sur la théorie des copules, les distributions marginales étant les DG ou les DGG.

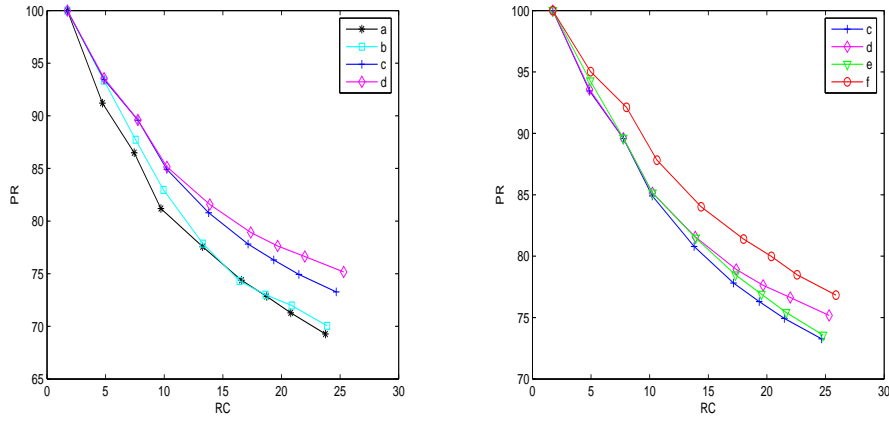


FIG. 5.9 – Courbes précision-rappel obtenues avec  $B = 3$ ,  $J = 3$ ,  $j_u = 1$  pour différents modèles paramétriques et différentes distances, (à gauche) modèles DG et DGG de lois marginales et (à droite) modèle DG pour les lois marginales.

## 5.5 Apport de l’algorithme de recherche arborescente

### 5.5.1 Performances et complexité opératoire de la recherche arborescente

Comme on l’a vu dans le chapitre 4, malgré sa simplicité et sa fiabilité, la recherche exhaustive engendre une complexité opératoire accrue. Dans le cas où la DSKL est utilisée comme distance, cela devient davantage compliqué puisque l’implémentation de la DSKL selon la méthode de Monte Carlo présente un coût de calcul élevé. La combinaison DSKL-recherche exhaustive n’est alors pas adéquate du point de vue pratique.

Pour remédier à ce problème, nous gardons toujours la DSKL comme mesure de similarité et nous utilisons comme algorithme de recherche, la recherche arborescente (décrite en détail dans 4.2) sur 3 étages de décomposition ( $J = 3, j_u = 1$ ). En premier lieu, nous utilisons l’algorithme de classification K-moyennes avec un nombre de classes  $K_3 = 8$  à l’échelle la plus grossière  $j = J = 3$  et un nombre de sous-classes  $\mathbf{k}_2[k] = \mathbf{k}_1[k] = 2$  aux échelles  $j = 2$  et  $j = j_u = 1$ . Par conséquent, les vecteurs des signatures à l’étage  $j = j_u = 1$  sont classés finalement en  $K_1 = 32$  sous-classes. Le nombre de canaux considérés est  $B = 3$  et les distributions marginales sont modélisées par des DG. La figure 5.10 permet de comparer les performances fournies par les recherche exhaustive et la recherche arborescente obtenues pour les différents cas suivants.

- **a.** La recherche exhaustive est utilisée et les dépendances inter-spectrales sont ignorées.
- **b.** La recherche arborescente est utilisée et les dépendances inter-spectrales sont ignorées.
- **c.** La recherche exhaustive est utilisée et les dépendances inter-spectrales sont modélisées par une copule gaussienne.

- **d.** La recherche arborescente est utilisée et les dépendances inter-spectrales sont modélisées par une copule gaussienne.

Dans les deux cas de l’approche univariée et multivariée, on constate, comme prévu pour toute méthode de recherche non-exhaustive, une légère baisse de performance pour la recherche arborescente. Cette perte en performances est compensée, comme le montre le tableau 5.5.1 par le gain considérable en complexité en terme de nombre de DSKL calculées pour différentes valeurs du nombre d’images retournées  $R$ .

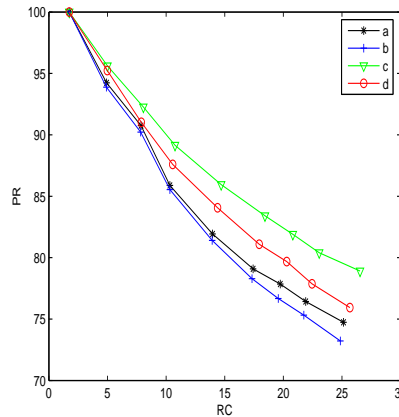


FIG. 5.10 – Courbes précision-rappel en pourcentage obtenues avec  $J = 3$ ,  $j_u = 1$  et la distance DSKL dans les deux cas de la recherche exhaustive ou arborescente.

TAB. 5.3 – Rapport moyen entre le nombre des DSKL calculées lors de la recherche arborescente et celui des DSKL calculées lors de la recherche exhaustive pour différentes valeurs du nombre  $R$  d’images retournées.

$R$	7	10	13	15	17	20
$B = 3$	0.186	0.183	0.180	0.178	0.176	0.173

### Influence de l’algorithme de classification

L’algorithme de classification K-moyennes présente l’inconvénient de classifier les signatures en un nombre prédéfini et fixe de classes. Nous avons alors suggéré comme décrit dans 4.2, d’appliquer un algorithme de classification non supervisé à savoir l’algorithme d’agglomération compétitive (AAC) avec comme nombre de classes initiales  $K_{max,3} = 8$  à l’échelle la plus grossière  $J = 3$  et un nombre de sous-classes initiales  $\mathbf{k}_{max,2}[k] = \mathbf{k}_{max,1}[k] = 2$  aux échelles  $j = 2$  et  $j = j_u = 1$ .

Pour différents nombres d'étages considérés par l'utilisateur  $j_u = 1, 2, 3$ , le tableau 5.4 donne le nombre final de sous-classes obtenues par l'algorithme de classification à l'échelle la plus fine  $j_u$ . Dans le cas de l'algorithme de classification K-moyennes, celui-ci reste, évidemment, le même que le nombre de initial de sous-classes prédéfini au début par l'utilisateur. Contrairement, dans le cas du AAC, le nombre final de sous-classes est réduit par la procédure d'agglomération. En plus, comme le montre le tableau 5.4, l'algorithme AAC abouti à de meilleurs performances en terme de précision et rappel que l'algorithme K-moyennes.

TAB. 5.4 – Nombre final de sous-classes, précision moyenne et rappel moyen obtenus avec la recherche arborescente pour les 2 algorithmes de classification K-moyennes et ACC pour  $R = 20$  et  $j_u = 1, 2, 3$ .

	$j_u = 3$	$j_u = 2$	$j_u = 1$
$K_j$ (K-moyennes)	8	16	32
$K_{j,\max}$ (AAC)	8	16	32
$K_j$ (AAC)	8	11	13
PR (K-moyennes)	79.67	84.31	83.59
PR (AAC)	79.77	86.46	87.82
RC (K-moyennes)	14.82	19.39	18.93
RC (AAC)	14.90	19.79	21.06

### 5.5.2 Influence du nombre d'échelles considérées

On s'est également proposé de comparer les performances de la recherche pour trois valeurs différentes de  $j_u = 3, 2, 1$  qui correspondent à un nombre d'échelles  $J - j_u + 1 = 1, 2, 3$ . D'après le tableau 5.4, on remarque que les performances de la recherche augmentent avec le nombre d'échelles considérées. Ainsi, si l'utilisateur n'est pas satisfait de la qualité des images en sortie à un étage  $j_u$  donné, il peut raffiner sa recherche en passant à l'échelle la plus fine et ainsi de suite.

## 5.6 Conclusion

Dans ce chapitre, nous avons validé nos contributions par des résultats expérimentaux conduits sur des images SPOT. Tout d'abord, nous avons mis en évidence le gain en performances obtenu par le modèle multivarié reposant sur la théorie des copules. Dans ce cadre, nous avons prouvé, tout d'abord, que la prise en compte des dépendances inter-spectrales par notre approche, assure de meilleurs résultats que l'approche univariée ainsi que d'autres approches multivariées telles que l'approche reposant sur une DGGM ou sur un prétraitement

préalable du type analyse en composantes indépendantes. Nous avons également prouvé que tenir compte simultanément des dépendances inter-spectrales et inter-échelles assure un gain en performances par rapport aux cas où toutes les deux ou l'une d'elles est ignorée. Ce gain est d'autant plus important que le nombre d'échelles considérées est grand. Ce gain dépend également du choix adéquat des familles de copules adoptées pour modéliser les dépendances inter-canaux et inter-échelle, surtout par la méthode bayésienne.

Nous avons ensuite étudié l'apport de l'utilisation de décompositions multirésolutions ayant des propriétés plus intéressantes que la S-TOD, en particulier le schéma de lifting vectoriel et la transformée en ondelettes  $M$ -bandes en arbre dual. Nous avons par la suite prouvé que les performances de la recherche sont influencées par le choix de la mesure de similarité. En effet la DSKL assure de meilleurs résultats que la DEN. Enfin, nous avons validé les avantages de notre algorithme de recherche arborescente à savoir le gain en complexité et l'échelonnabilité de la recherche. Nous avons également prouvé que les performances de cet algorithme dépendent du choix de l'algorithme de classification utilisé lors de la structuration de la base des index.

# Conclusion générale et perspectives

Cette thèse a été consacrée à la recherche d'images multispectrales dans le domaine transformé par des décompositions multirésolution. Nos contributions se sont articulées autour de 3 axes principaux suivants.

- L'extraction de descripteurs statistiques a pour rôle de décrire le plus "fidèlement" possible les propriétés statistiques de ces coefficients. Pour ceci, notre contribution a porté sur les deux points suivants : d'une part, le choix de la décomposition multirésolution et d'autre part, le choix d'un modèle multivarié pour modéliser la distribution conjointe des coefficients des composantes spectrales. Les paramètres de ce modèle ont été choisis comme signatures de l'image.
- Le choix d'une mesure de similarité adéquate pour comparer les vecteurs des signatures lors de la procédure de recherche.
- La conception d'un algorithme de recherche qui permet d'assurer un compromis entre les performances de recherche acceptables et une complexité opératoire raisonnable.

Pour ce faire, le premier chapitre a permis de dresser un panorama sur le fonctionnement des systèmes de recherche d'images par le contenu en mettant en évidence leur intérêt, leur principe. Nous avons également passé en revue les différents types de descripteurs rapportés dans la littérature ainsi que les moteurs de recherche d'images par le contenu les plus connus.

Dans le deuxième chapitre, nous avons procédé à une présentation des transformations en ondelettes et plus généralement des décompositions multirésolutions. Nous avons focalisé sur leurs propriétés intéressantes pour la compression et l'indexation d'images.

Le troisième chapitre a été consacré à la première partie de nos contributions se rapportant à l'étape d'extraction des signatures. Pour la première contribution, nous avons choisi comme décomposition la transformation en ondelettes dyadique séparable (S-TOD) appliquée indépendamment à chaque composante de l'image multispectrale. Par la suite, nous avons établi un modèle multivarié basé sur *la théorie des copules* pour modéliser la distribution conjointe des coefficients de toutes les composantes spectrales. Ce modèle a l'habileté de traduire les distributions marginales des coefficients d'ondelettes pour chaque canal spectral et à chaque



étage de décomposition ainsi que les dépendances inter-canaux et inter-échelles. Une seconde contribution a consisté à appliquer des décompositions multirésolutions ayant des propriétés plus intéressantes que la S-TOD. En particulier, le schéma de lifting vectoriel a permis d'exploiter les dépendances spectrales dès l'étape de décomposition. La 2D TOM-AD a été choisi pour extraire des signatures invariantes par rotation à partir de coefficients multirésolution riches en information directionnelle tout en gardant un facteur de redondance acceptable.

Dans le quatrième chapitre, nous avons présenté nos contributions concernant l'étape de la procédure de recherche. Nous avons commencé par définir les différentes mesures de similarité utilisées pour comparer les vecteurs des signatures, à savoir la Distance euclidienne normalisée et la distance symétrique de Kullback-Leibler (DSKL). Puis, nous avons présenté un nouvel algorithme de recherche ayant la particularité de permettre une organisation arborescente de la base des signatures en utilisant un algorithme de classification ainsi que d'assurer une recherche rapide et efficace selon une stratégie de parcours des échelles les plus grossières vers les échelles les plus fines.

Le cinquième chapitre a été consacré à la partie expérimentale. En utilisant des bases d'apprentissage et de test contenant des images satellitaires multispectrales SPOT3 et SPOT4 de plusieurs régions de la Tunisie, nous avons prouvé les résultats suivants.

- A décomposition multirésolution fixée, notre méthode multivariée reposant sur la théorie des copules assure de meilleures performances que d'autres méthodes multivariées présentes dans la littérature (loi DGGM ou prétraitement par analyse en composantes indépendantes des sous-bandes). Ces performances sont améliorées que la copule est choisie dans un cadre bayésien.
- La prise en compte simultanée par des copules des dépendances entre composantes inter-canaux et inter-échelles permet d'améliorer les performances par rapport aux cas où l'une d'elles est ignorée.
- Le choix de la décomposition influence grandement les performances que ce soit pour diminuer les redondances spectrales grâce au schéma de lifting vectoriel ou en permettant une recherche invariante par rotation grâce à la 2D TOM-AD.
- Le choix de la mesure de similarité pour comparer les signatures a une influence avérée sur les performances de la recherche. Ainsi, la DSKL assure-t-elle de meilleurs résultats que la DEN, étant adaptée à comparer des modèles statistiques.
- Le choix de l'algorithme de recherche est très important pour satisfaire un compromis entre performances de la recherche et complexité opératoire. Dans notre cas, la propriété d'échelonnabilité des coefficients d'ondelettes nous a été très utile pour concevoir un algorithme de recherche arborescente fiable et rapide.

Plusieurs perspectives sont envisageables à l'issue de notre travail. Nous pouvons d'abord étudier l'invariance par rapport à des transformations de l'image autres que la rotation. De plus, il serait intéressant d'explorer l'influence de la quantification des coefficients sur les performances du système d'indexation. Par ailleurs, jusqu'alors nous avons étudié l'influence des descripteurs issus de ces modélisations statistiques sur les performances du système d'accès. Ces dernières pourraient être améliorées si plusieurs types de descripteurs sont mis en jeu. Il s'avère alors prometteur d'étudier dans quelle mesure le couplage des descripteurs proposés dans cette thèse avec des signatures plus conventionnelles pourrait aboutir à un système de recherche plus performant. Enfin, une extension intéressante au cas d'image hyperspectrales pourrait être envisagée selon une stratégie similaire en prenant en compte le nombre élevé de composantes spectrales.

## Annexe A

# Estimation bayésienne de la famille de copule

Nous présentons ici une généralisation pour  $B \geq 2$  de la méthode d'estimation bayésienne initialement proposée dans [140] dans le cas bivarié  $B = 2$ .

### A.1 Position du problème

Soit  $(X_1, \dots, X_B)$ ,  $B$  variables aléatoires et  $\{C_{l;\theta_l}\}_{l=1}^L$  un ensemble de  $L$  familles de copules caractérisée chacune par un paramètre  $\theta_l$  de dimension  $p_l$ . Notre but consiste à trouver la “meilleure” famille de copules  $\{C_{l^*;\theta_l^*}\}$  qui caractérise le “mieux” la dépendance entre ces variables.

### A.2 Stratégie d'estimation bayésienne

Par analogie avec la méthode proposée dans [140] dans le cas bivarié, on a besoin d'un vecteur  $\mathbf{U} = (\mathbf{U}_1^T, \dots, \mathbf{U}_I^T)^T$  de  $I$  événements normalisés indépendants  $\mathbf{U}_1, \dots, \mathbf{U}_I \in [0, 1]^B$  calculés à partir de réalisations du vecteur de variables aléatoires  $(X_1, \dots, X_B)$ . En effet, la sélection va s'appuyer plutôt sur ces événements  $\mathbf{U}$  que sur les observations du vecteur  $(X_1, \dots, X_B)$ .

Il est clair que la fiabilité de cette méthode est fortement liée au choix de événements  $\mathbf{U}_1, \dots, \mathbf{U}_I$ . Un des choix les plus communs consiste à considérer des quantiles des  $B$  composantes. Ce choix a pour avantage d'assurer une robustesse au bruit pouvant entacher les observations disponibles.

Le problème de la sélection de la meilleure copule  $\{C_{l^*;\theta_l^*}\}$  peut être formalisé comme étant un problème de décision entre l'ensemble d'hypothèses  $\{H_l\}_{l=1}^L$  :

$H_l$  : Le vecteur de réalisations  $\mathbf{U}$  émane de la copule  $C_{l;\theta_l}$ .

L'estimation de la meilleure famille de copules se fait grâce à une stratégie de maximum a posteriori qui consiste à sélectionner la copule  $\{C_{l^*;\theta_l^*}\}$  qui maximise la probabilité conditionnelle  $P(H_l|\mathbf{U})$  de  $H_l$  connaissant les réalisations  $\mathbf{U}$  :

$$l^* = \arg_{l=1,\dots,L} \max P(H_l|\mathbf{U}). \quad (\text{A.1})$$

L'enjeu est donc d'obtenir une expression des probabilités conditionnelles  $P(H_l|\mathbf{U})$  ou encore de la probabilité a priori  $P(H_l)$  et de la vraisemblance  $P(\mathbf{U}|H_l)$  et ce, en vertu du théorème de Bayes :

$$P(H_l|\mathbf{U}) = P(\mathbf{U}|H_l)P(H_l). \quad (\text{A.2})$$

A cet égard, il s'avère judicieux de recourir à une représentation des copules  $\{C_{l;\theta_l}\}$  qui facilite l'expression de  $P(\mathbf{U}|H_l)P(H_l)$ . Un des paramètres les plus utilisées est le tau de Kendall est une mesure de dépendance entre variables aléatoires. Dans la cas bivarié, le tau de Kendall  $\tau_l$  n'est autre que la différence entre la probabilité de concordance  $P_c$  et celle de discordance  $P_d$  définies par :

$$\begin{aligned} P_c &= Pr\{(X_1 - X'_1)(X_2 - X'_2) > 0\} \\ P_d &= Pr\{(X_1 - X'_1)(X_2 - X'_2) < 0\} \end{aligned} \quad (\text{A.3})$$

où  $(X'_1, X'_2)$  est un couple de variables aléatoires identique à  $(X_1, X_2)$ . Le tau de Kendall  $\tau_l$  est défini comme suit :

$$\tau_l = P_c - P_d. \quad (\text{A.4})$$

Le tau de Kendall possède aussi l'avantage de s'exprimer simplement en fonction de la copule associée au couple de variables aléatoires  $(X_1, X_2)$  de la manière suivante :

$$\tau_l = 4 \left( \int \int_{[0,1]^2} 4C_l(x_1, x_2) dC(x_1, x_2) - 1 \right). \quad (\text{A.5})$$

Dans le cas multivarié, l'expression du tau de Kendall associé à une copule  $C_l$  s'obtient de la manière suivante :

$$\tau_l = \frac{1}{2^{B-1} - 1} \left( \int_{[0,1]^B} 2^B C_l(\mathbf{x}) dC_l(\mathbf{x}) - 1 \right). \quad (\text{A.6})$$

Cette équation indique qu'il existe une relation explicite entre  $\tau_l$  et  $\theta_l$ . Ainsi si le paramètre de la copule est scalaire ( $p_l = 1$ ), une reparamétrisation évidente de la copule se fait par le tau de Kendall

$$\tau_l = g(\theta_l). \quad (\text{A.7})$$

Dans le cas général, on peut recourir à d'autres mesures statistiques comme la dépendance de queue pour définir un nouveau paramètre  $\tau_l$  en fonction de  $\theta_l$  :

$$\boldsymbol{\tau}_l = \begin{pmatrix} \tau_{1,l} \\ \vdots \\ \tau_{p_l,l} \end{pmatrix} = \mathbf{g}_l(\boldsymbol{\theta}_l). \quad (\text{A.8})$$

L'avantage essentiel de cette reparamétrisation apparaît clairement puisqu'elle permet de normaliser en quelque sorte les copules. En effet, ces dernières seront exprimées en fonction de paramètres  $\tau_l$  de même nature.

Une fois cette représentation effectuée, l'enjeu est de d'introduire  $\tau_l$  comme une variable aléatoire auxiliaire dans un cadre bayésien (appelée aussi variable de nuisance) :

$$P(H_l|\mathbf{U}) = \int_{\Omega_l} P(H_l, \tau_l|\mathbf{U})d\tau_l \quad (\text{A.9})$$

où  $\Omega_l$  est le domaine de définition de  $\tau_l$ .

Grâce au théorème de Bayes et l'équation (A.9), le problème se réduit à trouver la copule  $\{C_{l^*}; \boldsymbol{\theta}_l^*\}$  qui maximise la quantité suivante :

$$P(H_l|\mathbf{U}) \propto \int_{\Omega_l} P(\mathbf{U}|H_l, \tau_l)P(H_l|\tau_l)P(\tau_l)d\tau_l. \quad (\text{A.10})$$

Cette optimisation passe par le choix des expressions des 3 probabilités apparaissant dans le membre de droite de l'équation (A.10).

– Vraisemblance :

Comme les  $I$  réalisations du vecteur  $\mathbf{U}$  sont considérées comme étant indépendantes, la vraisemblance peut s'écrire de la manière suivante :

$$P(\mathbf{U}|H_l, \tau_l) = \prod_{i=1}^I c_{l; \mathbf{g}_l^{-1}(\tau_l)}(\mathbf{U}_i). \quad (\text{A.11})$$

– Loi a priori sur la famille de copules (ou  $H_l$ ) :

On suppose qu'on ne dispose pas de connaissance antérieure supplémentaire sur les familles de copules ce qui se traduit par :

$$P(H_l|\tau_l) \propto 1_{\Omega_l}. \quad (\text{A.12})$$

– Loi a priori sur le paramètre  $\tau_l$  :

La connaissance a priori sur le paramètre  $\tau_l$  peut s'écrire de la manière suivante :

$$P(\tau_l) = \begin{cases} \frac{1}{\lambda(\Lambda_l)} & \text{si } \tau_l \in \Lambda_l \\ 0 & \text{sinon} \end{cases} \quad (\text{A.13})$$

où  $\lambda$  est la mesure de Lebesgue de l'intervalle  $\Lambda_l \subset \Omega_l$  de toutes les valeurs possibles de  $\tau_l$  qui sont considérées comme étant équiprobables. Le sous-intervalle  $\Lambda_l$  est introduit dans le but d'ajouter une éventuelle information préalable supplémentaire sur le paramètre  $\tau_l$ . Dans le cas où contraire,  $\Lambda_l$  se réduit à l'intervalle  $\Omega_l$  qui est le domaine de définition de  $\tau_l$ .

### A.3 Sélection de la meilleure copule

Finalement, on peut conclure que l'index  $l^*$  de la famille de copule qui traduit le mieux la dépendance entre les variables aléatoires  $X_1, \dots, X_B$  peut s'exprimer de la manière suivante :

$$l^* = \arg \max_{l=\{1, \dots, L\}} \frac{1}{\lambda(\Lambda_l)} \int_{\Omega_l \cap \Lambda_l} \prod_{i=1}^I c_{l; \mathbf{g}_l^{-1}(\boldsymbol{\tau}_l)}(\mathbf{U}_i) d\boldsymbol{\tau}_l. \quad (\text{A.14})$$

Une fois la famille de copule  $\{C_{l^*; \boldsymbol{\theta}_{l^*}}\}$  sélectionnée, le paramètre  $\boldsymbol{\theta}_{l^*}$  peut être estimé à partir de réalisations des variables  $X_1, \dots, X_B$  par une approche par maximum de vraisemblance ou une méthode des moments.

## Annexe B

# Généralisation du théorème de Sklar

La formulation initiale du théorème de Sklar concerne seulement le cas où les variables marginales sont *univariées*. Kolev *et al.* [143] ont proposé une extension de ce théorème au cas où les variables marginales sont *multivariées* et ceci de la manière suivante. On considère  $X_1, \dots, X_D$  variables aléatoires réparties en  $K$  vecteurs aléatoires disjoints  $\mathbf{X}_k = (X_{l_{k-1}+1}, \dots, X_{l_k})^\top$  de dimension  $d_k$ , avec  $2 \leq K \leq D$ ,  $l_0 = 0$ ,  $l_k = \sum_{j=1}^k d_j$  et  $\sum_{k=1}^K d_k = D$ . Kolev [143] a montré l'existence d'une copule  $K$ -dimensionnelle  $\tilde{C}$  et  $K$  copules  $\bar{C}_1, \dots, \bar{C}_K$  de dimensions respectives  $d_1, \dots, d_k$  telles que :

$$\forall (x_1, \dots, x_D) \in \mathbb{R}^D \quad F(x_1, \dots, x_D) = \tilde{C}(\bar{F}_1(\mathbf{X}_1), \dots, \bar{F}_K(\mathbf{X}_K)) \quad (\text{B.1})$$

où  $\bar{F}_1, \dots, \bar{F}_K$  sont les fonctions de répartition conjointes des vecteurs  $\mathbf{X}_1, \dots, \mathbf{X}_K$ . A leur tour, celles-ci s'expriment à l'aide du théorème de Sklar de la manière suivante :

$$\begin{aligned} \bar{F}_k(\mathbf{X}_k) &= \bar{F}_k(X_{l_{k-1}+1}, \dots, X_{l_k}) \\ &= \bar{C}_k(F_{l_{k-1}+1}(x_{l_{k-1}+1}), \dots, F_{l_k}(x_{l_k})) \end{aligned} \quad (\text{B.2})$$

où  $F_{l_{k-1}+1}, \dots, F_{l_k}$  sont les fonctions de répartition marginales des variables univariées  $X_{l_{k-1}+1}, \dots, X_{l_k}$  formant le  $k$ -ième vecteur aléatoires multivarié  $\mathbf{X}_k$ .

Ainsi, les dépendances entre les  $d_k$  variables univariées  $X_{l_{k-1}+1}, \dots, X_{l_k}$  à l'intérieur de la  $k$ -ième classe sont modélisées par la copule  $\bar{C}_k$  tandis que le lien entre les différentes classes ou vecteurs aléatoires  $\mathbf{X}_1, \dots, \mathbf{X}_K$  est traduit par la copule  $\tilde{C}$ .

## Annexe C

# Schéma de lifting vectoriel

A partir d'un signal multicomposantes d'approximation  $(a_j^{(b)})_{b=1}^B$  à un étage  $j = 0, \dots, J-1$ , le schéma de lifting vectoriel [118] permet d'obtenir un signal d'approximation  $(a_{j+1}^{(b)})_{b=1}^B$  et un signal détail  $(d_{j+1}^{(b)})_{b=1}^B$  et ceci de la manière suivante. La composante  $b_1$  est décomposée en mode spatial c'est-à-dire selon le schéma de lifting conventionnel tandis que les autres composantes  $b_i$  ( $i = 2, \dots, B$ ) sont décomposées selon les modes spatial et inter-spectral. En effet, à un étage  $j = 0, \dots, J-1$ , les coefficients d'indices pairs du signal  $a_j^{(b_i)}$  sont prédits à partir des coefficients d'indices impairs du même signal  $a_j^{(b_i)}$  (mode intra) mais aussi à partir de coefficients des composantes précédentes  $a_j^{(b_k)}$  avec  $k < i$ . Ensuite, le signal détail  $d_{j+1}^{(b_i)}$  correspondant au canal  $b_i$  à l'étage  $j+1$  est calculé comme étant l'erreur de prédiction comme suit pour tout  $\forall n = 1, \dots, N/2^{j+1}$ ,

$$d_{j+1}^{(b_i)}(n) = a_j^{(b_i)}(2n) - \left[ \sum_{\ell \in \mathcal{P}_{j,b_i}^{(b_i)}} p_{j,b_i}^{(b_i)}(\ell) a_j^{(b_i)}(2n - \ell) \right] - \left[ \sum_{k=1}^{i-1} \sum_{\ell \in \mathcal{P}_{j,b_i}^{(b_k)}} p_{j,b_i}^{(b_k)}(\ell) a_j^{(b_k)}(2n - \ell) \right] \quad (\text{C.1})$$

où  $\mathcal{P}_{j,b_i}^{(b_k)}$  est l'opérateur de prédiction du canal  $b_i$  à partir du canal  $b_k$  de poids  $p_{j,b_i}^{(b_k)}$ .

Selon la même logique que celle du schéma de lifting conventionnel, le signal approximation  $a_{j+1}^{(b_i)}$  est obtenu après modification des coefficients d'indices impairs du signal  $a_j^{(b_i)}$  par certains coefficients de détail des signaux  $d_{j+1}^{(b_k)}$  comme suit :

$$a_{j+1}^{(b_i)}(n) = a_j^{(b_i)}(2n-1) - \left[ \sum_{k=1}^B \sum_{\ell \in \mathcal{U}_{j,b_i}^{(b_k)}} u_{j,b_i}^{(b_k)}(\ell) d_{j+1}^{(b_k)}(n - \ell) + \sum_{k=1}^{i-1} \sum_{\ell \in \mathcal{U}'_{j,b_i}^{(b_k)}} u'_{j,b_i}{}^{(b_k)}(\ell) a_j^{(b_k)}(2n - 1 - \ell) \right] \quad (\text{C.2})$$

où  $\mathcal{U}_{j,b_i}^{(b_k)}$  et  $\mathcal{U}'_{j,b_i}{}^{(b_k)}$  sont les opérateurs de mise à jour du canal  $b_i$  à partir du canal  $b_k$  de poids respectifs  $u_{j,b_i}^{(b_k)}$  et  $u'_{j,b_i}{}^{(b_k)}$ .

Dans le cas d'une image multispectrale  $(I_0^{(b)})_{b=1}^B$ , le schéma de lifting vectoriel est appliqué sur l'intégralité des  $B$  composantes de l'image en opérant d'une manière séparable (sur les lignes puis sur les colonnes), ce qui fournit une image d'approximation multispectrale  $(I_J^{(b)})_{b=1}^B$  de



taille  $N_J \times N_J$  avec  $N_J = \frac{N}{2^J}$  à l'étage le plus grossier  $J$  et 3 sous-bandes détails multispectrales  $(w_{j,o}^{(b)})_{b=1}^B$  de taille  $N_j \times N_j \times B$  avec  $N_j = N/2^j$  orientées horizontalement ( $o = 1$ ), verticalement ( $o = 2$ ) et diagonalement ( $o = 3$ ) à chaque étage  $j = 1, \dots, J$ .

La décomposition DEC3 est un schéma de lifting vectoriel, considérée comme la généralisation de la décomposition 5/3. Elle est définie par les opérateurs de prédiction et de mise à jour suivants :

$$\mathcal{P}_{j,b_1}^{(b_1)} = \{-1, 1\}, \quad p_{j,b_1}^{(b_1)}(\pm 1) = \frac{1}{2}, \quad (\text{C.3})$$

$$\mathcal{P}_{j,b_2}^{(b_2)} = \mathcal{P}_{j,b_3}^{(b_3)} = \{-1, 1, 2\}, \quad \mathcal{P}_{j,b_2}^{(b_1)} = \mathcal{P}_{j,b_3}^{(b_2)} = \{-1, 0, 1\}, \quad (\text{C.4})$$

$$\mathcal{U}_{j,b_1}^{(b_1)} = \{0, 1\}, \quad \mathcal{U}'_{j,b_1}^{(b_1)} = \{1\}, \quad \mathcal{U}_{j,b_3}^{(b_3)} = \mathcal{U}_{j,b_2}^{(b_2)} = \{0, 1\}. \quad (\text{C.5})$$

## Annexe D

# Approximation du centroïde d'une classe au sens de la DSKL

Dans le cas de la DSKL, une approximation du centroïde d'une classe a été conçue [164] de la manière suivante.

Pour calculer le centroïde d'une classe donnée, les densités de probabilité  $f_{w_j^{i_{db}}}$  concernées sont représentées chacune par une séquence aléatoire  $\mathbf{s}_1^{i_{db}}, \dots, \mathbf{s}_L^{i_{db}}$  de longueur  $L$ . Le centroïde  $\mathbf{p}_j$  de la classe est alors représenté par la séquence aléatoire  $\mathbf{p}_1, \dots, \mathbf{p}_L$  de longueur  $L$  qui minimise la fonction objective suivante :

$$g(\mathbf{p}, \lambda) = d(\bar{\mathbf{s}}, \mathbf{p}) + d(\mathbf{p}, \tilde{\mathbf{s}}) + \lambda \left( \sum_{\ell=1}^L \mathbf{p}_\ell - 1 \right) \quad (\text{D.1})$$

où  $d$  est la divergence de Kullback-Leibler,  $\bar{\mathbf{s}}$  et  $\tilde{\mathbf{s}}$  représentent respectivement la moyenne arithmétique et la moyenne géométrique calculées sur toutes les séquences aléatoires  $\mathbf{s}^{i_{db}}$  de tous les éléments de la classe et  $\lambda$  est le multiplicateur de Lagrange.

# Bibliographie

- [1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions of Pattern Recognition and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, 2000.
- [2] N. Hervé, "Vers une description efficace du contenu visuel pour l'annotation automatique d'images," Thèse en sciences spécialité informatique de l'Université d'Orsay, Paris-Sud 11, France, 2009.
- [3] A. M. Puetz, K. Lee, and R. C. Olsen, "WorldView-2 data simulation and analysis results," *Proceeding of SPIE : Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, Orlando, FL, USA, 2009.
- [4] J. R. Jensen, "Remote sensing of the environment : an earth resource perspective," *Prentice-Hall series in geographic information science*, Upper Saddle River, NJ : Pearson Prentice Hall, 2007.
- [5] A. Bouillon, E. Breton, F. de Lussy, and R. Gachet, "SPOT5 geometric image quality," *IEEE International Geoscience and Remote Sensing Symposium*, Toulouse, France, 2003.
- [6] S. Baillarin, " Remote sensing image ground segment interoperability : Pleiades-HR case study," *IEEE International Geoscience and Remote Sensing Symposium*, Denver, Colorado, USA, 2006.
- [7] C. B. Jenkerson and B. C. Reed, "MODIS land data at the EROS data center DAAC," *International Geoscience and Remote Sensing Symposium*, Sydney, Australia, 2001.
- [8] T. Barclay , J. Gray, and D. Slutz, "Microsoft TerraServer : a spatial data warehouse," *ACM SIGMOD International Conference on Management of Data*, Dallas, Texas, USA, 2000.
- [9] K. Barnard, P. Duygulu, D. Forsyth, N. Freitas, D. M. Blei, and M. I. Jordan, "Matching words and pictures," *Journal of Machine Learning Research*, vol. 3, pp. 1107-1135, 2003.
- [10] F. Jing, M. Li, H. J. Zhang, and B. Zhang, "Keyword propagation for image retrieval," *International Symposium on Circuits and Systems*, Vancouver, Canada, 2004.

- [11] A. Haubold and A. Natsev, "Web-based information content and its application to concept-based video retrieval," *International Conference on Content-based Image and Video Retrieval*, Niagara Falls, Canada, 2008.
- [12] D. J. Harper J. M. Jose, and J. Furner, "Spatial querying for image retrieval : a useroriented evaluation," *International ACM SIGIR Conference on Research and Development in Information Retrieval*, Melbourne, Australia, 1998.
- [13] S. P. Wilson, J. Fauqueur, and N. Boujemaa, "Machine learning techniques for multimedia," *Implicit Versus Explicit Content Query, chapter Mental Search in Image Databases*, pp. 189-204, 2008.
- [14] A. DelBimbo and P. Pala, "Visual image retrieval by elastic matching of user sketches," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 121-132, 1997.
- [15] J. Fauqueur and N. Boujemaa, "Mental image search by boolean composition of region categories," *Multimedia Tools and Applications archive*, vol. 31, no. 31, pp. 95-117, 2006.
- [16] J. Fauqueur and N. Boujemaa, "Region-based image retrieval : Fast coarse segmentation and fine color description," *Journal of Visual Languages and Computing (JVLC), special issue on Visual Information Systems*, vol. 15, no. 1, pp. 69-95, 2004.
- [17] N. Vasconcelos and A. Lippman, "A unifying view of image similarity," *International Conference on Pattern Recognition*, Barcelona, Spain, 2000.
- [18] I. J. Cox, M. L. Miller, and T. P. Minka, "The bayesian image retrieval system, pichunter : Theory, implementation and psychological experiments," *IEEE Transactions on Image Processing*, vol. 9, no.1, pp. 20-37, 2000.
- [19] R. Kimchi and S. E. Palmer, "Form and texture in hierarchically constructed patterns," *Journal of Experimental Psychology : Human Perception and Performance* vol. 8, no. 4, pp. 521-535, 1982.
- [20] M. Hagedoorn and R. C. Veltkamp, "Reliable and efficient pattern matching using an affine invariant metric," *International Journal of Computer Vision*, vol. 31, no. 2, pp. 203-225, 1999.
- [21] D. Zhang and G. Lu, "A comparative study on shape retrieval using Fourier descriptors with different shape signatures," *Journal of Visual Communication and Image Representation*, vol. 14, pp. 41-60, 2003.
- [22] O. Martinez Bruno, L. Gustavo Nonato, M. A. Pazoti, and J. Batista, "Topological multi-contour decomposition for image analysis and image retrieval," *Pattern recognition letters*, vol. 29, no. 11, pp. 1675-1683, 2008.
- [23] C. Schmid and R. Mohr, "Local gray-value invariants for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530-535, 1997.

- [24] S. Derrode, M. Daoudi, and F. Ghorbel, "Invariant content-based image retrieval using a complete set of Fourier-Mellin descriptors," *IEEE International Conference on Multimedia Computing and Systems*, Florence, Italy, 1999.
- [25] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic, "Automatic and semi-automatic methods for image annotation and retrieval in query by image content (QBIC)," *Proc of SPIE Internatinal Conference on Storage and Retrieval for Image and Video Databases III*, San Jose, CA, USA, 1995.
- [26] C. L. Huang and D. H. Huang, "A Content-based image retrieval system," *Image and Vision Computing*, vol. 16, no. 3, pp. 149-163, 1998.
- [27] I. Bartolini, P. Ciaccia, and M. Patella, "WARP : accurate retrieval of shapes using phase of Fourier descriptors and time warping distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 141-147, 2005.
- [28] M. A. Charmi, S. Derrode, and F. Ghorbel, "Fourier-based geometric shape prior for snakes," *Pattern Recognition Letters*, vol 29, no. 7, pp. 897-904, 2008.
- [29] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *Computer Vision, Graphics and Image processing*, vol. 54, no. 5, pp. 438-460, 1992.
- [30] C. Ruberto and A. Morgera, "Moment-based techniques for image retrieval," *International Conference on Database and Expert Systems Application*, Torino, Italy, 2008.
- [31] M. K. Kamila, S. Mahapatra, and S. Nada, "Invariance image analysis using modified Zernike moments," *Pattern Recognition Letters*, vol. 26, no. 6, pp. 747-753, 2005.
- [32] F. Ghorbel, S. Derrode, R. Mezhoud, M. T. Bannour, and S. Dhahbi, "Image reconstruction from a complete set of similarity invariants extracted from complex moments," *Pattern Recognition Letters* vol. 27, no. 12, pp. 1361-1369, 2006.
- [33] F. Ghorbel, "A complete invariant description for gray-level images by the harmonic analysis approach," *Pattern Recognition Letters*, vol. 15, no. 10, pp. 1043-1051, 1994.
- [34] S. Derrode and F. Ghorbel, "Robust and efficient Fourier-Mellin transform approximations for gray level image reconstruction and complete invariant description," *Computer Vision and Image Understanding*, vol. 83, no. 1, pp. 57-78, 2001.
- [35] S. Derrode and F. Ghorbel, "Shape analysis and symmetry detection in gray-level objects using the analytical Fourier-Mellin representation," *Signal Processing*, vol. 84, no. 1, pp. 25-39, 2004.
- [36] S. W. Zucker, D. Terzopoulos, "Finding structure in co-occurrence matrices for texture analysis," *Computer Graphics and Image Processing*, vol. 12, no. 3, pp. 286-308, 1980.

- [37] M. Partio, B. Cramariuc, and M. Gabbouj, "An Ordinal co-occurrence matrix framework for texture retrieval," *EURASIP Journal on Image and Video Processing*, vol. 16, no. 1, pp. 24-37, 2007.
- [38] I. M. Elfadel and R. W. Picard, "Gibbs random fields, cooccurrences, and texture modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, pp. 24-37, 1994.
- [39] D. C. He and L. Wang, "Texture unit, texture spectrum and texture analysis," *IEEE Transactions on Geoscience and Remote sensing*, vol. 28, no. 4, pp. 509-512, 1990.
- [40] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786-804, 1979.
- [41] D. G. Sim, H. K. Kim, and D. I. Oh, "Translation, scale, and rotation invariant texture descriptor for texture-based image retrieval," *International Conference on Image Processing*, Vancouver, Canada, 2000.
- [42] J. Beck, A. Sutter, and R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *Computer vision, graphics, and image processing*, vol. 37, no. 2, pp. 299-325, 1987.
- [43] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol*, vol. 160, no. 1, pp. 106-154, 1962.
- [44] J. R. Smith and S. F. Chang, "Transform features for texture classification and discrimination in large image databases," *International Conference on Image Processing*, Austin, TX, USA, 1994.
- [45] N. G. Rao, V. V. Kumar, and V. V. Krishna, "Texture based image indexing and retrieval," *International Journal of Computer Science and Network Security*, vol. 9, no. 5, pp. 206-210, 2009.
- [46] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146-158, 2002.
- [47] R. J. Mathiassen, A. Skavhaug, and K. Bo, "Texture similarity measure using Kullback-Leibler divergence between Gamma distributions," *European Conference on Computer Vision*, Copenhagen, Denmark, 2002.
- [48] M. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models," *IEEE Transactions on Multimedia*, vol. 4, no. 4, pp. 517-527, 2002.
- [49] P. Pala and S. Santini "Image retrieval by shape and texture," *Journal of Pattern Recognition* vol. 32, no. 3, pp. 517-527, 1999.

- [50] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content : The QBIC System," *IEEE Computer*, vol. 28, no. 9, pp. 23-32, 1995.
- [51] J. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Gorowitz, R. Humphrey, R. Jain, and C. Shu, "Virage image search engine : an open framework for image management," *Proceedings of the SPIE, Storage and Retrieval for Image and Video Databases IV*, San Jose, CA, USA, 1996.
- [52] T. P. Minka and R. W. Picard, "Interactive learning with a society of models," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 1996.
- [53] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang, "Supporting similarity queries in MARS," *Proceedings of the ACM International Multimedia Conference*, Seattle, Washington, USA, 1997.
- [54] J. R. Smith and S. F. Chang, *Querying by color regions using the VisualSEEK content-based visual query system*, Maybury M.T. (ed.), Intelligent Multimedia Information Retrieval, California : AAAI Press, pp. 23-42, 1997.
- [55] W. Y. Ma and B. S. Manjunath, "Netra : A toolbox for navigating large image databases," *Multimedia Systems*, vol. 7, no. 3, pp. 184-198, 1999.
- [56] N. Boujemaa, J. Fauqueur, M. Ferecatu, F. Fleuret, V. Gouet, B. LeSaux, and H. Sahbi, "Ikona interactive specific and generic image retrieval," *MultiMedia Content Based Indexing and Retrieval Workshop*, Rocquencourt, France, 2001.
- [57] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLicity : Semantics-sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947-963, 2001.
- [58] Q. Iqbal and J. K. Aggarwal, "CIRES : a system for content-based retrieval in digital image libraries," *IEEE International Conference of Control, Automation, Robotics and Vision*, Singapore, 2002.
- [59] S. Kiranyaz, K. Caglar, E. Guldogan, O. Guldogan, M. Gabbouj "Muvis : A content-based multimedia indexing and retrieval framework," *International Symposium on Signal Processing and its Applications*, Paris, France, 2003.
- [60] S. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 11, pp. 674-693, 1989.
- [61] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909-996, 1988.

- [62] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform,” *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205-220, 1992.
- [63] A. Cohen, I. Daubechies, and J. C. Feauveau, “Biorthogonal bases of compactly supported wavelets,” *Communications on Pure and Applied Mathematics*, vol. 45, no. 5, pp. 485-560, 1990.
- [64] O. Rioul and P. Duhamel, “Fast algorithms for discrete and continuous wavelet transforms,” *IEEE transactions on information theory*, vol. 38, no. 2, pp. 569-586, 1992.
- [65] O. Rioul “A discrete-time multiresolution theory,” *IEEE transactions on signal processing*, vol. 41, no. 8, pp. 2591-2606, 1993.
- [66] J. Kovacevic and W. Sweldens, “Wavelet families of increasing order in arbitrary dimensions,” *IEEE transactions on image processing*, vol. 9, no. 3, pp. 480-496, 2000.
- [67] P. Steffen, P. Heller, R. A. Gopinath, C. S. Burrus, and R. A. Gopinath “Theory of regular  $M$ -Band wavelet bases, ”*IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3497–3511, 1993.
- [68] W. Sweldens, “The lifting scheme : a custom-design construction of biorthogonal wavelets,” *Applied and Computational Harmonic Analysis*, vol. 3, no. 2, pp. 186-200, 1996.
- [69] R.C. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, “Wavelet transforms that map integers to integers,” *Applied and Computational Harmonic Analysis*, vol. 5, no. 3, pp. 332-369, 1998.
- [70] E. P. Simoncelli and W. T. Freeman, “The steerable pyramid : a flexible architecture for multi-scale derivative computation,” *IEEE International Conference on Image Processing*, Washington, DC, USA, 1995.
- [71] E. P. Simoncelli and H. Farid, “Steerable wedge filters for local orientation analysis,” *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1377-1382, 1996.
- [72] B. Beferull-Lozano, H. Xie, and A. Ortega, “Rotation-invariant features based on steerable transforms with an application to distributed image classification,” *IEEE International Conference on Image Processing*, Barcelona, Spain, 2003.
- [73] G. Tzagkarakis, B. Beferull-Lozano, and P. Tsakalides, “Rotation-invariant texture retrieval with gaussianized steerable pyramids,” *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2702-2718, 2006.
- [74] G. Tzagkarakis, B. Beferull-Lozano, and P. Tsakalides, “Rotation-invariant texture retrieval via signature alignment based on steerable sub-gaussian Modeling,” *IEEE Transactions on Image Processing*, vol. 17, no. 17, pp. 1212-1225, 2008.
- [75] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, “The dual-tree complex wavelet transform,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123-151, 2005.



- [76] J. K. Romberg, H. Choi, R. G. Baraniuk, and N. G. Kingsbury, "Multiscale classification using complex wavelets and hidden markov tree models," *International Conference on Image Processing*, Vancouver, Canada, 2000.
- [77] N. G. Kingsbury, "Rotation-invariant matching of local features using dual-tree complex wavelets," *CIRM Mathematical Image Processing Meeting*, 2007.
- [78] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Rotation invariant texture image retrieval using rotated complex wavelet filters," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 36, no. 6, pp. 1273-1282, 2006.
- [79] A. Mumtaz and S. A. M. Gilani, K. Hameed, and T. Jameel, "Enhancing performance of image retrieval systems using dual tree complex wavelet transform and support vector machines," *Journal of Computing and Information Technology*, vol. 16, no. 1, pp. 57-68, 2008.
- [80] C. Chaux, L. Duval, and J. C. Pesquet, "Image analysis using a dual-tree M-band wavelet transform," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2397-2412, 2006.
- [81] G. Held, T. R. Marshall, "Data and image compression : tools and techniques," *Edition : 4th ed, Publisher : Chichester, England, New York : John Wiley*, 1996.
- [82] K. R. Rao and P. Yip, "Discrete cosine transform : algorithms, advantages, applications," *Academic Press Professional, USA*, 1990.
- [83] W. Dong, G. Shi, J. Xu, and X. U. Xian, "Adaptive nonseparable interpolation for image compression with directional wavelet transform," *IEEE Signal Processing Letters*, vol. 15, no. 1 pp. 233-236, 2008.
- [84] C. Eubanks, "Haar wavelets, image compression, and multi-resolution analysis," *Initial Report for Capstone Paper*, 2007.
- [85] T. Brahimi, A. Melit, and F. Khelifi, "An improved SPIHT algorithm for lossless image coding," *Digital Signal Processing*, vol. 19 , no. 2, pp. 220-228, 2009.
- [86] A. Bilgin, J. Sementilli, S. Fang, M. W. Marcellin, "Scalable image coding using reversible integer wavelet transforms," *IEEE Transactions on Image Processing*, vol. 9, no. 11, pp. 1972-1977, 1999.
- [87] D. S. Taubman, M. W. Marcellin, *JPEG2000 : Image compression fundamentals, standards and practice*, Kluwer, Boston, MA, 2002.
- [88] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system : an overview," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 4, pp. 1103-1127, 2000.
- [89] "JPEG. Coding of still pictures-JPEG 2000 part2" *ISO/IEC 15444-2*, 2001.
- [90] A. Vellaikal and C.-C. Kuo, "Joint spatial-spectral indexing for image retrieval," *IEEE International Conference on Image Processing*, Lausanne, Switzerland, 1996.

- [91] G. Feng and J. Jiang, "JPEG image retrieval based on features from DCT domain," *International Conference on Image and Video Retrieval*, London, UK, 2002.
- [92] F. Arnia, I. Iizuka, I. Masaaki, and F. H. Kiya, "DCT sign-based similarity measure for JPEG image retrieval," *IEEE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 90, no. 9, pp. 1976-1985, 2007.
- [93] J. L. Crowley, "A representation for visual information," *Technical Report CMU-RI-TR-82-07*, Robotics Institute, Carnegie Mellon University, 1981.
- [94] F. W. Campbell and J. J. Kulikowski, "Oriental selectivity of the human visual system," *Journal of Physiology*, vol. 187, pp. 437-445, 1966.
- [95] T. Chang and C. C. Jay, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 429-441, 1993.
- [96] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image indexing using moments and wavelets," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 3, pp. 557-565, 1996.
- [97] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux, "Wavelet optimization for content-based image retrieval in medical databases," *Medical Image Analysis*, vol. 14, no. 2, pp. 227-241, 2010.
- [98] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445-3462, 1993.
- [99] A. Said and W. A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243-250, 1996.
- [100] D. S. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158-1170, 2000.
- [101] S. T. Hsiang and J. W. Woods, "Embedded image coding using zeroblocks of sub-band/wavelet coefficients and context modeling," *IEEE International Symposium on Circuits and Systems*, Geneva, Switzerland, 2000.
- [102] C. Liu and M. Mandal, "Multiresolution image indexing based on embedded zerotree wavelet coding", *Canadian Conference on Electrical and Computer Engineering*, Halifax, Canada, pp. 430-434, 2000.
- [103] G. Voulgaris and J. Jiang, "Texture-based image retrieval in wavelets compressed domain", *International Conference on Image Processing*, Thessaloniki, Greece, 2001.
- [104] D. Liu, G. Liu, M. Yu, and Y. Wang, "An image retrieval method based on tree-structured wavelet transform", *International Conference on Computer Science and Software Engineering*, Wuhan, China, 2008.

- [105] Y. U. Tao, C. M. L. Ernest, and Y. T. Yuan, "A combination of fractal and wavelet for feature extraction", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 15, no. 8, pp. 1277-1298, 2001.
- [106] R. Methre and M. Ravindranath, "Exploring spatial information in spectral features for texture image retrieval", *International Journal of Computer and Network Security*, vol. 1, no. 3, pp. 46-49, 2009.
- [107] B. Luo, J. F. Aujol, Y. Gousseau, and S. Ladjal, "Indexing of satellite images with different resolutions by wavelet features," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1465-1472, 2008.
- [108] H. Yuan, X. P. Zhang, and L. Guan, "Content-based image retrieval using a Gaussian mixture model in the wavelet domain" *Proceeding of SPIE, Visual Communications and Image Processing*, Lugano, Switzerland, 2003.
- [109] R. Zhang, X.-P. Zhang, and L. Guan, "Wavelet-based texture retrieval using independent component analysis," *IEEE International Conference on Image Processing*, San Antonio, Texas, USA, 2007.
- [110] M. Kobayakwa, M. Hoshi, and T. Ohmori, "Robust texture image retrieval using hierarchical correlations of wavelet coefficients," *IEEE International Conference on Pattern Recognition*, Barcelona, Spain, 2000.
- [111] C. M. Pun and M. C. Lee, "Rotation invariant texture feature for content based image retrieval," *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland, 2002.
- [112] C. M. Pun, "Invariant content-based image retrieval by wavelet energy signatures," *IEEE International Conference on Acoustics, Speech and Signal Proc*, Hong-Kong, China, 2003.
- [113] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Rotationally invariant texture features using the dual-tree complex wavelet transform," *IEEE International Conference on Image Processing*, Vancouver, BC, Canada, 2000.
- [114] W. Dou, "Segmentation d'images multispectrales basée sur la fusion d'informations : application aux images IRM," *Thèse de doctorat de l'Université de Caen*, 2006.
- [115] S. D. Newsam and C. Kamath, "Retrieval using texture features in high resolution multi-spectral satellite imagery," *SPIE Defense and Security Symposium*, Orlando, Florida, USA, 2004.
- [116] L. Deren, X. Song, S. Haigang, and Z. Xiaodong, "Change detection based on spatial data mining," *ISPRS Workshop on Updating Geo-spatial Databases with Imagery*, Urumchi, Xingjiang, China, 2007.

- [117] F. Amato, C. Galdi, and G. Poggi, “Embedded zerotree wavelet coding of multispectral images,” *IEEE International Conference on Image Processing*, Washington, DC, USA, 1997.
- [118] A. Benazza-Benyahia, J. C. Pesquet, and M. Hamdi, “Vector lifting schemes for lossless coding and progressive archival of multispectral images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2011-2024, 2002.
- [119] A. Benazza-Benyahia and J. C. Pesquet, “Building robust wavelet estimators for multicomponent images using Stein’s principle,” *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1814-1830, 2005.
- [120] N. Khelil and A. Benazza-Benyahia, “Wavelet-based multivariate approach for multispectral image indexing,” *SPIE Conference on Wavelet Applications in Industrial Processing*, Philadelphia, USA, 2004.
- [121] M. Datcu and K. Seidel, “Human-centered concepts for exploration and understanding of earth observation images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 601-609, 2005.
- [122] G. Verdoolaege, S. De Backern, and P. Scheunders, “Multiscale color texture retrieval using the geodesic distance between multivariate generalized Gaussian models,” *IEEE International Conference on Image Processing*, San Diego, CA, USA, 2008.
- [123] R. B. Nelsen, *An introduction to copulas*, Lecture Notes in Statistics, vol. 139, Springer-Verlag, New York, 1999.
- [124] M. Fréchet, “Sur les tableaux de corrélation dont les marges sont données,” *Annales de l’Université Lyon*, Section A. 9, pp. 53-77, 1951.
- [125] A. Sklar, “Random variables, joint distribution functions and copulas,” *Kybernetika*, vol. 9, no. 6, pp. 449-460, 1973.
- [126] U. Cherubini, E. Luciano, and W. Vecchiato, *Copula methods in finance*, Wiley, 2004.
- [127] E. Bouye, V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, “Copulas for finance : a reading guide and some applications,” *GRO, Crédit Lyonnais*, 2000.
- [128] F. Flitti, Ch. Collet, and E. Slezak, “Astronomical multiband image fusion and restoration using pyramidal analysis and markovian segmentation,” *International Conference on Information Fusion*, Philadelphia, PA, USA, 2005.
- [129] N. Brunel, W. Piezcyński, and S. Derrode, “Copulas in vectorial hidden Markov chains for multicomponent image classification,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, Philadelphia, USA, 2005.
- [130] Y. Stitou, N. Lasmar, and Y. Berthoumieu, “Copulas based multivariate gamma modeling for texture classification,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, 2009.

- [131] T. S. Durrani and X. Zeng “Copula based divergence measures and their use in image registration,” *European signal processing conference*, Glasgow, Scotland, 2009.
- [132] X. Zeng and T. S. Durrani, “Band selection for hyperspectral images using copulas-based mutual information,” *IEEE Workshop on Statistical Signal Processing*, Cardiff, UK, 2009.
- [133] G. Mercier, “ Conditional copulas for change detection in heterogeneous remote sensing images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1428-1441, 2008.
- [134] S. Sakji and A. Benazza-Benyahia, “ Multivariate indexing of multichannel images,” *Procedure of the SPIE Conference on Wavelet Applications in Industrial Processing*, Boston, USA, 2007.
- [135] S. Sakji and A. Benazza-Benyahia, “Indexing of multichannel images in the wavelet transform domain,” *IEEE International Conference of Telecommunication Technology and Applications*, Damascus, Syria, 2008.
- [136] S. Sakji and A. Benazza-Benyahia, “Multivariate indexing of multichannel images based on the copula theory ,” *IEEE International Proceeding of Telecommunication and Applications*, Sousse, Tunisia, November 2008.
- [137] X. Chen and Y. Fan, “Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection,” *La Revue Canadienne de Statistique*, vol. 33, no. 3, pp. 389-414, 2005.
- [138] D. Pollard, “General chi-square goodness-of-fit tests with data-dependent cells,” *Probability Theory and Related Fields*, vol. 50, no. 3, pp. 317-331, 1979.
- [139] J. Dobric and F. Schmidt, “Testing goodness-of-fit for parametric families of copulas application to financial data,” *Communications in statistics, Simulation and computation*, vol. 34, no. 4, pp. 1053-1068, 2005.
- [140] D. Huard, G. Evin, and A C. Favre, “Bayesian copula selection,” *Computational Statistics and Data Analysis*, vol. 51, no. 2, pp. 809-822, 2006.
- [141] S. Sakji and A. Benazza-Benyahia, “Copula-based statistical models for multicomponent image retrieval using a Bayesian copula selection,” *International Symposium on Image and Signal Processing and Analysis*, Salzburg, Austria, 2009.
- [142] S. Sakji and A. Benazza-Benyahia, “Copula-based statistical models for multicomponent image retrieval in the wavelet transform domain,” *IEEE International Conference on Image Processing*, Cairo, Egypt, November 2009.
- [143] N. Kolev, U. Dos Anjos, and B. Vaz de M. Mendes, “Copulas : A review and recent developments,” *Stochastic Models*, vol. 22, no. 4, pp. 617-660, 2006.

- [144] S. Sakji and A. Benazza-Benyahia, "Multivariate indexing of multichannel images based on the vector lifting scheme," *IEEE International Geoscience and Remote Sensing Symposium*, Cape Town, South Africa, 2009.
- [145] S. Sakji and A. Benazza-Benyahia, "Rotation invariant multichannel image retrieval based on 2D dual-tree M-band wavelet decompositions," *International Symposium of Image Vision and Communication*, Bilbao, Spain, 2008.
- [146] S. Santini, R. Jain, "Similarity measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 871-883, 1999.
- [147] J. Zachary and S. S. Lengar, "Information theoretic similarity measures for content based image retrieval," *Journal of the American Society for Information Science and Technology*, vol. 52, no. 10, pp. 828-875, 2001.
- [148] J. D. Carroll and P. Arabie, "Multidimensional scaling," *Annual Review of Psychology*, vol. 31, pp. 607-649, 1980.
- [149] G. V. Wouwer, P. Scheunders, and D. V. Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 592-598, 1999.
- [150] G. H. Cha, "Non-metric similarity ranking for image retrieval," *International Conference on Database and Expert Systems Applications*, Krakow, Poland, 2006.
- [151] J. Yu, Q. Tian, J. Amores, and N. Sebe, "Toward robust distance metric analysis for similarity estimation," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, NY, USA, 2006.
- [152] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley, New-York, 1991.
- [153] S. K. Choy and C. S. Tong, "Statistical wavelet subband characterization based on generalized gamma density and its application in texture retrieval," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 281 - 289, 2010.
- [154] R. Kwitt and A. Uhl, "A joint model of complex wavelet coefficients for texture retrieval," *IEEE International Conference on Image Processing*, Cairo, Egypt, 2009.
- [155] S. Sakji and A. Benazza-Benyahia, "Fast scalable retrieval of multispectral Images with Kullback-Leibler divergence," *IEEE International Conference on Image Processing*, Hong-Kong, China, 2010.
- [156] A. Guttman, "R-trees : a dynamic index structure for spatial searching," *ACM SIGMOD International Conference on Management of Data*, Boston, USA, 1984.
- [157] N. Beckmann, P. Informatik, R. Schneider, and B. Seeger, "The R\*-tree : an efficient and robust access method for points and rectangles," *ACM SIGMOD international conference on Management of data*, Atlantic city, New Jersey, USA, 1990.

- [158] S. Berchtold, D. A. Keim, and H. P. Kriegel, "The X-tree : an index structure for high-dimensional data," *International Conference on Very Large Databases*, Bombay, India, 1996.
- [159] D. A. White and R. Jain, "Similarity indexing with the SS-tree," *International Conference on Data Engineering*, New Orleans, Louisiana, USA, 1996.
- [160] N. Katayama and S. Satoh, "The SR-tree : an index structure for high-dimensional nearest neighbor queries," *Proceedings of ACM SIGMOD : International Conference on Management of Data*, Tucson, Arizona, USA, 1997.
- [161] M. Majdoub, L. Fonteles, C. Benamar, and M. Antonini, "Fast indexing method for image retrieval using tree-structured lattices," *International Workshop on Content-Based Multimedia Indexing*, London, UK, 2008.
- [162] S. Sakji and A. Benazza-Benyahia, "A tree-structured image retrieval in the wavelet transform domain," *IEEE International Symposium of Signal Processing and Applications, ISSPA '07*, Sharjah, United Arab Emirates, 2007.
- [163] H. Frigui and R. Krishnapuram, "Clustering by competitive agglomeration," *Pattern Recognition*, vol. 30, no. 7, pp. 1109-1119, 1997.
- [164] R. Veldhuis, "The centroid of the symmetrical Kullback-Leibler distance," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 96-99, 2002.
- [165] N. Boujemaa, "On competitive unsupervised clustering," *Proceedings of the International Conference on Pattern Recognition*, Barcelona, September, Spain, 2000.
- [166] J .C. Bezdek , "Pattern recognition with fuzzy objective function algorithms," *Plenum Press*, 1981.