

THESE

Présentée

A L'UNIVERSITE PARIS XII VAL-DE-MARNE

Pour l'obtention

**DU DIPLOME DE DOCTEUR D'UNIVERSITE
EN SCIENCES**

Spécialité : Traitement d'images

Par

Georges L. LOUM

**SEGMENTATION
PYRAMIDALE DE TEXTURES
PAR DECOMPOSITION EN ONDELETTES**

soutenue le 25 Juillet 1996 devant le jury composé de :

Président	:	Jacques LEMOINE
Rapporteurs	:	Olivier MONGA Mohamed NAJIM
Directeur	:	Pierre PROVENT
Examineurs	:	Fernand MEYER Eric PETIT

*Chaque maison est construite par quelqu'un mais
celui qui a construit toutes choses, c'est Dieu.*

(Heb.3 v. 4, LA BIBLE)

A mon Dieu, le Dieu de Jésus-Christ

Remerciements

J'exprime ma profonde gratitude au Professeur Maurice Gaudaire pour m'avoir accueilli au sein du Laboratoire d'Etudes et de Recherche en Instrumentation Signaux et Systèmes (LERISS) de l'Université Paris XII Val-de-Marne.

Je témoigne toute ma reconnaissance à Monsieur Nahounou Bobouo, Directeur de l'INSET de Yamoussoukro, qui a bien voulu accepter que cette thèse soit préparée alternativement entre l'INSET de Yamoussoukro et l'Université Paris XII Val-de-Marne.

Je suis profondément reconnaissant à Monsieur Olivier Monga, Directeur de recherche à l'INRIA Rocquencourt et au Professeur Mohamed Najim de l'ENSERB de Bordeaux qui, malgré leur emploi du temps très chargé, ont accepté d'être les rapporteurs de ce travail.

Je remercie sincèrement Monsieur Fernand Meyer de l'Ecole Nationale Supérieure des Mines de Paris de participer à mon jury et de porter un intérêt à ce travail.

Je remercie le Professeur Eric Petit pour son soutien et ses conseils tout au long de la préparation de cette thèse.

Je suis fortement redevable à Jacques Lemoine qui n'a ménagé aucun effort sur le plan administratif pour que la préparation de cette thèse en alternance soit possible. Je lui suis reconnaissant pour la grande confiance qu'il a sue me témoigner et pour le regard critique qu'il a constamment jeté sur ce travail.

Je remercie très vivement Pierre Provent pour avoir dirigé cette thèse. Les échanges fructueux que nous avons eus et ses nombreux conseils m'ont été d'un apport considérable.

Je tiens à remercier tous les membres de l'équipe Image et Signal du LERISS pour leur amitié et leur franche collaboration.

Mes collègues et amis du Département Génie Electrique de l'INSET de Yamoussoukro trouveront ici l'expression de ma reconnaissance pour leur appui et leur soutien sans faille.

Que Mesdames Patricia Jamin et Annick Joncour soient remerciées pour leur grande sympathie, leur constante disponibilité et l'ambiance agréable qu'elles maintiennent au sein du laboratoire.

Je ne saurais oublier Mademoiselle Marianne Simonnot du CIES et tous les membres du Service des Bourses de la Coopération Française d'Abidjan, pour leurs aides précieuses.

Le soutien de ma femme Huguette tout au long de la préparation de cette thèse m'a été d'un grand réconfort. Je tiens ici à lui témoigner, ainsi qu'à ma fille Anne, à mes parents, et à mes amis, toute ma reconnaissance et mon profond attachement.

Enfin, s'il y a quelque satisfaction à tirer de ce travail, s'il y a quelque joie à éprouver au terme de ce travail, que toute la gloire et l'honneur en reviennent à Dieu sans qui, pas un mot de cette thèse n'aurait pu être écrit.

Table des Matières

Introduction	4
I Le point sur l'analyse de texture	5
1 Notion de texture	5
1.1 Définition de la notion de texture	5
1.1.1 La texture, une notion complexe	7
1.1.2 Les primitives	7
1.1.3 Organisation spatiale des primitives	8
1.1.4 Echelle d'analyse de la texture	8
1.1.5 Notion de voisinage	9
1.2 Attributs de texture.	9
1.2.1 Finalité d'une analyse de texture.	10
1.2.2 Détermination des attributs de texture	12
2 Texture et perception visuelle humaine	13
2.1 Le système visuel humain	15
2.2 Définition d'un ensemble de paramètres discriminants de texture	15
2.2.1 Position du problème	16
2.2.2 Paramètres discriminants de textures.	18
2.3 Les textons, éléments de base de la perception de la texture	19
2.4 Représentation de l'image rétinienne dans le cortex visuel	19
2.4.1 Décomposition multicanaux de l'information visuelle	
2.4.2 Modélisation de la décomposition multicanaux de l'information visuelle.	20 23
2.4.3 Intérêt du modèle multicanaux pour l'analyse de la texture	24
2.5 Conclusion	26
3 Méthodes classiques d'analyse de textures	27
3.1 Les méthodes structurelles.	27
3.1.1 Principe.	27
3.1.2 Extraction des primitives	29
3.1.3 Détermination des règles de placement des primitives	29
3.2 Les méthodes statistiques	30
3.2.1 Les méthodes de dépendance spatiale	33
3.2.2 Les méthodes d'analyse de texture fondées sur des modèles	36
3.2.3 Conclusion sur les méthodes statistiques.	37
3.3 Les méthodes d'analyse par filtrage spatial	37
3.3.1 Les transformations linéaires locales	39
3.3.2 Le Filtrage spatial multicanaux	41
3.3.3 Conclusion sur le filtrage spatial multicanaux	

II Ondelettes et Analyse de textures	43
4 Transformations en Ondelettes	44
4.1 Définition des ondelettes	45
4.1.1 Ondelettes continues	45
4.1.2 Transformée en ondelettes continue	46
4.1.3 Ondelettes orthogonales	48
4.2 Transformations en ondelettes et Analyse multirésolution	49
4.2.1 Définition des fonctions d'échelle	49
4.2.2 Fonctions d'échelle et fonctions ondelette associées	50
4.2.3 Décomposition en ondelettes orthogonales	52
4.2.4 Mise en oeuvre algorithmique de la décomposition	54
4.2.5 Extension de la TOD au cas bidimensionnel 2D	55
4.2.6 Comparaison de la TOD avec d'autres méthodes de représentation pyramidale	58
4.3 Autres types de transformations en ondelettes	59
4.3.1 Analyse multirésolution orthogonale en quinconce	59
4.3.2 Ondelettes biorthogonales	59
4.3.3 Ondelettes obliques ("wavelet frames")	60
4.3.4 Paquets d'ondelettes	61
4.4 Applications des ondelettes à l'analyse de textures	64
5 Caractérisation et classification de textures	67
5.1 Une nouvelle méthode de caractérisation de textures fondée sur la décomposition en ondelettes	68
5.1.1 Description de l'approche	68
5.1.2 Méthode de caractérisation des textures	70
5.1.3 Voisinage pyramidal	71
5.1.4 Définition du facteur de forme	72
5.1.5 Formule simplifiée et récursivité du calcul du facteur de forme	75
5.2 Choix de l'ondelette analysante	76
5.3 Caractérisation de textures à l'aide du facteur de forme	78
5.4 Algorithmes de classification de textures	83
5.4.1 Premier algorithme de classification	84
5.4.2 Deuxième algorithme de classification	93
5.5 Conclusion	101
6 Segmentation pyramidale de textures	103
6.1 Techniques de segmentation fondée sur une approche par ondelettes	104
6.2 Intérêt d'un voisinage pyramidal pour une segmentation multirésolution d'une image texturée	106
6.3 Segmentation pyramidale de textures	109
6.3.1 Principe de mise en oeuvre	109
6.3.2 Description de l'algorithme de segmentation	110
6.3.3 Expérimentations et résultats	116
6.4 Conclusion	129
Conclusion	132
Bibliographie	135

Liste des Figures

	6
1.1 Exemples de textures	9
1.2 Exemples de voisinages	10
1.3 Stratégie d'extraction d'attributs de texture	14
2.1 Schématisation du système visuel humain	15
2.2 Textures de même histogramme	20
2.3 Exemples de réseaux spatiaux	21
2.4 Modèle multicanaux du traitement visuel humain	31
3.1 Une image et deux de ses matrices de cooccurrences spatiales	38
3.2 Masques bidimensionnels de Laws les plus utilisés	46
4.1 Forme d'une cellule de résolution d'ondelettes dans l'espace temps-fréquence	47
4.2 Ondelette de Meyer	
4.3 Exemple de calcul de la transformée en ondelettes continue d'un signal avec l'ondelette de Haar	51
4.4 Fonction d'échelle associée à l'ondelette de Meyer	53
4.5 Schéma de décomposition en ondelettes d'un signal	55
4.6 TOD et structure du banc de filtres	
4.7 Décomposition en ondelettes au niveau de résolution 1 d'une mosaïque de textures	57
4.8 Schéma de calcul de la TOD d'une image	58
4.9 Construction d'une pyramide par décomposition en ondelettes d'une image	63
4.10 Arbres binaires correspondant aux paquets d'ondelettes	65
4.11 "The Tree-Structured Wavelet Transform"	69
5.1 Décomposition en séries d'ondelettes de Haar et de Fourier d'un signal $S(t)$	72
5.2 Exemple de voisinage pyramidal	
5.3 Images considérées pour le calcul du facteur de forme au niveau de résolution 3 et dans la direction $l=1$	76
5.4 Pixel (k) du niveau j et ses quatre correspondants (a,b,c,d) du niveau inférieur	79
5.5 Images test utilisées dans l'expérimentation de caractérisation	86
5.6 Texture naturelles utilisées dans les expérimentations de classification	89
5.7 Les 3 filtres de la DCT utilisés	98
5.8 Images de textures synthétiques	
6.1 Schéma de principe d'une segmentation de textures fondée sur une analyse par ondelettes	105
6.2 Exemples de support d'un voisinage plan et d'un voisinage pyramidal	116
6.3 Images test de la segmentation	118
6.4 Résultat de la segmentation de Image1	119
6.5 Résultat de la segmentation de Image1 moyennée	120
6.6 Résultat de la segmentation de Image2	120
6.7 Résultat de la segmentation de Image3	123
6.8 Résultats de la segmentation de Image1 à différents niveaux de résolutions	126
6.9 Résultats de la segmentation obtenus en utilisant un voisinage plan	127
6.10 Résultats de la segmentation obtenus avec deux ondelettes différentes	

Liste des Tableaux

5.1	Facteurs de forme moyen de différents niveaux de résolution et de différentes directions obtenus avec l'ondelette de Haar	80
5.2	Coefficients des filtres associés à l'ondelette de Daubechies à deux moments nuls	81
5.3	Facteurs de forme moyen de différents niveaux de résolution et de différentes directions obtenus avec l'ondelette de Daubechies	82
5.4	Résultats de la classification pour différents niveaux de résolution	87
5.5	Matrice de confusion des textures "sable fin", "laine", et "canevas"	88
5.6	Tableau récapitulatif des paramètres des trois méthodes implantées.	91
5.7	Résultats de la classification fournis par les trois méthodes	91
5.8	Variance des images basses résolution des textures test	97
5.9	Variance des images basses résolutions des textures synthétiques	99
5.10	Résultats du second algorithme de classification	101
6.1	Nombre et pourcentage des pixels ambigus de Image1 pour différents niveaux de résolution	121
6.2	Résultats de la segmentation des images test	122
6.3	Résultats de la segmentation réalisée à différents niveaux de résolution	125
6.4	Tableau récapitulatif des résultats obtenus avec deux ondelettes différentes	128

Graphique

5.1	Courbes d'évolution de la variance des images basses résolutions de quelques textures	97
-----	---	----

Segmentation pyramidale de textures par décomposition en ondelettes

Résumé

Cette thèse présente une méthode de segmentation pyramidale d'images d'aspect texturé par décomposition en ondelettes. D'une façon générale, tout processus de segmentation de textures comporte une étape de caractérisation qui précède l'étape de segmentation proprement dite.

Notre méthode de caractérisation de textures est fondée sur l'interprétation des coefficients de détails de la décomposition en ondelettes, comme décrivant les variations locales des niveaux de gris de la texture autour de leur valeur moyenne. Cette interprétation conduit à la définition d'un nouvel attribut, le "facteur de forme", qui se calcule sur un nouveau voisinage de type pyramidal.

Pour éprouver la pertinence de ce nouvel attribut, deux algorithmes de classification supervisée de textures sont proposés. Le premier exploite la représentation multirésolution de l'analyse par ondelettes pour réaliser une caractérisation efficace des textures à plusieurs niveaux de résolution. Le second détermine pour chaque texture, un niveau maximal de la décomposition en ondelettes, en définissant un seuil sur les variances des images d'approximation. Ce niveau est utilisé pour réaliser une préclassification, avant que le processus de classification ne soit conduit à son terme dans chaque classe constituée. Les résultats satisfaisants obtenus par ces deux algorithmes ont validé la méthode de caractérisation proposée et démontré la pertinence de l'attribut facteur de forme.

La méthode de segmentation présentée, tire profit de la forme pyramidale du voisinage sur lequel le facteur de forme est calculé. Cette forme constitue un véritable atout pour un processus de segmentation qui se fonde sur une représentation multirésolution des données. Elle permet de réduire sensiblement le nombre de pixels de l'image d'attributs représentant de zones inter-régions et minimise de ce fait, l'ambiguïté sur la localisation précise des frontières. De plus, les dimensions variables du voisinage pyramidal en fonction du niveau de résolution, suggèrent l'élaboration d'un processus de segmentation évoluant suivant une stratégie du plus grossier au plus fin. Les meilleurs prototypes de chaque classe de texture sont déterminés à l'aide d'un algorithme qui met en oeuvre un classificateur flou. Ces prototypes permettent de réaliser une segmentation grossière au niveau de résolution le plus élevé. Cette segmentation primaire est progressivement affinée lors de la descente de la pyramide en dirigeant le processus de segmentation vers les zones de singularités (zones frontalières ou bruitées). La méthode proposée fournit les meilleurs résultats et réalise une segmentation sans recouvrement des régions lorsque la décomposition de l'image est effectuée avec l'ondelette de Haar.

Mots clés :

Segmentation, Textures, Ondelettes, Voisinage pyramidal, Facteur de forme, Classification floue, Analyse multirésolution.

INTRODUCTION

INTRODUCTION

La segmentation d'une image est souvent considérée comme une étape préliminaire importante dans un processus de traitement d'images qui a pour vocation de reconnaître des objets ou d'interpréter une scène.

Elle consiste à partitionner l'image en différentes régions homogènes au sens d'un ou de plusieurs critères donnés, en vue d'obtenir une description compacte et exploitable de celle-ci. Ce partitionnement peut être réalisé en détectant les discontinuités locales (approche contours) ou en cherchant à identifier les zones homogènes (approches région).

L'homogénéité d'une région est souvent associée à la notion de texture. La texture qui est difficile à définir et à quantifier, est généralement caractérisée par un ensemble de paramètres ou attributs calculés dans le voisinage immédiat d'un point-image (appelé également pixel : picture élément). La pertinence de cette description paramétrique influence fortement la qualité de la segmentation.

Le problème de la segmentation des textures a été diversement abordé. Plusieurs états de l'art sont disponibles dans la littérature [Ree.93, Coc.95]. Les techniques de segmentation les plus récentes, intègrent les nouveaux outils mathématiques d'analyse qui possèdent des propriétés particulièrement adaptées à la segmentation des textures.

Les ondelettes par exemple, présentent plusieurs propriétés intéressantes pour ce type d'application. Elles réalisent une analyse multirésolution sensible à l'orientation qui peut être mise à profit, pour différencier les textures d'une image, par la taille et l'orientation de leurs éléments. En outre, elles constituent un bon modèle mathématique du système visuel humain

dont les performances sont remarquables dans la perception et la discrimination des textures. Ces propriétés justifient l'attention particulière dont les ondelettes sont l'objet depuis leur introduction.

Le travail de recherche que nous présentons, se fonde sur une décomposition en ondelettes pour réaliser la segmentation des textures.

L'application des ondelettes à la segmentation des textures a été initialement suggérée par Mallat [Mal.89a]. Les méthodes qui ont été développées par la suite, exploitent les informations des différents niveaux de résolution pour réaliser la caractérisation des textures puis la segmentation de l'image.

L'exploitation de ces informations est cependant réalisée en traitant indépendamment les différents niveaux de résolution; d'où les difficultés rencontrées pour fusionner l'ensemble des informations obtenues séparément à chaque niveau de résolution.

Nous proposons une nouvelle méthode de caractérisation de la texture qui se singularise par la prise en compte simultanée des informations contenues dans les différents niveaux de résolution de la décomposition en ondelettes. La combinaison de ces informations trouve un cadre justificatif dans une interprétation particulière des coefficients d'ondelettes.

L'approche adoptée impose la définition d'un nouvel attribut, le facteur de forme, et l'introduction d'un nouveau type de voisinage, à savoir le voisinage pyramidal, sur lequel le facteur de forme est calculé.

La localisation des frontières des régions texturées rend les problèmes de segmentation de textures difficiles. Les attributs extraits dans les zones proches des frontières sont généralement ambigus. Cela est dû au fait que dans ces zones, les voisinages définis pour la détermination des attributs recouvrent deux ou plusieurs régions de textures différentes. Ce phénomène est amplifié lorsque les données sont représentées à plusieurs niveaux de résolution.

Le voisinage pyramidal qui induit aux niveaux de résolution élevés, un support de taille moins importante par rapport à celle du voisinage plan classique, présente de ce fait, un grand intérêt pratique pour une segmentation de textures. Nous exploitons cet avantage ainsi que la pertinence du facteur de forme pour proposer un algorithme de segmentation pyramidal de textures.

Ce document est organisé en deux parties composées chacune de trois chapitres :

La première partie intitulée "Le point sur l'analyse de textures" propose une synthèse de notre recherche bibliographique sur l'analyse de textures. Le premier chapitre présente les problèmes rencontrés pour établir une définition précise de la notion complexe de la texture et évoque la stratégie générale adoptée pour sa caractérisation. La perception des textures par le système visuel humain fait l'objet du second chapitre. Ce chapitre propose une revue non exhaustive des principaux résultats des recherches relatives à la formidable capacité de l'homme à différencier et à reconnaître aisément différentes textures. Quant au troisième chapitre, il présente au lecteur les méthodes d'analyse de textures les plus représentatives de l'état de l'art actuel.

Le quatrième chapitre qui entame la seconde partie intitulée "Ondelettes et analyse de textures", est un chapitre introductif à la présentation de notre travail de recherche. Il présente la théorie des ondelettes et le lien avec l'analyse multirésolution. Le cinquième chapitre est consacré à la présentation de la méthode de caractérisation des textures proposée. La pertinence du facteur de forme et l'efficacité de la méthode proposée est vérifiée expérimentalement au moyen de deux algorithmes de classification différents qui tirent leur inspiration des propriétés de l'analyse par ondelettes. Dans le dernier et sixième chapitre, les propriétés du facteur de forme et de son voisinage de définition sont exploitées en vue de la segmentation des textures.

Ière PARTIE

LE POINT SUR L'ANALYSE DE TEXTURE

1

NOTION DE TEXTURE

La texture est un concept qui traduit un aspect visuel homogène de la surface d'un objet. Son analyse revêt un grand intérêt pratique pour de nombreuses applications. Elle est impliquée notamment dans l'imagerie médicale, l'aide au diagnostic, l'analyse du tissu urbain, la prospection des ressources terrestres, et le contrôle de qualité des produits industriels.

1.1 Définition de la notion de texture

1.1.1 La texture, une notion complexe

Des textures, dont quelques exemples sont montrés à la *figure 1.1*, peuvent être aisément distinguées dans des scènes naturelles ou dans divers types d'images (médicales, aériennes, satellitaires et industrielles).

Bien que facile à reconnaître, la texture est une notion complexe et difficile à définir. Sa complexité réside dans :

- *La grande variété de textures* : elles sont naturelles (sable, gazon), ou synthétiques (hachures, points, croix), structurées (mur de briques, moquette), ou aléatoires (peau d'orange, bruit apparaissant sur un écran de télévision),
- *La grande subjectivité rattachée à cette notion* : Les textures de sables, de siège, et de lierre de la *figure 1.1*, présentent pour l'observateur, des aspects différents qui font appel à des critères qualitatifs et fortement subjectifs que sont finesse, granularité, régularité, linéarité, etc... L'évaluation quantitative de ces critères nécessite une connaissance approfondie des mécanismes de la perception visuelle humaine. Ceux-ci sont malheureusement encore mal connus à l'heure actuelle.

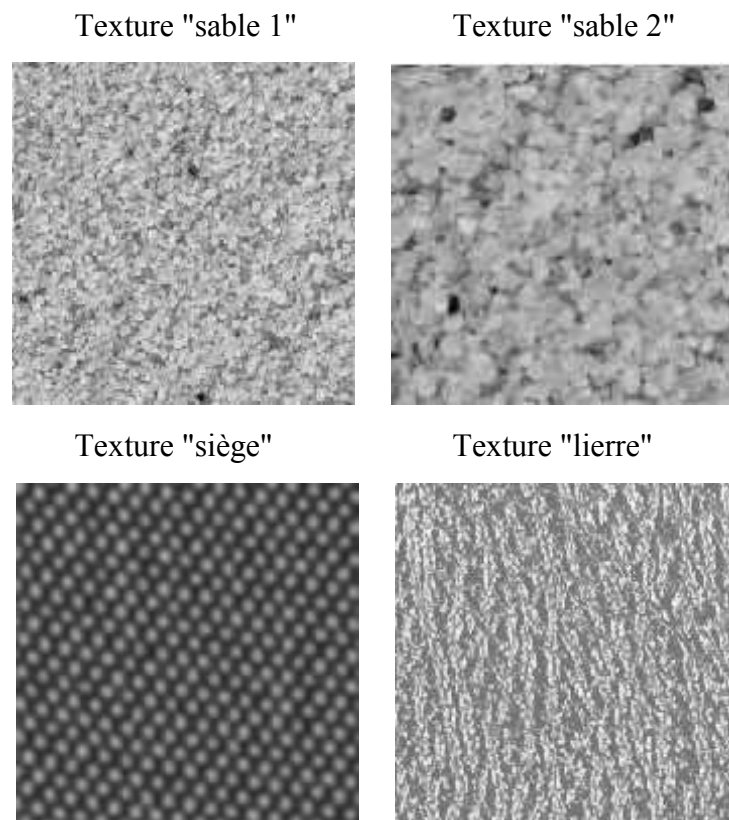


Figure 1.1 : Exemples de textures

Ces deux facteurs essentiels rendent difficile et quasiment impossible l'élaboration d'une définition formelle et universelle de la texture.

Les définitions publiées dans des ouvrages [Pra.91] et articles traitant du sujet, précisent de façon générale que *la texture d'une surface se traduit par une organisation*

spatiale plus ou moins homogène d'éléments simples appelés motifs ou encore primitives [Pos.87].

Ces définitions mettent en exergue ce que Haralick considère comme les deux dimensions de base de la texture à savoir, les primitives et leur organisation spatiale [Har.79]. Ces dimensions font respectivement intervenir les concepts de *tonalité* et de *texture*.

Le premier concept est rattaché à la description de primitives dites *tonales*. Celles-ci sont définies à partir des variations locales des tons de gris. Le second décrit la distribution spatiale de ces primitives sur une région plus étendue. Ces deux aspects sont toujours présents dans une image bien que l'un puisse prendre de l'ascendance sur l'autre.

1.1.2 Les primitives

Selon le type et la nature de la texture, les primitives diffèrent par leur taille, leur forme et leur orientation (*figure 1.1*).

Leur description est intimement liée à la résolution de l'image c'est-à-dire au nombre moyen de points qui les représente [Bal.82]. Si ce nombre est important, la primitive sera décrite de façon détaillée. Dans le cas contraire, il sera impossible de la décrire correctement. Il est en effet plus facile, vue la quantité d'informations dont on dispose, de décrire avec beaucoup plus d'exactitude les primitives de la texture "sable 2" de la *figure 1.1*, que celles de la texture "sable 1".

1.1.3 Organisation spatiale des primitives

L'organisation spatiale des primitives peut obéir à une loi déterministe. Le motif de base de la texture est alors répété, de façon plus ou moins régulière, dans différentes directions de l'espace. C'est le cas des primitives de la texture "tissu siège" de la *figure 1.1*. Ces textures sont plus ou moins structurées et quelques fois périodiques.

L'organisation des primitives peut également obéir à une loi probabiliste, à l'instar de celle de la texture "sable 1". Ce type de textures présente un aspect désordonné, voire aléatoire.

Les primitives de certaines textures en revanche, n'obéissent à aucune de ces deux lois organisationnelles. Elles sont "peu ordonnées", et présentent à la fois un aspect aléatoire et déterministe. La texture "sable 2" en est un exemple.

1.1.4 Echelle d'analyse de la texture

Si l'on observe la texture "sable 2" en s'éloignant progressivement, on peut remarquer que ses grains grossiers apparaissent de plus en plus fins et de plus en plus ressemblants à ceux de la texture "sable 1".

Cette expérience montre que l'aspect d'une texture change selon l'échelle à laquelle elle est observée. En réalité, l'aspect structuré ou aléatoire d'une texture, correspond à une vision macroscopique ou microscopique de cette dernière.

Il apparait donc important, pour toute analyse de texture, de choisir une échelle appropriée, qui soit en adéquation avec la taille des éléments de texture.

Indiquons à toutes fins utiles, que la notion d'échelle est différente de celle de résolution : la première est rattachée à la représentation de l'image, tandis que la seconde est liée à la quantité d'informations présente dans l'image.

1.1.5 Notion de voisinage

Les différents types de textures possèdent une propriété essentielle qui leur est commune : l'impression visuelle que l'on éprouve en observant chacune d'elles, reste la même quelle que soit la partie observée.

Cette homogénéité particulière que présente en ses différentes régions une texture donnée, permet de lier la notion de texture à celle d'une information locale de voisinage.

Ainsi, la texture sera décrite par un ensemble de paramètres caractéristiques (ou d'attributs) mesurés ou calculés dans le voisinage immédiat (*figure 1.2*) de chaque point-image (pixel) $x(k,l)$. Cet ensemble d'attributs est généralement regroupé sous la forme d'un vecteur appelé vecteur d'attributs.

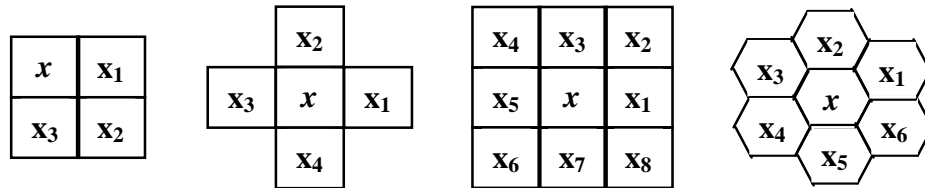


Figure 1.2 : Exemples de voisinages (les tailles peuvent être plus grandes)

1.2 Attributs de texture

1.2.1 Finalité d'une analyse de texture

La finalité d'une analyse de texture est multiple. Elle peut être :

- soit une caractérisation de texture qui conduit à la détermination d'un ensemble de mesures pertinentes, caractéristiques d'une texture donnée,
- soit une classification de textures qui permet la différenciation et la reconnaissance d'une surface donnée parmi un ensemble prédéfini de textures,
- soit encore une segmentation d'image texturée qui consiste en un partitionnement de l'image en régions de texture homogènes.

Le résultat de chacune de ces opérations, est fortement conditionné par la détermination et le choix des attributs de texture.

1.2.2 Détermination des attributs de texture

La stratégie d'extraction de ces attributs a été initialement suggérée par Haralick, Shanmugan, et Dinstein [Har.73]. Elle est schématisée par la *figure 1.3*.

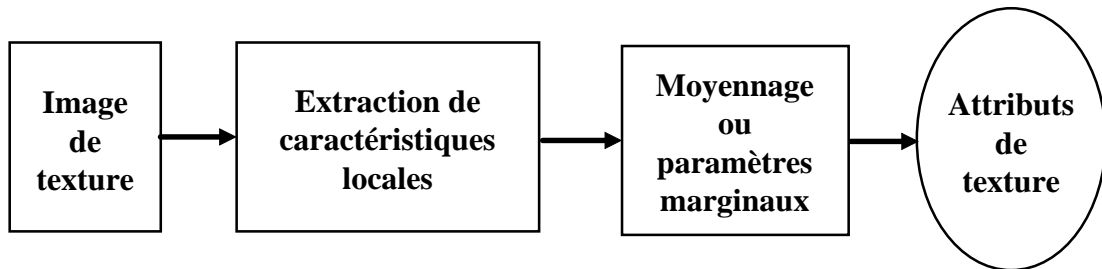


Figure 1.3 : *Stratégie d'extraction d'attributs de texture*

La première étape de cette stratégie consiste à réaliser une analyse locale qui permet d'extraire un ensemble de paramètres représentatifs de la structure de la texture. Ces paramètres constituent les points d'entrée d'une seconde phase nécessaire pour la prise en compte de la répartition spatiale des différents tons de gris. Au cours de cette phase, des mesures globales marginales sont calculées sur un voisinage plus important. Ces mesures fournissent un ensemble fini d'attributs de texture.

Ces attributs rendent compte de la répartition et de la disposition des niveaux de gris sur la surface texturée considérée et permettent de la caractériser. Ils sont en général, la traduction d'une propriété géométrique, probabiliste, ou fonctionnelle telle que l'orientation, le contraste, l'homogénéité, la rugosité, la variance,....

Pour éviter les temps de calcul élevés, et garantir la qualité de l'analyse, les attributs doivent être les plus pertinents possibles et en nombre réduit [Kas.86].

En général, le choix des attributs dépend de l'image à analyser, de l'application envisagée, des connaissances disponibles a priori, et de la vitesse d'exécution recherchée. C'est pourquoi il est difficile de trouver un ensemble d'attributs qui satisfasse tous les cas d'analyse. Plusieurs auteurs ont essayé d'apporter une solution à ce problème en proposant des algorithmes de sélection automatique des "meilleurs" attributs pour une banque d'images donnée [He.88]. Toutefois, il est nécessaire d'indiquer que pour un problème d'analyse donné,

seuls les résultats obtenus (a posteriori malheureusement!), peuvent témoigner de la justesse du choix effectué.

Bien qu'il ne semble pas se dégager de règles qui régissent le choix des attributs, il est généralement admis qu'un bon ensemble d'attributs, doit :

- être la résultante d'une analyse locale et d'une analyse globale,
- contenir des informations structurelles et statistiques,
- prendre en considération les aspects de la perception visuelle humaine.

La première composante prend en compte les éléments de texture et leur contexte spatial. La seconde rend compte des variations des tons de gris sur des voisinages de tailles différentes. Enfin, si l'on désire avoir les attributs les plus pertinents, il est tout naturel de s'inspirer des mécanismes de reconnaissance du système visuel humain.

Ces mécanismes sont, il est vrai, d'une grande complexité et restent encore imparfaitement connus. Cependant, les résultats déjà obtenus dans divers domaines (psychophysique, physiologie, neurobiologie), ont notamment conduit à une meilleure compréhension du mode de perception de la texture par le système visuel humain. La présentation de quelques uns de ces résultats, fait l'objet du chapitre suivant.

2

TEXTURE ET PERCEPTION VISUELLE HUMAINE

Plusieurs chercheurs ont essayé et essaient encore de sonder les mécanismes complexes de la perception visuelle humaine.

Leur principal objectif est d'étudier la capacité de l'homme à reconnaître, différencier, et interpréter, avec une facilité apparente, les différentes régions d'une image. L'enjeu de l'acquisition d'une telle connaissance pour le traitement d'images, est la mise au point de systèmes d'analyse qui vus de l'extérieur, auraient un fonctionnement similaire à celui du système visuel humain.

Les travaux réalisés à ce jour bien qu'encore incomplets, ont apporté quelques éclaircissements sur le processus de traitement de l'information visuelle. Certains de ces travaux ont notamment permis de préciser quels sont les mécanismes de perception et d'analyse de la texture qui joue un rôle important dans l'interprétation d'une scène ou d'une image.

Ce chapitre qui débute par une description schématique de l'appareil visuel humain, présente quelques uns des principaux résultats des recherches consacrées à la perception de la texture.

2.1 Le système visuel humain

La constitution de l'oeil est comparable à celle d'un capteur. Le cristallin est un système optique convergent. Il forme sur la rétine, surface concave photosensible, l'image des objets placés devant lui. L'intensité lumineuse qui arrive sur la rétine dépend de l'ouverture de l'iris. Ce dernier est un diaphragme dont le centre est percé d'un orifice dilatable appelé pupille.

La rétine renferme un grand nombre de cellules détectrices qui fournissent à chaque instant au cerveau des informations correspondant à l'intensité, à la couleur et à la position des sources lumineuses.

Il existe trois populations de cellules au niveau de la rétine : les cellules photoréceptrices, les cellules bipolaires et les cellules ganglionnaires. Les cellules photoréceptrices (sensibles à la lumière) sont de deux types : les *cônes* et les *bâtonnets*.

Les cônes possèdent une liaison individuelle avec le cerveau. Cette liaison est réalisée au niveau de la fovéa où chaque cône est relié à une fibre nerveuse. Ils favorisent les informations les plus précises en résolution et en sensibilité (à cause de leur grande densité dans la fovéa), et jouent un rôle important dans la vision diurne. Ils sont également responsables de la vision des couleurs.

Les bâtonnets, environ vingt fois plus nombreux que les cônes, se situent essentiellement en dehors de la fovéa et sont regroupés entre eux avant d'être reliés au cerveau. Ils sont sensibles aux faibles niveaux lumineux, d'où leur importance dans la vision nocturne. Aux faibles luminosités, les bâtonnets fournissent des informations concernant la brillance moyenne des objets et non leurs couleurs.

Les informations recueillies par les bâtonnets et les cônes, sont acheminées pour traitement vers les cellules bipolaires puis ganglionnaires. Elles traversent ensuite le nerf optique et les corps genouillés latéraux droit et gauche, avant d'être transmises au cortex strié (*figure 2.1*).

L'action des cellules bipolaires et ganglionnaires est complexe à définir. Notons simplement que les cellules ganglionnaires jouent un rôle très important dans la perception des contrastes et dans la détermination de la direction des mouvements.

Le cortex est constitué de cellules simples et complexes [Hub.62]. Les premières jouent comme nous le verrons plus loin, un rôle crucial dans le traitement visuel primaire. Les secondes sont impliquées dans la détection du mouvement car elles répondent favorablement à des flash ou à des stimuli mobiles dans une direction donnée.

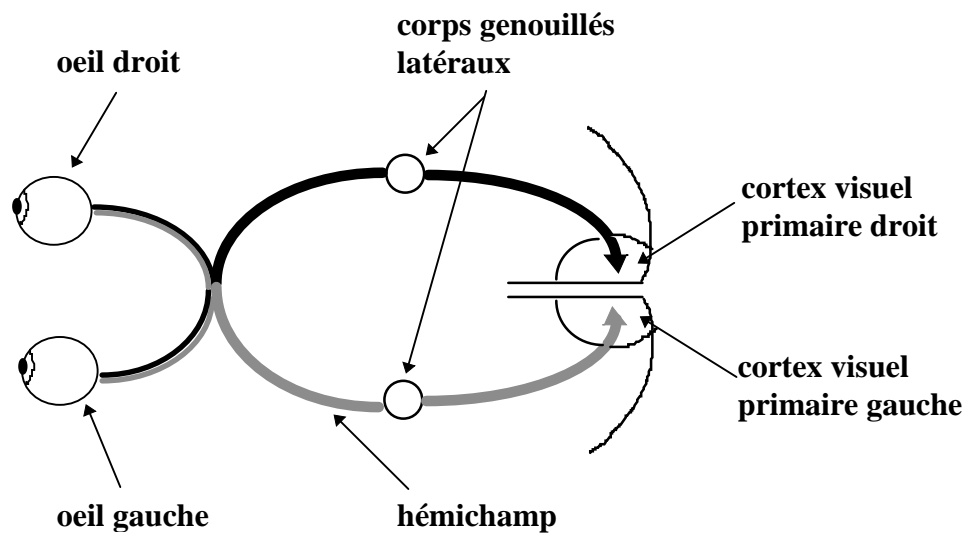


Figure 2.1 : Schématisation du système visuel humain

Tout au long de son acheminement de la rétine au cortex, l'information subit un traitement primaire de "bas niveau" réalisé par différents groupes de neurones repartis en différentes couches. A l'issue de ce traitement, qui a pour objectif l'extraction de primitives pertinentes, l'information visuelle est intégrée à un niveau supérieur du chemin visuel, pour y subir un traitement plus élaboré dit de "haut niveau". Ce dernier traitement permet d'aboutir à l'interprétation de la scène visuelle.

Les différentes parties du système visuel (système optique, cellules photoréceptrices, groupes de neurones, ...) constituent des modules spécialisés qui traitent l'information de façon séquentielle ou parallèle, avec plusieurs interactions entre eux. Modularité et

parallélisme sont donc les deux caractéristiques de base de l'architecture fonctionnelle du système visuel.

Ce système au demeurant complexe, présente des performances remarquables particulièrement en ce qui concerne la vision de la texture. C'est donc à juste titre qu'il a suscité la curiosité de certains chercheurs qui ont entrepris de comprendre et de modéliser les mécanismes qui lui permettent, entre autres, de percevoir et de discriminer aussi aisément les textures. Julesz est l'un des premiers à s'être intéressé à la perception humaine de la texture. En 1962, à partir d'expériences perceptuelles humaines, il a cherché à déterminer des paramètres statistiques discriminants de texture.

2.2 Définition d'un ensemble de paramètres discriminants de texture

2.2.1 Position du problème

Une image de texture peut être caractérisée par une description statistique de la distribution de ses niveaux de gris. Les statistiques du premier ordre tel que l'histogramme, ne sont pas suffisantes pour les caractériser. Elles ne prennent pas en compte les informations de connexité qui sont importantes pour la discrimination des textures. L'exemple de deux textures présentées à la *figure 2.2* est significatif à ce sujet. Bien qu'elles soient visuellement différentes, ces textures ont le même histogramme (c'est-à-dire que l'ensemble de leurs moments (moyenne, variance,...) sont identiques jusqu'à l'ordre N).

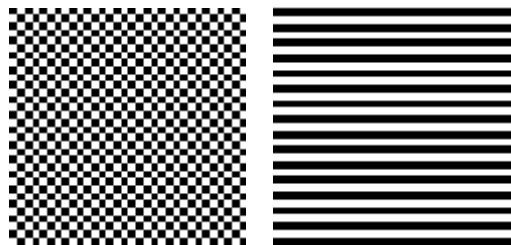


Figure 2.2 : Textures de même histogramme

Cet exemple montre que :

- il est indispensable voire nécessaire, pour discriminer deux textures, de s'intéresser aux statistiques d'ordre supérieur, qui décrivent les relations existantes entre paires ou entre groupe de pixels.
- la texture ne peut pas être décrite par un seul point-image. Par conséquent, il est utile de tenir compte dans l'évaluation des paramètres de texture, de la dépendance spatiale entre les niveaux de gris dans un voisinage donné.

Le problème à résoudre est de déterminer l'ordre des statistiques auquel le système visuel est sensible, et de préciser l'étendue du voisinage (local ou global) dans lequel ce système considère de tels statistiques.

2.2.2 Paramètres discriminants de textures

En 1962, Julesz s'est intéressé aux statistiques du second ordre pour la discrimination des textures. Il a montré expérimentalement que l'homme ne peut discriminer sans effort des textures de distributions de probabilité égales jusqu'à l'ordre deux (premier et second ordre), et différentes pour les ordres supérieures [Jul.62].

Cette conjecture a été nuancée par Pratt *et al.* [Pra.78], qui ont précisé que la discrimination était possible pour des textures contenant des structures déterministes particulières. Plus tard, Julesz a lui-même fourni des contre-exemples de sa conjecture.

A la recherche d'un ensemble de paramètres discriminants, Gagalowicz s'est intéressé aux textures microscopiques. Il les a modélisées par des processus stochastiques. Son objectif était de déterminer les paramètres de ces processus de telle sorte que les textures de synthèse générées, soient proches visuellement des textures originales [Gag.83]. Les résultats obtenus sont particulièrement intéressants car ils apportent une solution au problème de la mesure de la texture. Les conclusions de son étude sont les suivantes :

- le système visuel humain réalise une analyse locale de la texture,
- l'homme ne peut pas discriminer visuellement deux textures naturelles possédant les mêmes histogrammes locaux et les mêmes fonctions d'autocovariance,

- l'oeil ne prend pas en compte les dépendances spatiales entre points éloignés; le voisinage dans lequel il effectue une mesure d'organisation (respectivement de désorganisation) spatiale, correspond à quelques degrés d'angle solide sur la fovéa (9° environ).

Il est utile de préciser ici, l'importance du qualificatif "naturel" (par opposition à synthétique) utilisé dans le second point de la conclusion. Gagalowicz a en effet lui-même montré qu'il existait des textures de synthèse visuellement différentes bien que possédant mêmes autocovariance et histogramme.

Ainsi donc, l'ensemble des paramètres discriminants d'une texture naturelle, peut être limité aux statistiques locales du premier et du second ordre.

- ***Conclusion : nouvelle définition de la texture***

Les travaux de Julesz et de Gagalowicz indiquent d'une part que la perception de la texture repose sur une analyse locale, et d'autre part, que les statistiques du premier et du second ordre jouent un rôle important dans la discrimination des textures. Ces résultats permettent de lier la notion de texture à celle d'une information locale de voisinage. Ce qui conduit à une nouvelle définition de la texture : *Une texture est une région d'une image pour laquelle on peut définir une fenêtre de dimensions minimales, telle qu'une observation au travers de celle-ci se traduit par une perception (impression) visuelle identique pour toutes les translations possibles de cette fenêtre à l'intérieur de la région considérée* [Uns.84].

Cette définition intègre la propriété d'invariance en translation de la texture et met l'accent sur l'existence d'une fenêtre minimale à l'intérieur de laquelle une texture donnée peut être caractérisée. La définition de la taille de cette fenêtre pose, en d'autres termes, le problème de l'échelle de perception et d'analyse de la texture. Une solution possible et dont l'intérêt a été mis en évidence par plusieurs chercheurs [Ros.71, Mar.82], est d'effectuer l'analyse de l'image à différentes échelles. La représentation multi-échelles permet en quelque sorte d'observer la texture à différentes distances. Elle est en outre compatible avec le traitement visuel humain de l'information, puisque des expériences ont montré que l'image rétinienne est représentée dans le cortex à différentes échelles spatiales, rangées de la plus grossière (basse fréquence spatiale), à la plus fine (haute fréquence spatiale) [Wil.91].

D'autres expériences entreprises en physiologie de la vision ont permis d'établir l'existence de deux modes de vision humaine :

- le *mode de vision préattentive* qui opère en parallèle de façon quasi instantanée, et qui couvre un large champ visuel, et,
- le *mode de vision focale* qui est couplée avec une interprétation de type cognitive avec le cerveau. Ce mode de vision permet d'analyser la scène de façon systématique par un lent déplacement de la zone focale.

Ces deux modes de vision jouent un rôle essentiel dans la perception des textures. Ils ont été reliés à la définition de la notion de textons que Julesz a introduit en 1981 [Jul.81], suite à des expérimentations psychophysiques de perception de textures synthétiques macroscopiques.

2.3 Les textons, éléments de base de la perception de la texture

Julesz définit les *textons* comme étant les caractéristiques fondamentales de la texture. Selon lui, ces textons constituent les fonctions de base de la décomposition des textures par le système visuel humain. Ces fonctions sont reconnaissables par leur taille, leur forme allongée (ellipse, segment de ligne, coins, taches, etc...), leur placement ou leur orientation spatiale.

Julesz soutient que le traitement visuel de la texture fait intervenir la détection et le calcul de la densité des textons, et que la discrimination de deux textures s'opère par une comparaison des densités relatives des textons.

La détection des textons relève du mode de vision préattentive. Ce mode qui ne tient pas compte du couplage des textons, dirige le mode de vision attentive vers les zones où les gradients de textons apparaissent.

Plusieurs expérimentations portent à croire que les textons sont vraisemblablement considérés par le système visuel humain comme des caractéristiques importantes pour le traitement de l'information visuelle.

La théorie des textons de Julesz peut dans un certain sens, être rapprochée de la théorie du "bas-niveau" visuel de Marr (*the primal sketch*) dans laquelle l'image texturée est organisée symboliquement, à plusieurs échelles, en termes de primitives (barres, taches compactes, bords) [Mar.82].

La représentation de l'image dans le cortex visuel est d'ailleurs un sujet qui a particulièrement préoccupé de nombreuses équipes de recherche, depuis que des études ont suggéré que l'information visuelle subit un traitement comparable à un filtrage passe-bande lors de son acheminement vers le cerveau.

2.4 Représentation de l'image rétinienne dans le cortex visuel

2.4.1 Décomposition multicanaux de l'information visuelle

Les travaux pionniers de Campbell et Robson [Cam.68] portant sur l'analyse de la voie visuelle, ont permis d'établir que *le système visuel humain réalisait* dans sa première phase de traitement, *une analyse linéaire grâce à des canaux sensibles chacun à une gamme limitée de fréquences spatiales*.

Les expérimentations psychophysiques qui ont permis de suggérer une telle décomposition sont décrites dans la littérature [Bus.87]. Elles utilisent des réseaux spatiaux (*figure 2.3*) comme stimuli, et sont basées sur des techniques d'adaptation.

D'autres recherches ont permis de mettre en évidence la sélectivité d'orientation des canaux spatiaux révélés par Campbell et Robson. Des études physiologiques effectuées sur des chats et des singes qui ont un cortex visuel similaire à celui de l'homme, ont montré que les cellules simples de ces animaux, répondent à des stimuli orientés et périodiques [Hub.62, Val.82]. Il a été également établi que la sensibilité visuelle de l'homme dépend fortement de l'orientation spatiale. Cette sensibilité est maximale pour des raies horizontales et verticales. Elle décroît lorsque les raies sont obliques, pour atteindre le minimum pour une orientation de 45°.

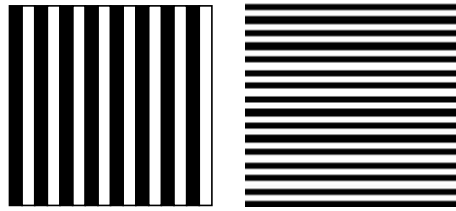


Figure 2.3 : Exemples de réseaux spatiaux

Pendant longtemps, on s'est interrogé sur le modèle mathématique qui pourrait correctement décrire la décomposition multicanaux effectuée par le système visuel.

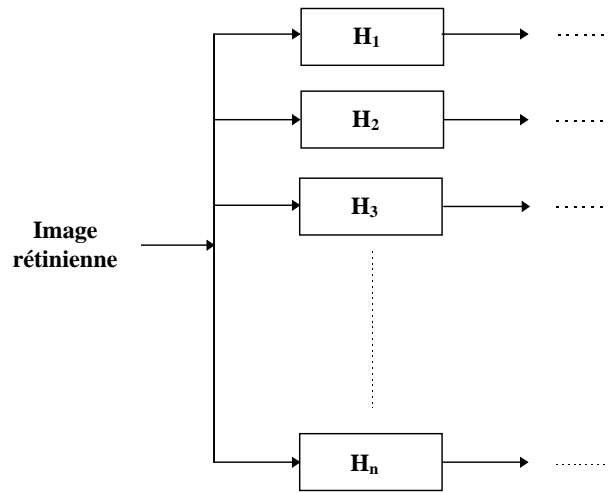
L'analyse de Fourier qui avait été initialement proposée, a suscité de nombreuses réserves. Cela parce que cette proposition entre directement en conflit avec la théorie qui précise que le traitement de l'information rétinienne réalisé par le cortex visuel, est basé sur une détection de caractéristiques (fronts, barres) de différentes tailles et de différentes orientations [Hub.62]. D'autre part, l'analyse de Fourier n'a aucune signification physiologique. La vision humaine en effet, ne prend en compte qu'une portion d'espace finie relativement limitée.

Il s'avèrera donc nécessaire d'opter pour un autre modèle. Le développement récent de nouveaux outils mathématiques en a inspiré plusieurs.

2.4.2 Modélisation de la décomposition multicanaux de l'information visuelle

La décomposition multicanaux de l'information visuelle est l'oeuvre des cellules simples du cortex visuel. Chacune de ces cellules peut être caractérisée par sa réponse à des spots lumineux (réponse impulsionnelle). Le support de cette réponse est le champ récepteur de la cellule.

Certains chercheurs ont mesuré les réponses impulsionnelles de cellules corticales simples. Ces réponses dépendent de l'orientation spatiale des stimuli. Maffei et Fiorentini [MaF.76] ont montré qu'elles sont raisonnablement linéaires et qu'elles peuvent être modélisées par des filtres linéaires.



H_i : convolution avec un filtre passe-bande orienté H_i

Figure 2.4 : Modèle multicanaux du traitement visuel humain

Ce résultat suggère que la décomposition de l'image par le système visuel humain en un ensemble de canaux parallèles, indépendants, et d'orientations particulières, peut être modélisé par des bancs de filtres linéaires passe-bande ayant chacun une sélectivité d'orientation spatiale particulière (figure 2.4).

En 1980, Marcelja a montré que le profil des champs récepteurs des cellules simples du cortex visuel des mammifères, se rapprochait des fonctions de Gabor à une variable (1-D) [Mar.80]. Daugman [Dau.80] qui a par la suite étendu le modèle au cas bidimensionnel (2-D), a montré que ces fonctions minimisent conjointement les incertitudes sur les localisations dans les domaines spatial et fréquentiel [Dau.85].

Les fonctions de Gabor sont le produit d'une gaussienne par une fonction sinus. Les fonctions utilisées en vision (2D), sont définies par :

$$G(x, y) = \exp \left[- \left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2} \right) \right] \times \exp(j2\pi(u_0x + v_0y))$$

où (x_0, y_0) représente le point de fixation de l'observateur dans l'espace visuel bidimensionnel (2D), \dagger_x et \dagger_y l'échelle de la fonction de Gabor suivant les axes x et y respectivement, et u_0 et v_0 les fréquences de modulation.

Le choix des fonctions de Gabor comme modèle, a été conforté par les résultats d'autres travaux et principalement par ceux de Pollen et Roner [Pol.83]. Ceux-ci ont prouvé l'existence dans le cortex strié, de cellules contiguës, calées sur la même orientation et à la même fréquence spatiale, mais en quadrature de phase. Ils ont montré que ces cellules simples du cortex visuel, symétriquement paires et impaires pouvaient être modélisées par un couple de filtres de Gabor symétriques.

Des objections ont cependant été émises contre ce modèle pour deux raisons principales. La première découle de récents travaux qui ont montré que contrairement aux fonctions de Gabor, les cellules simples du cortex visuel humain ne minimisaient pas conjointement les localisations spatiales et fréquentielles [Sto.90]. La seconde et la plus significative, est que les canaux des cellules simples ont une bande passante constante en échelle logarithmique [Web.85]. Ce qui n'est pas le cas des bandes fréquentielles des filtres de Gabor, du fait de la taille constante des fenêtres gaussiennes intervenant dans la définition des fonctions.

D'autres modèles qui possèdent, à un degré ou à un autre, un lien avec les fonctions de Gabor, ont été proposés comme alternatives. Citons par exemple, les dérivées d'une gaussienne (chapeau mexicain) par Marr [Mar.82], les différences d'une gaussienne par Malik et Perona [Mal.90], et les ondelettes qui ont été dans le principe, initialement proposées par Mallat [Mal.89b].

Les ondelettes qui sont utilisées pour analyser des signaux présentant des gammes de variations très différentes les unes des autres, semblent correspondre davantage aux réalités physiologiques du traitement visuel humain. Gaudart *et al* [Gau.96] ont récemment montré qu'il était possible à partir d'une ondelette mère, de construire chacun des champs récepteurs tels que les neurophysiologistes les ont décrits. Il apparaît de ce fait que les ondelettes constituent un "bon" modèle pour la vision humaine.

Il semble peu probable cependant, qu'un modèle mathématique soit approprié pour décrire parfaitement le traitement biologique complexe réalisé dans le cortex visuel humain. Ce, d'autant plus que la description de ce traitement reste encore incomplète. Par exemple, on ne sait ni le nombre de canaux impliqués dans cette décomposition et ceux d'entre eux qui y jouent un rôle prépondérant, ni comment le cerveau humain combine la sortie des différents filtres pour effectuer des jugements visuels tels que la reconnaissance ou la discrimination. En outre, le traitement de l'information par des filtres agissant de façon indépendante, semble être remis en cause par plusieurs expérimentations qui ont démontré de significatives interactions entre les différents mécanismes [Olz.92]. Toutefois, ces différents modèles proposés ont l'avantage d'apporter un formalisme algorithmique crédible pour l'élaboration de méthodes d'analyse d'images.

2.4.3 Intérêt du modèle multicanaux pour l'analyse de la texture

Les différents travaux qui ont permis d'établir que le système visuel réalisait une décomposition multicanaux, ont contribué à approfondir les connaissances sur les mécanismes de perception humaine de la texture.

En effet, chacune des images filtrées résultant du traitement visuel primaire, contient des variations d'intensité sur une bande donnée de fréquence et d'orientation particulières. Ces variations mettent à coup sûr en évidence, certaines caractéristiques de l'image originale à savoir, régularité, granularité, et directionnalité qui sont indispensables au système visuel pour entres autres, discriminer et reconnaître différentes textures.

Cette décomposition de l'image rétinienne inspire et motive le choix d'un modèle de filtrage multicanaux pour l'analyse de textures. L'idée sous-jacente à une telle approche, est d'analyser l'image de textures à partir de la caractérisation de la distribution spatiale de l'énergie des différents canaux filtrés.

Comme l'indique Bovik *et al.* [Bov.90], le modèle multicanaux semble être un cadre unificateur pour les deux théories, antagonistes à première vue, qui ont été initialement émises pour expliquer le traitement visuel humain. *La détection de caractéristiques* trouve son compte en effet, dans les propriétés de localisation spatiale des différents canaux qui

permettent d'extraire des caractéristiques locales particulières des éléments de texture, tandis que l'étroitesse des canaux fréquentiels garantit la prise en compte de l'analyse globale que présuppose *la décomposition de Fourier*.

2.5 Conclusion

Les avancées effectuées dans la compréhension des mécanismes de fonctionnement du système visuel humain, ont levé le voile sur certains aspects de la perception humaine de la texture. Celle-ci est réalisée localement et est fortement liée à l'orientation, à la taille des éléments de texture, ainsi qu'à la distance qui les sépare. Elle repose également sur une décomposition multicanaux de l'information visuelle qui permet de mettre en évidence les caractéristiques particulières de la texture.

Les différents résultats obtenus, ont induit et/ou légitimé le développement de diverses approches et méthodes d'analyse de textures. Les travaux de Julesz et de Gagalowicz ont permis de faire le point sur la discrimination et la perception de textures naturelles. Ces travaux sont également à la base du développement des méthodes probabilistes. La théorie des textures apporte une justification à la stratégie d'extraction d'attributs schématisée à la figure 1.2. Enfin, le modèle de filtrage multicanaux est motivé par les récentes découvertes sur le fonctionnement du système visuel humain.

Dans cette thèse, nous adopterons comme méthode d'analyse de textures, l'analyse par ondelettes qui est un modèle de filtrage multicanaux. Le principal objectif n'est pas d'utiliser le modèle le plus probable du traitement biologique cortical, mais plutôt de recourir à un outil pratique qui possède des propriétés intéressantes d'analyse, et qui plus est, trouve la justification de sa mise en oeuvre dans le traitement effectué par le système visuel.

Avant d'aborder la théorie des ondelettes, nous présentons, au chapitre suivant, les principales méthodes d'analyse de textures qui ont été proposées. Ces méthodes ont été logiquement influencées par la progression de la recherche en physiologie du système visuel humain. Ainsi, si les premières méthodes d'analyse de textures mises en oeuvre se sont trouvées relativement éloignées d'un modèle fiable de la vision humaine, celles qui ont été introduites récemment,

tendent d'intégrer les connaissances acquises sur les mécanismes de la vision humaine même si ceux-ci sont encore loin d'être maîtrisés.

3

METHODES CLASSIQUES D'ANALYSE DE TEXTURES

Ce chapitre est consacré à la présentation de quelques méthodes d'analyse de texture. Ce terme regroupe l'ensemble des techniques d'extraction d'attributs globaux, utilisables pour caractériser une texture.

Les nombreuses techniques qui ont déjà été développées, utilisent des approches, des modèles et des descripteurs aussi différents les uns que les autres. Certaines de ces techniques sont fondées sur des formulations mathématiques solides. D'autres sont purement heuristiques, mais relèvent néanmoins d'une recherche intuitive de caractéristiques pertinentes, pouvant permettre d'analyser efficacement les images texturées [He.92]. Quelques unes parmi elles opèrent à un niveau symbolique en organisant l'image texturée en termes de primitives. Les autres caractérisent directement la texture à partir des niveaux de gris de l'image.

Ces différentes méthodes d'analyse que l'absence d'une définition précise de la texture justifie, font l'objet d'une littérature abondante. Les articles de synthèse de Haralick [Har.79] et de Van Gool *et al.* [Goo.85] en donnent un bon aperçu.

Ces méthodes peuvent être classées selon l'approche adoptée ou selon le type de textures à caractériser. Sans recourir à une classification précise de ces différentes techniques,

nous présentons dans ce qui suit, celles qui nous semblent les plus représentatives de l'état de l'art actuel.

3.1 Les méthodes structurelles

3.1.1 Principe

La régularité qui caractérise certaines textures, implique l'adoption de méthodes structurelles pour leur analyse et leur synthèse. Le principe de ces méthodes est de caractériser la texture par des primitives élémentaires et par leur arrangement spatial selon des règles de placement qui peuvent être déterministes pour des textures périodiques et très bien structurées, ou aléatoires pour des textures plus chaotiques.

3.1.2 Extraction des primitives

Les primitives d'une même texture peuvent être de tailles différentes. Ce qui rend leur extraction complexe. Le processus d'extraction peut être conduit de deux façons différentes.

La première possibilité est de regrouper des pixels voisins et similaires pour obtenir la primitive de base. Elle est illustrée par la méthode d'extraction de primitives mise en oeuvre par Matsuyama *et al.* [Mat.82]. Cette méthode qui se fonde sur l'hypothèse qu'une primitive est constituée de pixels adjacents ayant des niveaux de gris similaires, utilise une technique de croissance de région pour extraire les différentes primitives de l'image.

En adoptant la même approche, certains auteurs ont proposé d'autres techniques d'extraction de primitives. Citons les travaux de Connors et d'Harlow [Con.80a] qui utilisent certains descripteurs (comme le contraste) de la méthode statistique des matrices de cooccurrences que nous développerons plus loin, pour déterminer la taille et la forme des primitives.

La seconde possibilité consiste à déterminer en premier lieu la structure de la texture avant de procéder à l'extraction de la primitive.

Matsuyama *et al.* [Mat.83] utilisent pour cela, la Transformée de Fourier pour trouver la structure de la texture en recherchant les deux vecteurs du plan fréquentiel qui représentent la périodicité de la texture. Cette donnée leur permet de déterminer la maille dans laquelle s'insère l'élément de base de la texture.

Vilnrotter *et al.* [Vil.86] quant à eux, partent de l'idée que les informations essentielles de la structure d'une texture sont contenues dans la description des contours de ses différents éléments. Ils utilisent donc une analyse statistique des caractéristiques de contours pour déterminer la structure de la texture, et pour contrôler ensuite le processus d'extraction de ses primitives.

Enfin, Kim et Park [Kim.92] déterminent la structure de la texture à partir de deux directions privilégiées qu'ils obtiennent en effectuant une projection de l'image texturée sur plusieurs axes d'orientations différentes. Les primitives sont ensuite capturées de façon beaucoup plus précise dans une grille ayant la forme d'un parallélogramme.

La première approche qualifiée d'ascendante, est simple dans son principe et dans sa mise en oeuvre. Cependant elle fournit des résultats peu satisfaisants, particulièrement lorsque la texture est fine ou dégradée par un bruit aléatoire, ou lorsque la variabilité de la structure de la texture est importante. La seconde approche (l'approche descendante) s'est révélée plus robuste que la première. Elle est néanmoins plus complexe à mettre en oeuvre et plus coûteuse en temps de calcul.

On peut envisager dans une certaine mesure, de caractériser les primitives au lieu de les extraire. Cela est réalisable en appliquant la morphologie mathématique à l'analyse de texture. La granulométrie morphologique permet en effet, d'utiliser un groupe d'éléments structurants de différentes tailles pour décomposer l'image texturée en une série d'images représentatives de primitives de textures, suivant leur forme et leur taille [Coc.95].

3.1.3 Détermination des règles de placement des primitives

La stratégie de détermination des règles de placement des primitives est généralement liée à la technique utilisée pour l'extraction de la primitive.

Matsuyama *et al.* [Mat.83] qui ont recours à une analyse de Fourier pour extraire les primitives, déterminent leurs règles de placement par une analyse dans le plan de phase de Fourier. Certains auteurs comme Vilnrotter *et al.* [Vil.86] et Matsuyama *et al.* [Mat.82], déterminent ces règles en calculant les vecteurs différence ou distance entre les différents motifs de la texture.

Cette détermination des règles de placement est par contre réalisée de façon implicite, lorsque l'on réussit à capturer la structure de base de la texture dans un parallélogramme qui épouse sa forme [Kim.92]. Dans ce cas, la périodicité élémentaire de la texture est décrite par la forme du parallélogramme, et les vecteurs directeurs de ses côtés, permettent d'établir les règles de placement des primitives de la texture.

D'une façon générale, les méthodes structurelles sont complexes à mettre en oeuvre. Leurs performances sont en outre fortement conditionnées par la variabilité spatiale de la structure de la texture. Ce qui restreint leur champ d'application à l'analyse de textures périodiques et faiblement bruitées qui, dans la réalité, ne se rencontrent que très peu.

3.2 Les méthodes statistiques

Les textures qui présentent un aspect aléatoire, suggèrent le développement de méthodes statistiques pour leur analyse. Ces dernières peuvent être réparties en deux groupes : les méthodes de dépendance spatiale et les méthodes fondées sur des modèles tels que les champs de Markov, les fractals, les modèles autorégressifs, etc....

3.2.1 Les méthodes de dépendance spatiale

Ces méthodes s'inscrivent dans le contexte général de l'analyse de texture schématisée à la *figure 1.2*, et permettent d'évaluer certaines propriétés de la texture telles que la granularité et

la directivité. Elles procèdent par une analyse sélective de l'image suivant une direction d'observation donnée, et reposent toutes sur l'évaluation de mesures statistiques invariantes par translation.

- **Les matrices de cooccurrence**

La méthode des "matrices de cooccurrence" a été introduite par Haralick [Har.79]. Elle se réfère aux statistiques d'ordre deux et trouve une justification dans la conjecture de Julesz. Cette méthode construit un ensemble de matrices de cooccurrence de niveau de gris à partir desquelles seront extraites plusieurs valeurs de caractéristiques.

Soient deux pixels $I(x_1, y_1)=i$ et $I(x_2, y_2)=j$ d'une image $I(x, y)$ qui possède N niveaux de gris. La construction des matrices de cooccurrence de taille $N \times N$ est basée sur l'estimation des fonctions de densité de probabilité conditionnelle conjointes du second ordre $F(i, j/d, \theta)$ de la paire de pixels (i, j) séparés par une distance d suivant une orientation relative θ ($0^\circ, 45^\circ, 90^\circ$, et 135°).

L'exemple de la *figure 3.1* donne les matrices $F(i, j/1, 0^\circ)$ et $F(i, j/2, 135^\circ)$ obtenues à partir de l'image I pour deux distances (1 et 2) et deux orientations (0° et 135°) différentes.

Les matrices de cooccurrence conduisent à une quantité d'informations beaucoup plus volumineuse que le nombre de pixels de l'image. Pour réduire la masse d'informations, Haralick a proposé 14 indices ou attributs qui correspondent à des caractères descriptifs de la texture.

Ces descripteurs sont fortement corrélés entre eux et ne sont jamais utilisés ensemble. Les attributs les plus fréquemment utilisés sont : les mesures de contraste, d'homogénéité, d'entropie, de corrélation, et de moment d'inertie. L'expression de la mesure de contraste est donnée ci-dessous à titre d'exemple.

$$\text{Contraste : } Cont = \sum_{p=1}^N p^2 \sum_{i=1}^N \sum_{\substack{j=1 \\ |i-j|=p}}^N \frac{F(i, j)}{R}$$

où R est une constante de normalisation, et p sert à pondérer chaque terme de la matrice de cooccurrence par sa distance à la diagonale. On peut constater que la mesure de contraste est élevée pour les termes éloignés de la diagonale de la matrice.

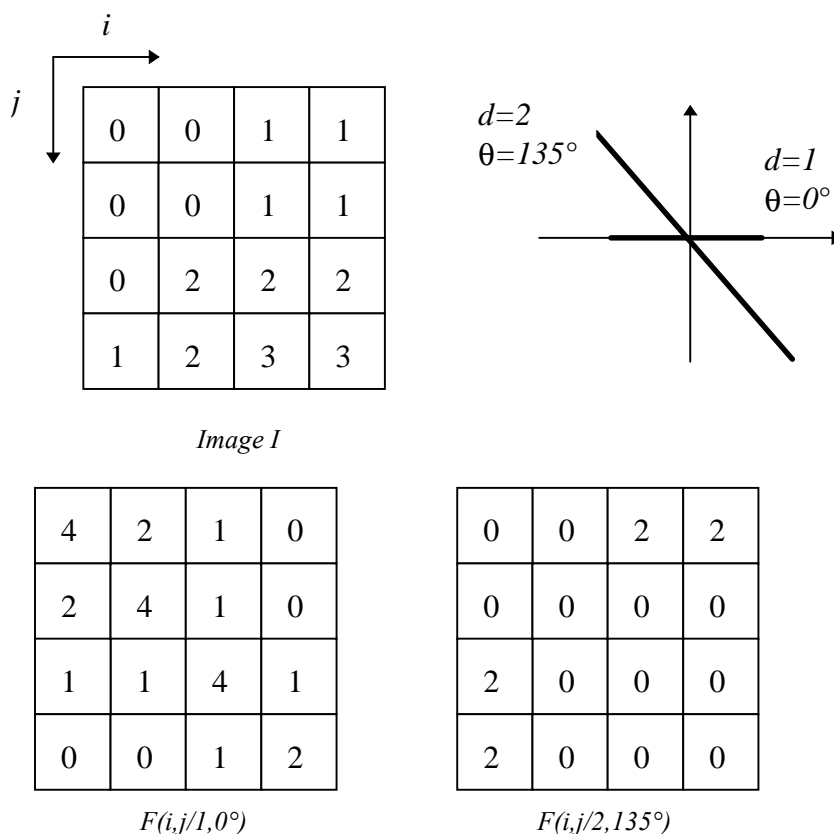


Figure 3.1 : Une image et deux de ses matrices de cooccurrences spatiales

La méthode d'Haralick qui est l'une des plus complètes en terme de quantité d'informations, a été utilisée dans des applications variées. Des études comparatives [Con.80b] ont prouvé sa supériorité sur plusieurs méthodes. Cependant, ses utilisateurs lui reprochent son temps de calcul excessif, la grande place mémoire nécessaire à sa mise en oeuvre, et sa grande sensibilité à l'échelle d'analyse.

- **Autres méthodes de dépendance spatiale**

Dans une optique semblable à celle d'Haralick d'autres méthodes de dépendance spatiale ont été proposées.

Citons par exemple, celles qui utilisent des statistiques d'ordre plus élevé pour caractériser les textures [Tsa.92, Reu.95], et celle mise en oeuvre par Galloway [Gal.75] qui totalise dans plusieurs directions θ , les fréquences d'occurrences $F(I, L, \theta)$ de segments de pixels de longueur L et de niveau de gris I constant. Les matrices d'isosegments qu'il obtient, lui permettent de calculer un ensemble d'attributs de textures qui soulignent principalement la présence de lignes courtes ou de lignes longues, et l'uniformité des longueurs ou des niveaux de gris.

Il est également intéressant d'indiquer qu'en adoptant une toute autre approche, Unser [Uns.86a] a décrit une alternative à la méthode d'Haralick. Sa méthode s'appuie sur la diagonalisation de la matrice de covariance d'une paire de pixels et donne lieu à une caractérisation des propriétés de texture avec un nombre moindre d'attributs.

Dans son étude, Unser montre que dans le cas gaussien, les sommes et les différences de deux pixels représentent les axes principaux de la densité de probabilité conjointe d'ordre deux. Sa méthode qui nécessite beaucoup moins de place mémoire, est tout autant efficace que celle d'Haralick. Elle permet en effet d'obtenir un grand nombre d'attributs calculés à partir des matrices de cooccurrences. Par exemple, le contraste s'obtient à partir de l'histogramme des différences.

Les attributs proposés par les méthodes de dépendance spatiale sont dans leur grande majorité plus ou moins heuristiques et n'ont en général que très peu de corrélation avec les caractéristiques utilisés par le système visuel humain pour discriminer les textures. D'autres méthodes statistiques basées sur des modèles mathématiques ont été récemment introduites avec succès en analyse d'images et notamment en analyse de textures. L'intérêt qu'elles ont suscité, s'explique par le fait que les images de synthèse réalisées à l'aide des modèles définis, ont un aspect visuel très proche de textures naturelles.

3.2.2 Les méthodes d'analyse de texture fondées sur des modèles

Le principe de ces méthodes est de représenter la texture par un processus aléatoire bidimensionnel correspondant à un modèle choisi, puis d'estimer les paramètres de ce modèle

afin de les utiliser pour caractériser la texture. Les modèles utilisés sont généralement des champs aléatoires ou des fractals.

- **Modélisation de la texture par des champs aléatoires**

Les modèles probabilistes ont connu un rapide développement depuis leur introduction récente en analyse d'images.

De nombreux modèles : champs de Markov, modèle autorégressif [Mao.92], ..., ont été proposés pour décrire la texture. Parmi ces modèles, les champs de Markov ont suscité beaucoup d'intérêt et ont fait l'objet de nombreux travaux.

La diversité des modélisations qu'ils offrent (champ de Markov, champ markovien gaussien, champ de Markov caché,...), leur capacité de généralisation et leur robustesse, ont contribué à leur succès aussi bien en analyse [Cro.83] qu'en synthèse [Seb.89], ou en classification [Che.85] et en segmentation de textures [Man.91, Ker.94].

L'analyse d'une texture décrite par le modèle markovien repose sur le choix d'une structure spatiale (une topologie) qui va représenter le plus fidèlement possible la texture étudiée.

L'approche classique consiste à imposer un type de modèle de voisinage local (un site), sur lequel chaque pixel est étroitement lié à ses voisins, puis, à estimer les paramètres de ce modèle.

Le critère d'estimation retenu est souvent celui du Maximum A Posteriori (MAP). Etant donné une réalisation y du champ d'observation Y (la texture), on cherche la réalisation \hat{x} , du champ de primitives X , la plus probable au sens de la distribution a posteriori. L'application de la stratégie bayésienne permet d'obtenir la formulation simple suivante :

$$\hat{x} = \arg \max_x (P(Y = y / X = x) P(X = x))$$

Grâce à l'interprétation en termes de distribution de Gibbs, l'estimation bayésienne conduit à la minimisation globale d'une fonction d'énergie.

Rechercher le minimum global de la fonction d'énergie revient à résoudre un problème d'optimisation globale. Pour cela, on peut utiliser des algorithmes de relaxation stochastiques ("recuit stochastique simulé") ou déterministes (Iterated Conditional Mode (ICM)) [Kir.83, Coc.95]. Les premiers fournissent de bons résultats mais nécessitent le plus souvent un espace de configuration extrêmement vaste qui donne lieu malheureusement à des optimisations longues et coûteuses. Les seconds bien que rapidement convergents, se laissent souvent piéger dans d'éventuels minima locaux. Ces algorithmes d'optimisation constituent l'un des points faibles de la modélisation markovienne.

Une seconde faiblesse apparaît dans la géométrie fixe du motif textural élémentaire qui présuppose la stationnarité de la texture, et ne permet d'appréhender qu'un seul facteur d'échelle.

Pour pallier ces faiblesses, des méthodes fondées sur des coopérations entre modèles markoviens et méthodes d'analyse hiérarchique (multigrilles, multiéchelles, multirésolution), ont été développées [Bou.91, Per.92]. Les modèles hiérarchiques markoviens répondent ainsi au besoin d'une hiérarchisation du traitement statistique de l'information, et conduisent à une amélioration sensible des performances globales des algorithmes d'optimisation. Les résultats satisfaisants qu'ils fournissent, expliquent leur utilisation croissante dans des applications variées.

- **Modélisation de la texture par les fractals**

La géométrie fractale permet l'étude d'ensembles que la géométrie euclidienne ne peut appréhender. Un objet fractal est défini comme un ensemble dont la dimension fractale est fractionnaire et strictement supérieure à sa dimension topologique.

Les fractals ont présenté un intérêt pour l'analyse de texture depuis que Pentland [Pen.84] a trouvé une corrélation importante entre la dimension fractale, et l'estimation visuelle de la rugosité d'une surface.

L'un des modèles fractals utilisés en analyse de texture est le mouvement Brownien fractionnaire fBm (fractional Brownian motion). Selon ce modèle, la texture est considérée comme un signal aléatoire $T(x)$, avec x appartenant à un espace vectoriel de dimension topologique D_T .

La dimension fractale D_F de la texture est donnée par :

$$D_F = D_T + 1 - H$$

où H est le paramètre principal du fBm ($0 < H < 1$)

D_F rend compte de la rugosité apparente de la surface : plus H est élevé, plus la texture paraît lisse.

La dimension fractale qui est facile à calculer, peut donc être utilisée comme paramètre pour la caractérisation des textures, comme ce fut le cas dans [Pel.84, Pen.84]. Seulement, elle est insuffisante pour caractériser à elle seule, un grand nombre de textures dont par exemple celles qui ont des motifs identiques mais différemment orientés. Cela s'explique par le fait que l'analyse fractale est invariante aux transformations d'échelles, de translations et de rotations. Ce qui n'est pas le cas de la texture, qui n'est en définitive fractale que sur une gamme limitée d'échelle [Pel.84].

On adjoint donc à la dimension fractale, des dimensions d'ordre élevé ou des mesures, comme la corrélation ou la lacunarité qui représente la variance de la mesure de la dimension fractale, ou encore la signature fractale qui est la variation locale de la mesure de la dimension fractale [Pel.84]. Une autre proposition qui a été faite, est de mesurer le paramètre fractal H suivant plusieurs directions de déplacement, et sur un certain nombre de niveaux de la pyramide [Kpa.91]. Cette approche n'est envisageable que pour des textures naturelles pour lesquelles, l'hypothèse d'isotropie des modèles fractals n'est pas vérifiée.

Les recherches actuelles s'orientent plutôt vers une analyse multifractale [Ber.91, Lev.92, Lev.95]. Les auteurs qui la développent remettent en cause le bien fondé de

l'application de la dimension fractale en analyse d'images. En effet, cette application qui assimile l'intensité lumineuse à une dimension géométrique, considère que l'image bidimensionnelle (2D) de niveaux de gris, est équivalente à une surface tridimensionnelle (3D).

L'approche multifractale permet de corriger cette erreur en considérant l'intensité lumineuse associée à chaque pixel comme une mesure fonction d'un paramètre α qui peut servir à caractériser les comportements "singuliers" des niveaux de gris. Ainsi, les informations sur le comportement de la mesure (fournies par le spectre des valeurs α), et celles sur la répartition géométrique des valeurs de ce comportement, peuvent être avantageusement mises à profit pour caractériser la texture.

Les attributs fractals ne peuvent se calculer que sur des fenêtres ou des régions de taille importante. Leur temps de calcul est généralement élevé. Par contre, ils permettent souvent de bien caractériser la texture.

3.2.3 Conclusion sur les méthodes statistiques

L'application des méthodes statistiques en analyse de texture a connu un accroissement remarquable ces dernières années. Ces méthodes bénéficient des retombées d'une coopération avec d'autres méthodes (hiérarchiques, mais aussi multicanaux [Che.95]), qui améliorent nettement leurs performances et leur permettent d'aborder avec de nouveaux outils des problèmes classiques en traitement d'images.

Le champ d'application de ces méthodes qui était restreint à des textures aléatoires, s'est étendu aux textures naturelles qui présentent un aspect à la fois structuré et aléatoire. Pour ce type de textures, les méthodes d'analyse par filtrage spatial sont plus indiquées. Ces dernières permettent en effet, de considérer les distributions spatiales de niveaux de gris sur des voisinages de tailles plus importantes que celles des voisinages mis en oeuvre dans les méthodes statistiques classiques.

3.3 Les méthodes d'analyse par filtrage spatial

Le filtrage spatial peut être envisagé avec deux objectifs différents.

On peut chercher à mettre en évidence les structures particulières d'une image texturée en la décomposant par un banc de filtres en une série d'images filtrées. Les informations contenues dans ces images peuvent être alors directement exploitées pour l'analyse de l'image. Les méthodes de filtrage par transformations linéaires locales utilisent une telle approche.

On peut également envisager d'utiliser un ensemble de filtres multicanaux localisés, avec comme objectif principal de rechercher des propriétés caractéristiques de l'image à partir de l'analyse des réponses fréquentielles des canaux mis en oeuvre. C'est le but que vise les méthode de filtrage spatial multicanaux.

3.3.1 Les transformations linéaires locales

Le principe de ces techniques est de caractériser les propriétés locales de la texture par un ensemble d'attributs calculés après avoir convolué l'image originale avec des masques (filtres) de taille réduite.

Laws a été le premier à proposer une telle technique de caractérisation de textures [Law.79]. L'image est d'abord convoluée avec un ensemble de p filtres à réponse impulsionnelle finie (micromasques bidimensionnels de taille 3×3 ou 5×5), puis des attributs appelées "énergies de texture" sont calculées sur un macromasque de taille plus importante (15×15 à 31×31).

Un ensemble de micromasques a été proposé de façon ad-hoc par Laws. Ceux-ci permettent de détecter des structures particulières de la texture (lignes, forme en V,...). Malgré leur nature empirique, ces masques sont simples et faciles à mettre en oeuvre. La *figure 3.2* présente les quatre micromasques les plus utilisés [Law.79].

$$\begin{aligned}
\text{E5L5} &= \begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 8 & 12 & 8 & 2 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} & \text{L5S5} &= \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix} \\
&\text{détecteur de ligne horizontale} & & \text{détecteur de ligne verticale} \\
\\
\text{R5R5} &= \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ -4 & 16 & -24 & 16 & -4 \\ 6 & -24 & 36 & -24 & 6 \\ -4 & 16 & -24 & 16 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix} & \text{E5S5} &= \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -2 & 0 & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & -4 & 0 & 2 \\ 1 & 0 & -2 & 0 & 1 \end{bmatrix} \\
&\text{détecteur de point de haute fréquence} & & \text{détecteur de forme en V}
\end{aligned}$$

Figure 3.2 : Masques bidimensionnels de Laws les plus utilisés

Le vecteur d'attributs associé à chaque pixel est formé de p énergies de textures $f_n(i,j)$, (pour $n=1, \dots, p$) définies comme suit :

Soit W le macromasque de taille $(2m+1) \times (2m+1)$, centré autour d'un pixel $x_n(i,j)$ de l'image obtenue à la sortie du filtrage par le n ième micromasque ($n=1, \dots, p$). L'énergie de texture $f_n(i,j)$ associée à ce pixel, est donnée dans sa formulation simplifiée par :

$$f_n(i, j) = \frac{1}{(2m+1)^2} \sum_W |x_n(k, l) - m(i, j)|$$

$$\text{avec } m(i, j) = \frac{1}{(2m+1)^2} \sum_W x_n(k, l)$$

La méthode de Laws qui a connu plusieurs déclinaisons [Pie.83, Wan.86], s'intègre dans le cadre théorique que Unser a défini en établissant un lien entre l'application des bancs de filtres et une transformation linéaire des valeurs de niveaux de gris de pixels voisins [Uns.86b].

Cette transformation linéaire locale peut être reliée à l'étude statistique classique. Elle permet en effet d'accéder à des statistiques d'ordre élevé par l'intermédiaire de statistiques d'ordre 1. Cela, parce que chaque filtre mis en oeuvre par cette transformation, combine en une valeur unique, les niveaux de gris d'un groupe de pixels adjacents.

Dans le cas d'une image composée de deux textures, la transformation linéaire locale donne des solutions optimales lorsqu'elle est déterminée de façon à diagonaliser la matrice de covariance de ces textures. Cette propriété est obtenue avec la Transformation de Karhunen-Loève (TKL). Cette transformation assure la décorrélation des images filtrées, et concentre l'"énergie de textures" sur les images filtrées associées aux plus grandes valeurs propres.

La capacité de la TKL, à extraire les différents constituants d'une texture donnée, a été démontrée par Ade [Ade.83]. Poursuivant son idée, Provent [Pro.91] a montré que le nombre d'images filtrées (ou de canaux) possédant une énergie de textures significative, dépendait de la texture. Comme ces énergies se répartissent différemment dans les canaux selon la texture, il est possible de caractériser les textures de façon pertinente, et de réaliser une segmentation efficace des régions texturées d'une image.

L'absence d'algorithmes rapides de calcul de la TKL, lui fait préférer d'autres transformations dites sous-optimales, telles que la DCT qui, dans de nombreux cas, donne des résultats très comparables.

3.3.2 Le filtrage spatial multicanaux

Le filtrage multicanaux permet de décomposer l'image originale par des filtres (ou des canaux) accordés chacun sur une bande étroite de fréquence et d'orientation particulières. Cette décomposition fournit différentes images filtrées. Ces images contiennent diverses caractéristiques de l'image originale qui peuvent être utilisées pour l'analyse de textures.

Le filtrage multicanaux a été initialement mis en oeuvre à l'aide de la classique transformation de Fourier [Baj.73]. Les résultats mitigés obtenus alors, ont été nettement améliorés avec l'implantation de nouvelles transformations.

La Transformation de Wigner-Ville (TWV) qui s'est révélée être un outil intéressant pour l'analyse des signaux 1D, a été utilisée en analyse d'images sous sa version lissée en raison de la présence de termes d'inférences qu'elle induit entre les différentes composantes issues du filtrage [Fla.84]. Cette version dénommée Pseudo Distribution de Wigner, a été

appliquée à l'analyse de textures par Reed et Wechsler [Ree.90] qui l'ont utilisé dans un problème de segmentation d'images.

La Transformation de Gabor qui peut être interprétée comme une version lissée de la TWV [Jac.88], a été plus largement utilisée du fait qu'elle a été proposée comme modèle de description du traitement visuel humain.

Turner [Tur.86] a établi l'utilité des filtres de Gabor en analyse de textures en montrant que le spectre de puissance des canaux renfermait des informations importantes pour la discrimination des textures.

Ces filtres ont été diversement appliqués à la segmentation de textures. Les techniques existantes utilisent un banc de filtres soit à valeur complexes [Bov.90], soit à valeurs réelles et symétriques [Jai.91], ou constituant des paires de filtres conjugués en quadrature [Tan.90]. Les processus de segmentation de ces différentes techniques se basent toutes sur une information qui présente une unité de nature fréquentielle.

Récemment, certains auteurs ont utilisé les informations de phase [Buf.90] (ou de la dérivée de la phase [Yum.94]), pour segmenter des textures. Les résultats obtenus sont tout à fait prometteurs.

Indiquons, à la charge des méthodes d'analyse de textures fondées sur la Transformation de Gabor, qu'elles sont difficiles à mettre en oeuvre. En effet, la panoplie de canaux (d'orientations, de bandes passantes et de fréquences centrales diverses) qu'il est possible d'obtenir avec cette transformation, est impressionnante. Ces nombreux canaux qui ne renferment pas tous des informations utiles, constituent une véritable charge de calcul. La difficulté réside donc dans le choix du nombre de canaux utiles et de leurs caractéristiques.

Pour résoudre ce problème, Bovic et *al.* [Bov.90] ont appliqué un seuillage sur le spectre de puissance globale de l'image pour détecter les pics de fréquence. Ces fréquences sont par la suite utilisées comme les fréquences centrales des canaux utiles, et leur nombre fixe celui des canaux utiles. Cette idée de base a connu diverses améliorations par la suite [Tan.95, Dun.95]. Toutefois, aucune solution satisfaisante n'a encore été trouvée. Les

recherches actuelles tendent vers la résolution de ce problème et vers le développement d'algorithmes rapides pour la mise en oeuvre du filtrage de Gabor.

3.3.3 Conclusion sur le filtrage spatial multicanaux

La méthode de filtrage par transformations linéaires est plus facile à mettre en oeuvre et semble être plus performante que celle du filtrage multicanaux.

Une étude comparative entreprise par Ng *et al.* [Ng.92], a révélé que les filtres de Gabor sont inadaptés pour les textures stochastiques qui ne présentent pas une fréquence ou une orientation dominante. En outre, pour des textures structurées, les performances de ces filtres bien qu'étant nettement meilleures, restaient néanmoins en deçà de celles des transformations linéaires locales.

Le fait que certains masques de Laws soient similaires à des filtres discrets de Gabor tronqués, donne à penser qu'il existe cependant un lien entre ces deux méthodes. Ng *et al.* ont montré que bien que fondamentalement différentes, ces méthodes ont effectivement une structure de représentation identique.

La Transformée en ondelettes qui peut être interprétée aussi bien en terme fréquentiel qu'en terme de filtrage linéaire, semble pouvoir constituer le point de rencontre de ces deux approches. Les propriétés remarquables de cette transformée, ainsi que la souplesse de sa mise en oeuvre, sont discutées dans la seconde partie où sera également décrite l'approche que nous avons adoptée pour son application à l'analyse de textures.

IIème PARTIE

ONDELETTES ET ANALYSE DE TEXTURES

4

TRANSFORMATIONS EN ONDELETTES

L'analyse de Fourier introduite au début du 19^e siècle par Joseph Fourier, permet de repérer les composantes physiques d'un signal grâce à une analyse spectrale réalisée à l'aide des séries ou de l'intégrale de Fourier. Toutefois l'étendue infinie de ses fonctions de base, la rend inadaptée pour l'analyse des signaux bien localisés en temps et donc infiniment étalés en fréquence.

Pour pallier cette défaillance de la méthode de Fourier, Dennis Gabor dans les années 1940, a proposé d'analyser le signal "par bloc" à l'aide d'une fenêtre de pondération gaussienne qui glisse sur le graphe du signal. Cette méthode d'analyse connue sous le nom de *Transformée de Fourier à fenêtre glissante* ou de *Transformée de Gabor* (TG), permet de définir le spectre fréquentiel du signal sur un ensemble de plages temporelles localisées. Il est ainsi possible de suivre l'évolution dans le temps du contenu spectral du signal et de réaliser par conséquent, une analyse *temps-fréquence* du signal.

A l'usage, cette nouvelle méthode d'analyse bien que fructueuse dans plusieurs applications, a révélé une insuffisance : la longueur fixe de la fenêtre glissante limite la résolution temporelle et ne permet pas d'analyser simultanément des phénomènes dont les échelles de temps sont différentes.

C'est pourquoi le géophysicien Jean Morlet a proposé en 1983, alors qu'il s'intéressait à l'étude des signaux sismiques, une nouvelle méthode d'analyse : l'*analyse par ondelettes*. L'idée sous-jacente est de décomposer le signal à analyser sur une famille particulière de fonctions obtenues par translation et dilatation d'une fonction prototype appelée ondelette mère. Ce procédé permet de réaliser une analyse temporelle fine du signal lorsque l'ondelette est contractée et une analyse fréquentielle fine lorsqu'elle est dilatée. Ce réglage progressif de la forme de l'ondelette permet d'effectuer un zoom sur n'importe quelle singularité intéressante du signal.

Une telle approche était déjà apparue au début du 20^e siècle à travers les recherches de Haar, Franklin et Littlewood-Paley pour suppléer la Transformation de Fourier dans l'examen de la régularité et des propriétés locales d'une fonction [Coh.92a].

4.1 Définition des ondelettes

4.1.1 Ondelettes continues

Les ondelettes sont des fonctions $\psi_{a,b}(t)$ générées à partir d'une fonction $\psi(t)$, par translations (paramètre b , $b \in \mathbf{R}$) et par dilatations (paramètre a , $a \neq 0$) [Gro.84] dont la forme générale est :

$$\psi_{a,b}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad (4.1)$$

La quantité b est le paramètre de localisation temporelle, tandis que $1/a$ est le paramètre fréquentiel.

Quand a est petit, la résolution est fine dans le domaine temporel et grossière dans le domaine fréquentiel (*figure 4.1*). Les ondelettes étroites que l'on obtient dans ce cas, permettent d'extraire les détails fins du signal. Par contre les valeurs élevées de a donnent lieu à des ondelettes larges et plates qui mettent en évidence les traits grossiers du signal.

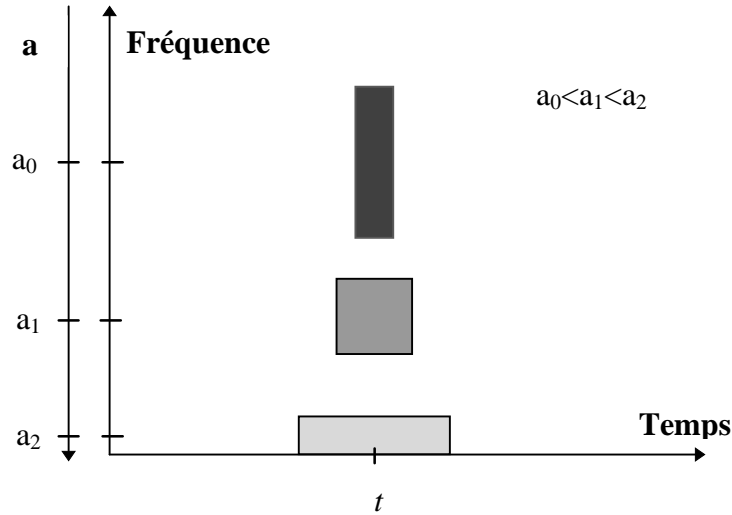


Figure 4.1 : Forme d'une cellule de résolution d'ondelettes dans l'espace temps-fréquence

4.1.2 Transformée en ondelettes continue

Un signal $S(t)$ est décomposé par $\psi_{a,b}(t)$ en :

$$D_{a,b} = \int_{-\infty}^{+\infty} S(t) \psi_{a,b}^*(t) dt \quad (4.2)$$

avec $\psi_{a,b}^*$ conjugué de $\psi_{a,b}$, $a \neq 0$ et $b \in \mathbf{R}$.

L'ensemble des coefficients $(D_{a,b})$ constitue la transformée en ondelettes continue.

Si la fonction ondelette mère $\psi(t)$ satisfait la condition d'admissibilité suivante :

$$C_{\psi} = \int_{-\infty}^{+\infty} (|\Psi(\omega)|^2 / |\omega|) d\omega < \infty$$

(où Ψ désigne la Transformée de Fourier (TF) de ψ),

le signal $S(t)$ peut être reconstruit. La transformée en ondelettes continue inverse est alors définie par :

$$S(t) = \frac{1}{C_{\psi}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} D_{a,b} \psi_{a,b}(t) \frac{da db}{a^2}$$

La condition d'admissibilité impose à $\psi(t)$ d'être oscillante. La *figure 4.2* donne un exemple de fonction $\psi(t)$ proposée par Meyer [Mey.90].

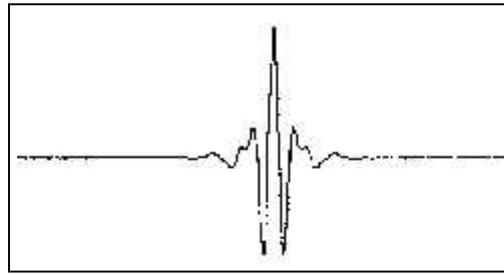


Figure 4.2 : Ondelette de Meyer

La transformée en ondelettes permet de caractériser la régularité ou l'irrégularité du signal. Cette propriété intéressante qui fait de cette transformée un outil performant pour l'analyse des signaux, est illustrée par la *figure 4.3*.

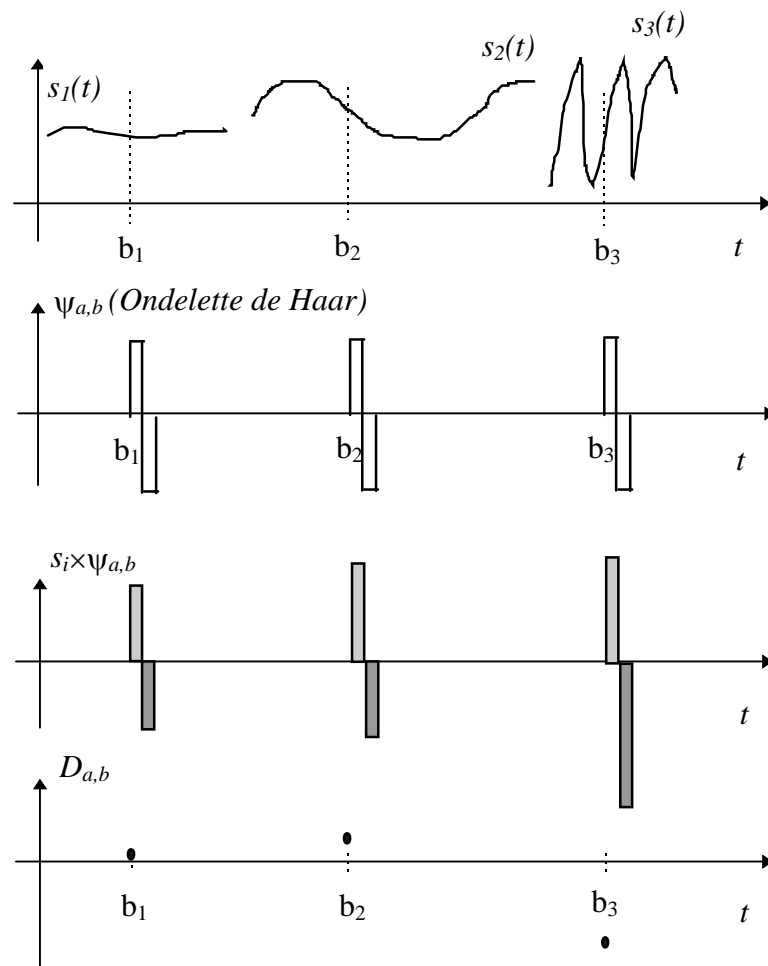


Figure 4.3 : Exemple de calcul de la transformée en ondelettes continue d'un signal avec l'ondelette de Haar.

On constate, en valeur absolue, que :

- Si le signal est à très basse fréquence (cas de $s_1(t)$), le coefficient $|D_{a,b}|$ prend des valeurs très faibles.
- Si le signal est à basse fréquence (comme $s_2(t)$), la valeur du coefficient $|D_{a,b}|$ est relativement faible.
- Par contre, si le signal est de haute fréquence (à l'image de $s_3(t)$), le coefficient $|D_{a,b}|$ prend une valeur plus élevée. Cette valeur est particulièrement importante dans le cas où le paramètre fréquentiel $1/a$ de l'ondelette est de l'ordre de grandeur de la fréquence centrale du signal.

4.1.3 Ondelettes orthogonales

La famille de fonctions $\psi_{a,b}(t)$ constitue une base d'ondelettes orthonormale de l'espace $L^2(\mathbf{R})$ des fonctions de carré intégrable sur \mathbf{R} , si les paramètres a et b sont discrétisés suivant : $a = 2^j$ et $b = k2^j$ [Mey.90].

Cette propriété permet :

- d'une part de calculer à une échelle j donnée, les coefficients d'ondelettes $d_{j,k}$ d'un signal $S(t)$ par :

$$d_{j,k} = \langle S(t), \psi_{j,k}(t) \rangle = \int \psi_{j,k}(t) S(t) dt \quad (4.3)$$

$$\text{avec } \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \text{ et } k \in \mathbf{Z}, \quad (4.4)$$

- et d'autre part de reconstruire $S(t)$ par :

$$S(t) = \sum_k d_{j,k} \psi_{j,k}(t) \quad (4.5)$$

Un exemple classique de base orthonormale d'ondelettes est celle qui est générée par la fonction de Haar. Cette fonction est définie par :

$$\psi(t) = \begin{cases} 1 & \text{pour } 0 \leq t \leq 1/2 \\ -1 & \text{pour } 1/2 \leq t \leq 1 \\ 0 & \text{ailleurs} \end{cases}$$

La discrétisation du paramètre a entraîne logiquement l'introduction du concept d'échelle en lieu et place de la fréquence. L'utilisation de la notion d'échelle permet dans une certaine mesure, de contourner la limite de résolution préjudiciable à la Transformation de Gabor, et qui est due au principe d'incertitude temps-fréquence [Ale.92].

L'application du concept d'échelle à l'analyse par ondelettes prend tout son sens quand cette analyse est vue au travers d'une analyse multirésolution. S'appuyant sur la théorie des ondelettes orthonormales, Meyer [Mey.90] et Mallat [Mal.89a] ont donc développé l'analyse multirésolution orthogonale par ondelettes.

4.2 Transformations en ondelettes et Analyse multirésolution

Une analyse multirésolution permet de représenter le signal comme une limite de ses approximations successives, où chaque approximation est une version lissée de la précédente.

4.2.1 Définition des fonctions d'échelle

L'analyse multirésolution *dyadique* qui sera ci-dessous développée, tient son appellation du fait qu'un facteur de résolution 2 lie deux échelles consécutives.

Elle est définie par une suite croissante de sous-espaces vectoriels fermés emboîtés $\{V_j\}_{j \in \mathbb{Z}}$ vérifiant les propriétés :

- $\bigcap_{j \in \mathbb{Z}} V_j = 0$, $\bigcup_{j \in \mathbb{Z}} V_j = L^2(R)$,
- $S(t) \in V_j \Leftrightarrow S(2t) \in V_{j+1}$.
- $\exists \phi(t) \in V_0$ tel que $\{\phi(t-k)\}_{k \in \mathbb{Z}}$ constituent une base orthonormée de V_0 .

De façon à introduire la notion de multirésolution, on génère à chaque niveau de résolution j et à partir de $\phi(t)$, une famille de fonctions d'échelle $\{\phi_{j,k}(t)\}_{k \in \mathbb{Z}}$ définies par :

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (4.6)$$

Cette famille $\{\phi_{j,k}(t)\}$ constitue pour chaque j , une base orthonormale du sous-espace vectoriel V_j .

Du fait que $\phi \in V_0$ et que $V_0 \subset V_1$, il existe une séquence $\{h_k\}$ telle que la fonction d'échelle ϕ satisfait l'équation de dilatation ("*The two scales difference equation*") suivante :

$$\phi(t) = \sqrt{2} \sum_k h_k \phi(2t - k) \quad (4.7)$$

On montre que les coefficients h_k sont ceux d'un filtre $H(\omega)$ définie par :

$$H(\omega) = \sum_{k=-\infty}^{+\infty} h_k e^{-ik\omega},$$

Ce qui conduit à l'équation :

$$\Phi(2\omega) = H(\omega)\Phi(\omega) \quad (4.8)$$

$H(\omega)$ satisfait les propriétés suivantes :

- $H(0) = 1$, (4.9)
- $\forall \omega \in \mathbb{R}, |H(\omega)|^2 + |H(\omega + \pi)|^2 = 1$, (propriété des filtres conjugués)
- $\exists \rho > 0 / \forall \omega \in \mathbb{R}, |\omega| < \pi/2 \Rightarrow |H(\omega)| > \rho$

4.2.2 Fonctions d'échelle et fonctions ondelette associées

Les espaces $\{V_j\}_{j \in \mathbb{Z}}$ étant emboîtés et fermés, pour tout indice j , il existe un unique espace W_j inclus dans V_{j+1} et orthogonal à V_j , tel que V_{j+1} soit la somme directe (notée \oplus) de V_j et W_j :

$$V_{j+1} = V_j \oplus W_j \quad (4.10)$$

L'espace W_j complémentaire de V_j , contient la différence d'information entre une approximation au niveau de résolution $j+1$, et une approximation au niveau de résolution j .

Les fonctions ondelettes $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ précédemment définies, constituent à l'échelle j , une base orthonormale de W_j , si elles sont liées aux fonctions d'échelle qui leur sont associées par :

$$\psi(t) = \sqrt{2} \sum_k g_k \phi(2t - k) \quad (4.11)$$

$$\text{avec } g_k = (-1)^k h_{1-k}$$

Ce qui donne par transformation de Fourier :

$$\Psi(2\omega) = G(\omega)\Phi(\omega) \quad (4.12)$$

$$\text{avec } G(\omega) = H^*(\omega + \pi)e^{-i\omega} \quad (4.13)$$

où H^* est le complexe conjugué de H .

La satisfaction de cette condition permet à un niveau de résolution donnée, d'associer une fonction d'ondelette à une fonction d'échelle.

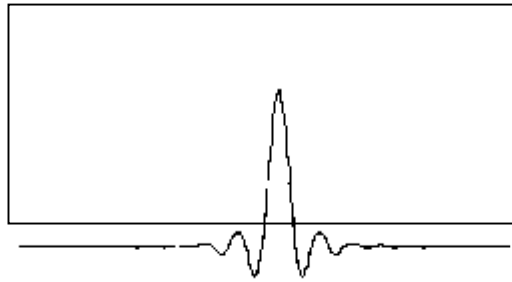


Figure 4.4 : Fonction d'échelle associée à l'ondelette de Meyer

La fonction d'échelle $\phi(t)$ associée à l'ondelette de Meyer de la *figure 4.2*, est présentée à la *figure 4.4*. Plusieurs autres exemples d'ondelettes et de fonctions d'échelles associées peuvent être trouvés dans la littérature traitant des ondelettes [Dau.88, Mal.89a, Coh.92a].

4.2.3 Décomposition en ondelettes orthogonales

La décomposition en séries d'ondelettes (SO) orthogonales, ou simplement en ondelettes, s'effectue en projetant le signal sur les bases orthonormales de fonctions d'échelle et d'ondelettes.

La projection du signal $S(t)$ sur les espaces V_j et W_j se décompose respectivement sur les bases $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$ et $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ en les signaux d'expressions S_j et D_j :

$$S_j = \sum_k s_{j,k} \phi_{j,k} \text{ et } D_j = \sum_k d_{j,k} \psi_{j,k} \quad (4.14)$$

Le signal S_j est interprété comme l'approximation du signal $S(t)$, tandis que D_j est vu comme un signal de détail qui renferme les informations nécessaires à la reconstruction du signal d'entrée.

On peut de même envisager que puisque $V_{j+1} = V_j + W_j$, un signal S_{j+1} de V_{j+1} , se décompose en ses composantes sur V_j et W_j :

$$S_{j+1} = S_j + D_j \quad (4.15)$$

L'analyse multirésolution peut être par la suite obtenue en itérant cette décomposition de base. Il suffit à chaque étape de considérer le signal d'approximation qui vient d'être calculé, comme point d'entrée de la décomposition à effectuer (*figure 4.5*).

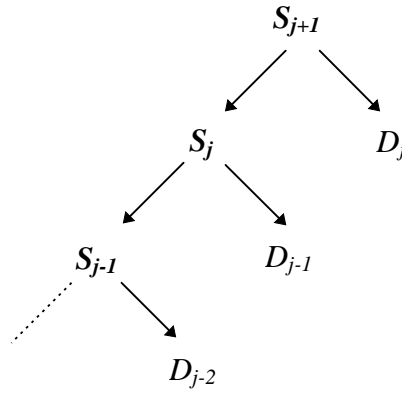


Figure 4.5 : Schéma de décomposition en ondelettes d'un signal

Si l'on suppose que le signal d'entrée est un signal de résolution 2^0 , la décomposition en ondelettes du signal sur J niveaux, fournit une approximation grossière de ce signal au niveau de résolution J , et des signaux de détails aux différents niveaux de résolutions j (avec $1 \leq j \leq J$). Cette décomposition qui est complète et orthogonale, s'écrit :

$$\begin{aligned}
 S_0(t) &= \sum_k s_{0,k} \phi_{0,k}(t) \\
 S_0(t) &= \sum_k (s_{-J,k} \phi_{-J,k}(t) + \sum_{j=1}^J (d_{-j,k} \psi_{-j,k}(t))) \quad (4.16)
 \end{aligned}$$

Cette décomposition est unique et s'effectue sans perte et sans redondance d'information. Son expression montre le caractère multirésolution de l'analyse. Le passage d'un niveau à un autre, correspond à une progression d'un octave dans le plan de Fourier. Cette propriété résulte de la décomposition dyadique de l'analyse.

Le calcul pratique des coefficients de la décomposition s'effectue au moyen d'un algorithme performant que Mallat a proposé et qui est fondé sur la convolution du signal avec des filtres conjugués en quadrature (CQF) [Mal.89a].

4.2.4 Mise en oeuvre algorithmique de la décomposition

L'implémentation pratique des séries d'ondelettes présuppose la discrétisation du signal d'entrée. Cette contrainte permet d'introduire la notion de Transformée en Ondelettes Discrètes (TOD) dont le parallélisme avec les SO a été démontré par Rioul [Rio.93a].

Mallat [Mal.89a] a montré qu'en considérant le signal d'entrée S_n comme la suite des coefficients d'échelle $s_{0,n}$, au niveau de résolution 0,

- les coefficients d'ondelettes s'obtiennent de façon itérative par les relations ($j \leq 0$) :

$$\begin{aligned} s_{j-1,n} &= \sum_k h_{k-2n} s_{j,k} \\ d_{j-1,n} &= \sum_k g_{k-2n} s_{j,k} \end{aligned} \quad (4.17)$$

- et que le calcul de ces coefficients peut être réalisé à partir d'un banc de filtres présentant des propriétés similaires aux filtres conjugués en quadrature (CQF) que Smith et Barnwell ont antérieurement introduit [Smi.86].

Ces résultats permettent d'interpréter la TOD comme un filtrage linéaire par banc de filtres itérés en octaves : les signaux de coefficients $s_{j,n}$ passent à travers une paire de filtres H et G et sont décomposés en deux sous-bandes. Les signaux résultants de cette opération sont ensuite décimés ou sous-échantillonnés (un échantillon sur deux est conservé) (*figure 4.6*). Le processus peut être ensuite itéré jusqu'au niveau de résolution désiré.

H et G qui sont les TF de h et g , correspondent respectivement à des filtres passe-bas et passe-haut. Ils (ou plus exactement leurs para-conjugués) permettent également de réaliser la synthèse du signal (*figure 4.6*). La formule de reconstruction exacte du signal s'écrit :

$$s_{j,n} = \sum_k (h_{n-2k} s_{j-1,k} + g_{n-2k} c_{j-1,k}) \quad (4.18)$$

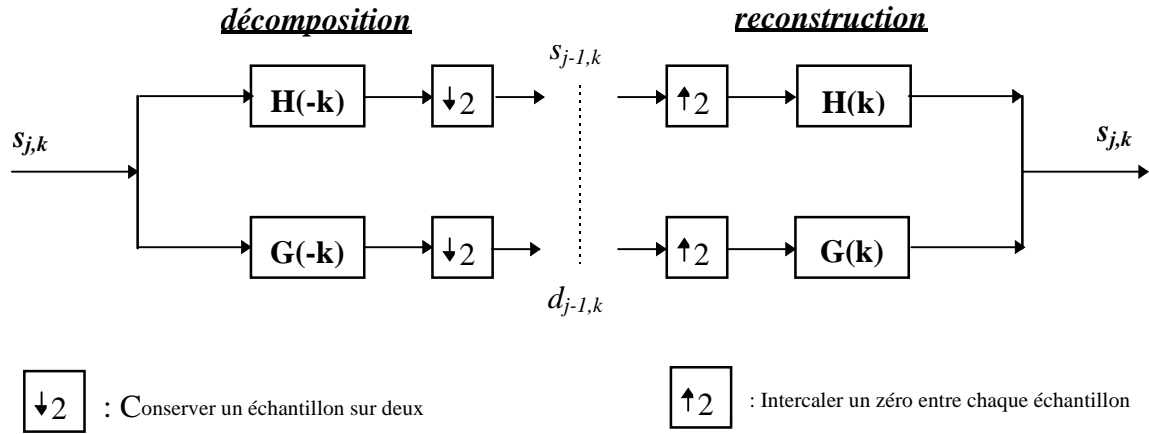


Figure 4.6 : TOD et structure du banc de filtres

4.2.5 Extension de la TOD au cas bidimensionnel 2D (cas des images)

La décomposition d'une image I_{j+1} de niveau de résolution $j+1$ donne au niveau de résolution j , une image I_j et trois images de détails D_j^1 , D_j^2 , D_j^3 , qui correspondent aux pertes d'informations suivant les directions horizontale (indice 1), verticale (indice 2) et diagonale (indice 3).

$$I_{j+1} = I_j + D_j^1 + D_j^2 + D_j^3 \quad (4.19)$$

$$\text{avec } I_{-j} = \sum_k i_{-j,k} \phi_{-j,k} \quad \text{et} \quad D_j^i = \sum_k d_{j,k}^i \psi_{j,k}^i \quad (4.20)$$

pour $i=1, 2$ et 3 , et $k=(k_x, k_y) \in \mathbb{Z}^2$.

La *figure 4.7* présente les quatre images obtenues en décomposant au niveau de résolution 1, une image de mosaïque de textures.

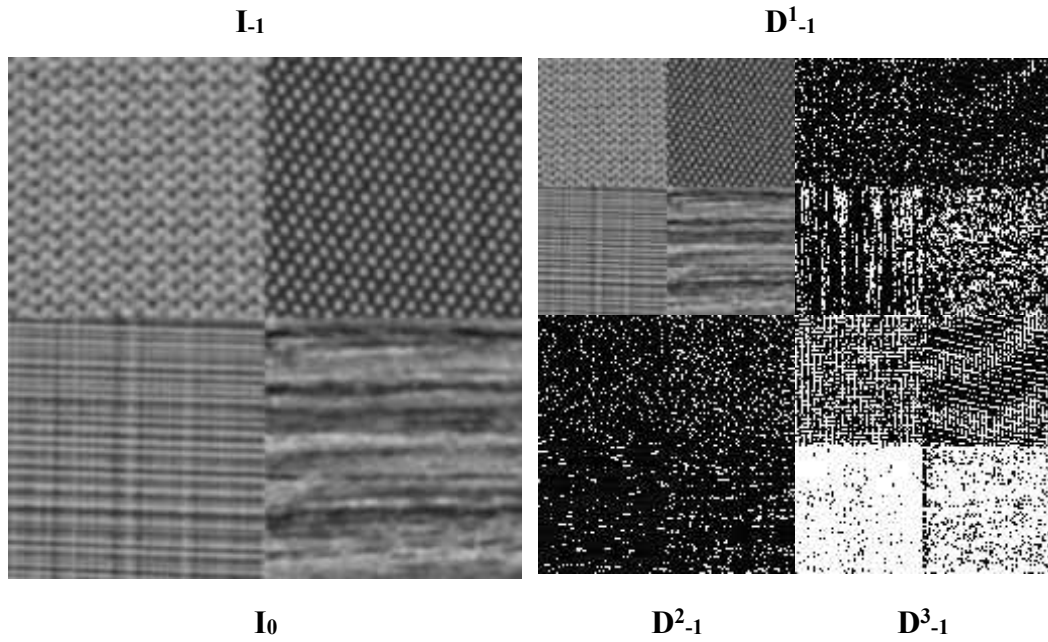


Figure 4.7 : Décomposition en ondelettes au niveau de résolution 1 d'une mosaïque de textures

La mise en oeuvre des filtres se fait suivant le schéma de base de décomposition de Mallat décrit précédemment (*figure 4.6*), en filtrant d'abord les lignes, puis les colonnes (*figure 4.8*), et sous l'hypothèse que H et G sont des filtres séparables. Ce qui revient à réaliser la TOD au moyen de quatre filtres définis ci-dessous, liés aux fonctions d'échelle et d'ondelettes par :

$$\Phi(2\omega_x, 2\omega_y) = H(\omega_x, \omega_y)\Phi(\omega_x, \omega_y) \text{ et } \Psi^i(2\omega_x, 2\omega_y) = H^i(\omega_x, \omega_y)\Phi(\omega_x, \omega_y)$$

avec :

$$H(\omega_x, \omega_y) = H(\omega_x)H(\omega_y), \quad H^2(\omega_x, \omega_y) = G(\omega_x)H(\omega_y),$$

$$H^1(\omega_x, \omega_y) = H(\omega_x)G(\omega_y), \quad H^3(\omega_x, \omega_y) = G(\omega_x)G(\omega_y), \quad (4.21)$$

et : $G(\omega) = H^*(\omega + \pi)e^{-i\omega}$.

Les trois filtres H^i réagissent préférentiellement à trois orientations différentes: 0° pour H^1 , 90° pour H^2 , et 45° pour H^3 . D'autres orientations peuvent être obtenues en utilisant des filtres non séparables. La dilatation (paramètre j) effectuée dans chaque dimension, est alors remplacée par une rotation et une dilatation réalisée simultanément dans les deux dimensions. Ce type de transformation en ondelettes est connu sous le nom de Transformation Circulaire [Ant.93].

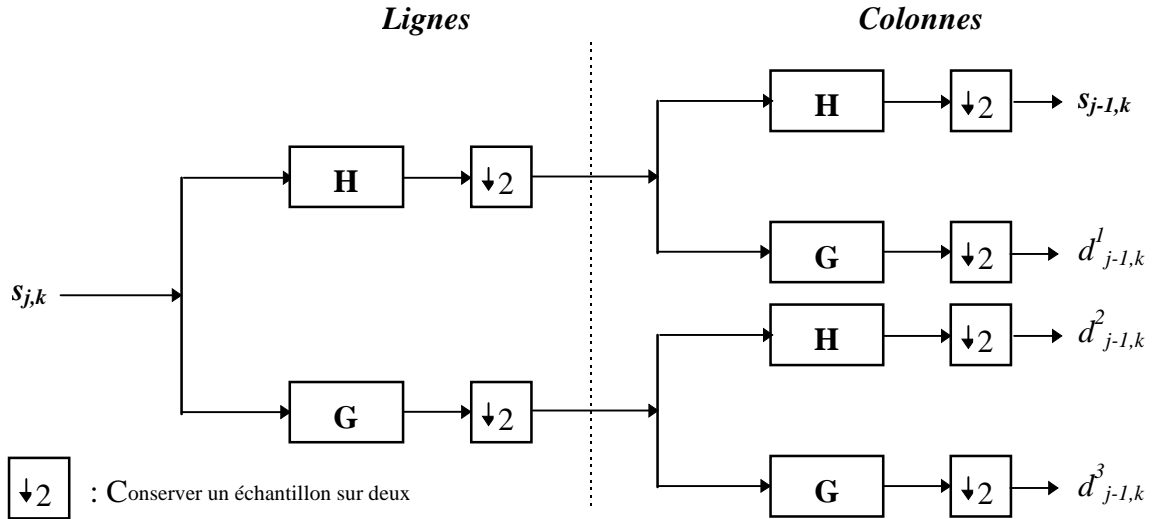


Figure 4.8 : Schéma de calcul de la TOD d'une image

La sélectivité à l'orientation des ondelettes, permet d'interpréter la TOD comme une transformation associée à une décomposition de l'image en canaux fréquentiels d'orientations sélectives et de bandes passantes constantes sur une échelle logarithmique [Mal.89b]. Ce qui lui a valu d'être proposée comme modèle de description du traitement visuel humain.

L'application de l'algorithme de Mallat pour le calcul de la TOD, conduit dans le cas 2D, à une représentation pyramidale. La pyramide est une structure où l'image est représentée à plusieurs niveaux de résolutions. Si l'image initiale de taille $2^m \times 2^m$ constitue le niveau de résolution 0 de la pyramide, au niveau n , la taille de l'image devient $2^{m-n} \times 2^{m-n}$; le facteur de résolution entre deux niveaux étant égal à 2. La représentation pyramidale par la TOD s'obtient en itérant le pas de base de la décomposition sur plusieurs niveaux de résolution (figure 4.9).

Les images de détails faiblement corrélées extraites à chaque niveau de résolution, permettent de reconstruire l'image d'entrée (en suréchantillonnant convenablement les différentes images des différentes résolutions) :

$$I_0 \leftrightarrow I_{-J} + D_{-J}^1 + D_{-J}^2 + D_{-J}^3 + D_{-J+1}^1 + D_{-J+1}^2 + D_{-J+1}^3 + \dots + D_{-1}^1 + D_{-1}^2 + D_{-1}^3$$

où le symbole + a le sens de *et* plutôt que celui de l'addition.

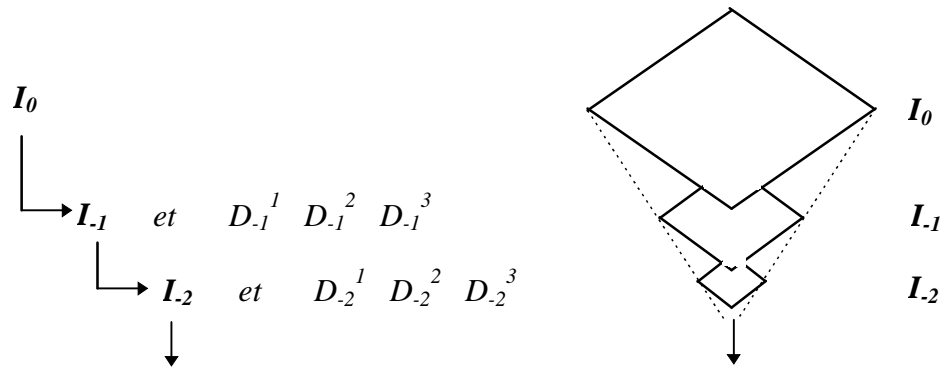


Figure 4.9 : Construction d'une pyramide par décomposition en ondelettes d'une image

4.2.6 Comparaison de la TOD avec d'autres méthodes de représentation pyramidale

La première méthode de représentation pyramidale est celle de Tanimoto et Pavlidis [Tan.75] qui ont proposé de générer la pyramide par un sous échantillonnage itératif de l'image d'entrée. Leur méthode présente l'inconvénient de fournir des approximations successives un peu trop éloignées de l'image traitée. De plus, aucune information inter-résolution ne permet de revenir à la résolution de départ.

En vue d'améliorer ce premier algorithme, Burt [Bur.81] a proposé de convoluer l'image d'un niveau de résolution donnée par une fonction gaussienne, afin d'obtenir l'image de niveau inférieur. Le choix de la fonction gaussienne se justifiait par les résultats des travaux de Marr [Mar.82] sur la vision humaine. La pyramide construite à partir de la méthode de Burt est connue sous le nom de "pyramide gaussienne".

Pour obtenir les détails inter-résolutions, Burt et Adelson [Bur.83] effectuent une différence entre deux niveaux de la pyramide gaussienne et construisent ainsi la "pyramide laplacienne". Cette dernière renferme à chaque niveau, des détails perdus entre deux résolutions consécutives. Malheureusement ces détails sont corrélés entre eux. Il est donc difficile de différencier les informations dues à la redondance de la représentation, de celles contenues dans l'image d'entrée. En outre, la représentation pyramidale de Burt comporte plus de pixels que l'image originale.

Les pyramides construites à partir de la TOD pour des bases d'ondelettes orthogonales, ne présentent pas ces inconvénients. La non corrélation des images de détails et la non redondance d'informations garantissent une reconstruction exacte et marquent la supériorité de la TOD sur les précédentes techniques de représentation pyramidale. Cette supériorité se trouve confirmée par la facilité avec laquelle la construction de la pyramide par ondelettes est mise en oeuvre.

4.3 Autres types de transformations en ondelettes

L'analyse par ondelettes a été l'objet d'une grande attention de la part des chercheurs depuis son introduction. Cet intérêt s'est traduit par l'émergence d'idées nouvelles autour du concept de base. Le développement détaillé de ces idées qui sont succinctement présentées ici, peut être trouvé dans la littérature [Rio.91, Coh.92a, Chu.92].

4.3.1 Analyse multirésolution orthogonale en quinconce

Un autre type d'analyse multirésolution orthogonale pour les images a été développé par Feauveau [Fea.90]. Il s'agit de l'analyse multirésolution en quinconce qui est caractérisée par un facteur de résolution de $\sqrt{2}$. Cette analyse intéressante, décompose l'image en deux fois plus d'échelles que la précédente. Elle permet de ce fait d'accéder à une représentation plus fine de l'image sur les différents niveaux de résolution. Les filtres associés à cette analyse multirésolution sont non séparables et non sélectifs à l'orientation.

4.3.2 Ondelettes biorthogonales

La propriété d'orthogonalité des bases d'ondelettes qui ont été précédemment présentées, conduit à la construction de filtres complexes et asymétriques, et prive les ondelettes associées à ces filtres de propriétés comme la régularité et la symétrie qui sont intéressantes pour le traitement du signal et des images. Pour s'affranchir de la contrainte limitative d'orthogonalité, Cohen, *et al* [Coh.92b] ont introduit les bases d'ondelettes biorthogonales. Cette approche offre une plus grande flexibilité dans le choix des fonctions de base d'ondelettes.

Le principe est d'utiliser une famille de fonctions d'ondelettes formant une base non orthonormale pour l'analyse, et une autre pour la synthèse. Les fonctions de base de ces deux familles sont néanmoins mutuellement orthogonales, sauf lorsque leurs indices coïncident.

On considère donc deux fonctions ondelettes duales $\tilde{\Psi}$ et $\tilde{\Psi}$ telles que le signal S se décompose suivant :

$$S = \sum_{j,k} \langle S, \tilde{\Psi}_{j,k} \rangle \Psi_{j,k} = \sum_{j,k} \langle S, \Psi_{j,k} \rangle \tilde{\Psi}_{j,k} \quad (4.23)$$

La mise en oeuvre pratique de cette décomposition s'appuie sur des schémas pyramidaux du même type que ceux décrits précédemment. La construction des bancs de filtres associés aux ondelettes biorthogonales, a préoccupé plusieurs chercheurs [Vet.92, Coh.92a, Rio.93b]. Ceux-ci ont exploré les relations entre ondelettes, bancs de filtres, et analyse multirésolution. Ils ont ainsi pu mettre au point plusieurs filtres présentant d'intéressantes caractéristiques (symétrie, support compact, régularité des ondelettes associées). De plus, ces filtres sont nettement moins complexes que ceux mis en oeuvre dans le cas orthogonal.

4.3.3 Ondelettes obliques ("wavelet frames")

Les conditions restrictives imposées sur la définition de la fonction $h(t)$ (équation 4.7), qui assurent la reconstruction exacte du signal décomposé en série d'ondelettes orthogonales, peuvent être allégées au prix d'une représentation redondante du signal. C'est ce qui se produit lorsqu'on restreint à une grille discrète les valeurs des paramètres a et b de l'équation (4.1) de l'ondelette continue. Dans ce cas, on fixe un pas de dilatation $a_0 > 1$, et un pas de translation $b_0 \neq 0$.

La famille d'ondelettes obliques que l'on obtient s'écrit alors :

$$\Psi_{m,n}(t) = a_0^{-m/2} \Psi(a_0^{-m} t - kb_0)$$

ce qui laisse apparaître une discrétisation de a et de b respectivement par a^m_0 et $kb_0a^m_0$.

Notons que les ondelettes orthogonales sont obtenues pour $a_0=2$ et $b_0=1$.

Les ondelettes obliques ont été initialement employées par Morlet pour traiter numériquement les signaux sismiques. Daubechies a par la suite développé la théorie relative à ce type d'ondelettes [Dau.90]. Dans cette théorie, le signal est décomposé sur une base surcomplète d'ondelettes à structure oblique, appelée "frame" en anglais. Cette base est définie de telle sorte que le rapport de l'énergie des coefficients $d_{j,k}$ d'ondelettes et de l'énergie E_s du signal $S(t)$, soit borné par deux réels A et B , avec $A>0$, $B>0$, et A,B indépendants de $S(t)$:

$$A \leq \frac{\sum_{j,k} |d_{j,k}|^2}{E_s} \leq B \quad (4.24)$$

La précision de la reconstruction du signal dépend de l'intervalle $[A,B]$. Dans le cas où $A = B = 1$, la famille d'ondelettes est un "tight frame". Elle se comporte alors comme une base orthonormale, sans que ces ondelettes soient nécessairement linéairement indépendantes. Dans ce cas, le calcul des coefficients d'ondelettes peut être effectuée de façon itérative en utilisant un banc de filtres sans sous-échantillonner la sortie des filtres [Uns.95].

On montre que la reconstruction du signal peut être exacte dans le cas général, si on utilise pour la synthèse, des fonctions duales de celles de l'analyse. Ce qui établit un lien avec les ondelettes biorthogonales.

Une des propriétés intéressantes des ondelettes obliques est qu'elles ne délocalisent pas l'information au cours de la décomposition qu'elles mettent en oeuvre. Ce qui n'est pas le cas des ondelettes orthogonales, à cause du sous-échantillonnage qui intervient dans le schéma de la décomposition.

4.3.4 Paquets d'ondelettes

La théorie des paquets d'ondelettes a quand à elle, été motivée par la spécificité des signaux "naturels" (parole, acoustique, ..., textures) qui ont leur spectre davantage localisé

dans les fréquences moyennes. Or la TO décompose le signal par une série de canaux fréquentiels de bandes passantes situées principalement dans les basses fréquences (puisque la décomposition est itérée à partir des images basses fréquences).

Les paquets d'ondelettes qui constituent une extension multibande de la TO pyramidale ont été proposés comme alternative pour l'analyse de ces signaux. Dans cette nouvelle théorie, le concept de base d'ondelettes a été généralisé pour inclure une librairie d'ondes modulées (les paquets d'ondelettes). Cette librairie $\{W_n\}_{n=0}^{\infty}$ peut être générée à partir d'une seule fonction $W_0(t)$ par :

$$\begin{aligned} W_{2n}(t) &= \sqrt{2} \sum_k h(k) W_n(2t - k) \\ W_{2n+1}(t) &= \sqrt{2} \sum_k g(k) W_n(2t - k) \end{aligned} \quad (4.25)$$

où $W_0(t)$ peut être identifié à la fonction $\phi(t)$ et $W_1(t)$ à l'ondelette mère $\psi(t)$.

La librairie peut être ensuite définie comme une collection de bases orthonormales composée de fonctions de la forme : $W_n(2^j x - k)$ où $j, k \in \mathbf{Z}$, et $n \in \mathbf{N}$. Chaque élément de la librairie est déterminé par un sous-ensemble d'indices : un paramètre d'échelle j , un paramètre de localisation k , et un paramètre d'oscillation n .

La décomposition par paquets d'ondelettes se réalise comme dans la formulation de base mais s'effectue, dans le cas d'un signal par exemple, à la fois sur le signal d'approximation et sur le signal de détail. Ce qui conduit à un arbre structuré binaire (*figure 4.10 [a]*). Un intérêt évident de cette approche pour le traitement du signal, est la possibilité d'effectuer une décomposition fréquentielle adaptative [Wic.92].

Dernièrement, l'idée d'une généralisation de la théorie des paquets d'ondelettes a été émise [Her.93, Smi.95]. Elle consiste à effectuer une décomposition par paquets d'ondelettes aussi bien adaptée fréquentiellement que temporellement (ou spatialement) (*figure 4.10 [b] et [c]*).

La décomposition spatiale est réalisée en soumettant l'image à une division de type quad-tree. Le non recouvrement des partitions de l'image obtenues, permet de traiter indépendamment chaque sous-image. Quand à la décomposition fréquentielle, elle est réalisée à partir du banc de filtres miroirs conjugués.

Cette approche permet d'orienter la décomposition de l'image suivant les zones d'intérêt de l'image qui peuvent être des zones de non stationnarité ou des régions à contenu fréquentiel local important.

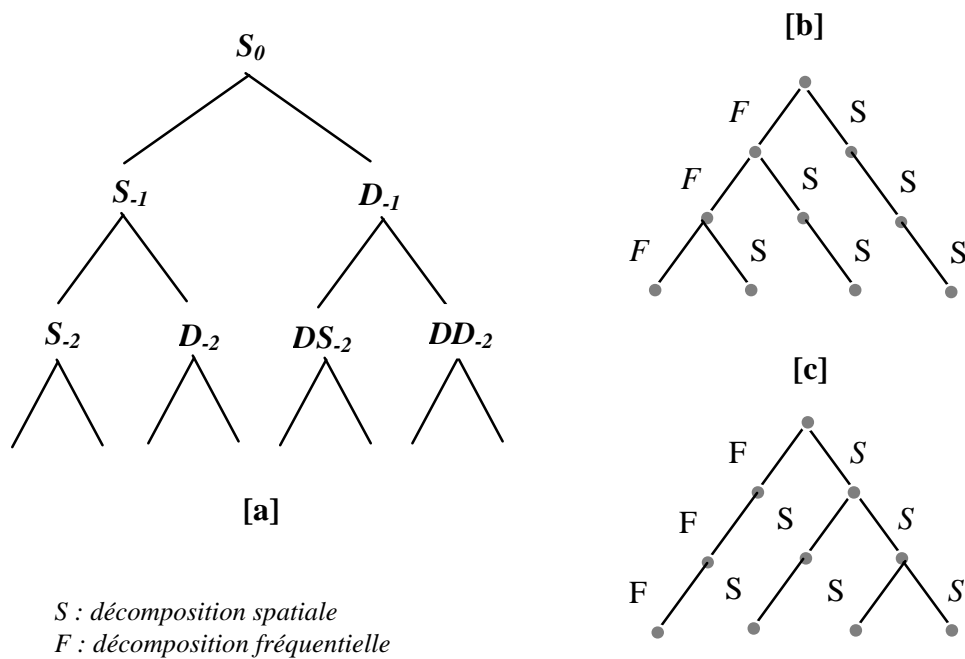


Figure 4.10 : Arbres binaires correspondant aux paquets d'ondelettes
 [a] : paquets d'ondelettes standard
 [b et c] : généralisation des paquets d'ondelettes.

4.4 Applications des ondelettes à l'analyse de textures

Les différents développements qu'ont connus les ondelettes, indiquent le dynamisme des recherches entreprises autour de ce concept. Les chercheurs ont en effet très vite perçu les possibilités énormes de cet outil d'analyse qui s'est imposé comme une alternative aux méthodes existantes et qui a connu de nombreuses applications dans des domaines très variés [Mey.87, Sen.93, Mat.90].

En analyse d'images, les ondelettes ont été principalement appliquées à la compression d'images, à la détection de contours, et à un degré moindre, à l'analyse de textures.

Dans ce dernier cas, du fait de l'*analyse multirésolution* et *multicanaux* qu'elles réalisent, de l'*analyse de la régularité* de l'image qu'elles permettent, et du *traitement du système visuel humain* qu'elles peuvent décrire, les ondelettes apparaissent comme un outil naturel et attractif pour la caractérisation et la segmentation des textures. En outre l'algorithme de calcul rapide dont la TOD bénéficie, fait d'elle un *outil de calcul de choix*.

Mallat [Mal.89a] a été le premier à indiquer en 1989, après avoir présenté la théorie de la représentation en ondelettes orthogonales, que les ondelettes constituent un outil mathématique intéressant pour l'analyse de la texture. Dans son article, il montre qu'il est possible de calculer simplement la dimension fractale d'une texture à partir de sa représentation en ondelettes.

Plusieurs techniques d'analyse de texture par ondelettes ont par la suite été proposées. Elles diffèrent principalement par le type d'ondelettes utilisé dont le choix est souvent guidé par le désir de privilégier les propriétés rattachées soit, à l'orthogonalité des bases d'ondelettes et à leur simplicité de mise en oeuvre [Kis.94, Sal.95], soit, à la description invariante par translation des ondelettes obliques [Uns.95], ou à la décomposition adaptative réalisée par les paquets d'ondelettes [Cha.93, Lai.93].

Les techniques présentées reposent en général, sur la détermination d'un ensemble de propriétés statistiques (énergie, entropie, dispersion, ...) calculées à chaque niveau de la décomposition sur un voisinage défini. Ces statistiques étant déterminées sur chacune des quatre sous-images, on obtient à chaque niveau un quadruplet de caractéristiques. Le vecteur attribut de textures est finalement constitué par les J quadruplets formés aux J différents niveaux de la décomposition.

Cette méthode de caractérisation fournit des résultats intéressants. Cependant, elle conduit à un vecteur de taille imposante, pour peu que la décomposition soit portée à un niveau élevé et que le nombre de statistiques calculé soit supérieur à 1.

Chang et Kuo [Cha.93] ont proposé une méthode d'analyse de textures originale. Leur approche conduit à un nouveau type de décomposition en ondelettes : "The Tree-Structured Wavelet Decomposition" (TSWD). Celui-ci apparaît comme une simplification de la décomposition par paquets d'ondelettes et une optimisation de la décomposition classique.

Dans cette nouvelle approche, l'image décomposée n'est pas nécessairement l'image de basse résolution, (comme c'est le cas dans la décomposition classique) ou les quatre sous-images du niveau courant (comme dans le cas de la décomposition par paquets d'ondelettes). La décomposition est uniquement itérée à partir de la ou des sous-images qui renferment le plus d'informations au sens d'un critère d'énergie (*figure 4.11*).

Cette décomposition prend fin lorsque tous les canaux du niveau courant contiennent des informations insignifiantes. En pratique, le niveau maximal a été fixé à 4.

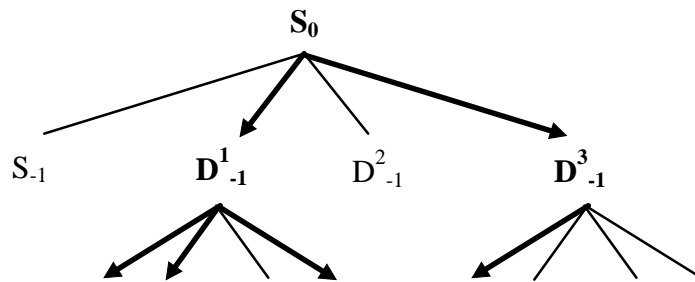


Figure 4.11 : "The Tree-Structured Wavelet Transform"

La texture est ensuite simplement caractérisée par la filiation des canaux décomposés et par une combinaison de l'énergie des canaux les plus dominants des lignées obtenues.

Signalons également que pour augmenter la sélectivité angulaire de la décomposition en ondelettes d'une image, Vautrot et Bonnet [Vau.95] ont proposé de réaliser un filtrage angulaire de l'image d'entrée avec une série de filtres de type gaussien privilégiant chacun, une orientation particulière. Les images obtenues à la sortie de ce filtrage sont ensuite décomposées en ondelettes. Ce procédé fournit des sous-images de différentes orientations, et améliore les performances des systèmes d'analyse de texture par ondelettes.

La segmentation des textures qui constitue l'objectif final de cette thèse, implique un niveau de caractérisation des textures. Ce niveau de reconnaissance occupe une place importante dans la procédure de segmentation. En effet, la qualité des performances de l'opération de segmentation est fortement subordonnée à la pertinence des attributs utilisés pour caractériser les différentes textures.

La méthode de caractérisation de textures que nous avons développée et qui se fonde sur une décomposition en ondelettes, est présentée au chapitre suivant. Elle se différencie principalement par la définition d'un nouvel attribut à caractère multirésolution, appelé *le facteur de forme*, et par le *voisinage* de type *pyramidal* sur lequel cet attribut est calculé.

5

CARACTERISATION ET CLASSIFICATION DE TEXTURES

Une approche efficace de caractérisation des textures, consiste à mesurer les variations des niveaux de gris dans le voisinage immédiat de chacun de ses pixels. Ces fluctuations présentent en effet, une certaine périodicité et des amplitudes qui diffèrent selon la texture.

Les méthodes actuelles d'analyse de textures fondées sur une décomposition en ondelettes, soumettent dans une première phase, l'image de texture au banc de filtres associé à la Transformée en Ondelettes Discrètes (TOD). Ensuite, en vue de caractériser la texture, elles calculent un ensemble de propriétés statistiques sur un voisinage de taille $n \times n$, défini sur chacune des images de la décomposition.

Le voisinage défini dans ces méthodes, est un voisinage "plan" qui comme nous le verrons au chapitre suivant, se révèle peu adapté pour un processus de segmentation de textures opérant sur des données représentées à plusieurs niveaux de résolution.

Nous montrons dans ce chapitre, qu'à la faveur d'une interprétation particulière des coefficients d'ondelettes, il est possible de caractériser la texture par ses variations de niveaux

de gris, en calculant les attributs sur un nouveau type de voisinage à savoir le voisinage "pyramidal". Pour ce faire, nous proposons un nouvel attribut de texture, appelé "facteur de forme", qui permet de réaliser une caractérisation multirésolution de la texture. La pertinence de cet attribut est éprouvée dans deux algorithmes de classification différents qui tirent leur inspiration des propriétés liées à la décomposition par ondelettes.

5.1 Une nouvelle méthode de caractérisation de texture fondée sur la décomposition en ondelettes

5.1.1 Description de la méthode

Par souci de clarté, nous avons choisi de présenter notre méthode en la considérant dans un premier temps, dans le cas monodimensionnel. L'extension à deux dimensions (2D), pour l'analyse d'images de textures, en sera facilitée.

Rappelons que la décomposition en séries d'ondelettes d'un signal $S_0(t)$ au niveau de résolution J , s'écrit (pour $1 < j < J$):

$$S_0(t) = S_{-J}(t) + D_{-J}(t) + D_{-J+1}(t) + D_{-J+2}(t) + \dots + D_{-2}(t) + D_{-1}(t) \quad (5.1)$$

$$\text{avec : } S_{-J}(t) = \sum_k s_{-J,k} \phi_{-J,k}(t) \quad \text{et} \quad D_{-j}(t) = \sum_k d_{-j,k} \psi_{-j,k}(t)$$

Les coefficients $s_{-J,k}$ du signal $S_{-J}(t)$, sont le résultat d'un filtrage passe-bas. Ils peuvent donc être interprétés comme des moyennes locales du signal calculées sur un intervalle de longueur L . Le signal $S_{-J}(t)$ est donc une fonction en escalier dont la taille des marches vaut L (figure 5.1 [A1]).

L'équation (5.1) peut également s'écrire :

$$S_0(t) - S_{-J}(t) = D_{-J}(t) + D_{-J+1}(t) + D_{-J+2}(t) + \dots + D_{-2}(t) + D_{-1}(t) \quad (5.2)$$

Ce qui permet de dire que la somme des signaux de détail caractérise les fluctuations du signal $S_0(t)$ autour du signal $S_{-J}(t)$. Ces fluctuations sont représentées sur le graphe [A2] de la Figure

5.1 dans le cas simple d'une décomposition en ondelettes de Haar.

Ainsi, l'ensemble des coefficients $\{d_{-j,k}\}_{1 \leq k \leq J}$, décrit sur les différents intervalles L , les variations du signal $S_0(t)$ autour des différentes valeurs moyennes locales $s_{-j,k}$.

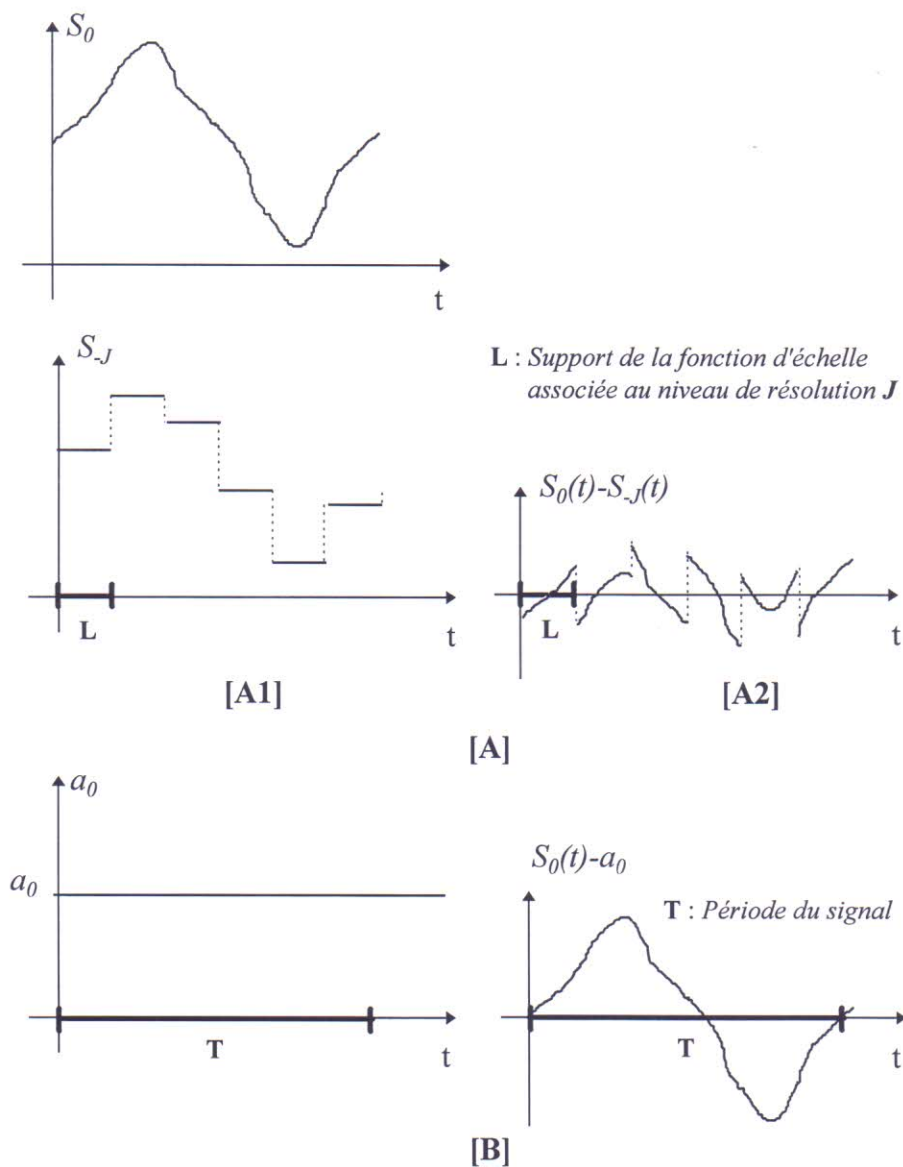


Figure 5.1 : Décomposition en séries [A] d'ondelettes de Haar et [B] de Fourier d'un signal $S(t)$

Cette description est unique dans le cas où la décomposition du signal s'effectue sur une base orthonormale d'ondelettes.

Le signal peut en définitive, être caractérisé par ses variations autour des différentes moyennes locales, en réalisant une combinaison de l'ensemble des coefficients de détail de la décomposition en ondelettes.

Cette interprétation peut également être appliquée à la décomposition en séries de Fourier d'un signal périodique. Dans ce cas, l'intervalle L correspond à la période T du signal, et les termes de pulsation d'expression générale $n(2\pi/T)$ (avec n entier >0) permettent de caractériser les variations du signal autour du terme moyen a_0 (figure 5.1 [BJ]).

L'un des avantages de l'analyse par ondelettes sur celle de Fourier est la grande souplesse qu'offre les ondelettes pour le choix de la longueur de l'intervalle L . Dans la décomposition en séries de Fourier, l'intervalle L qui correspond à la période du signal, est fixé d'emblée par les caractéristiques de ce signal. Tandis que dans le cas de la décomposition en ondelettes cet intervalle est indépendant des propriétés du signal étudié. Sa longueur est fixée par deux paramètres : le support de la fonction d'échelle et le niveau de la décomposition considéré.

Le choix de l'intervalle L par le niveau de la décomposition correspond à la réalisation de l'analyse multirésolution par TOD dont l'intérêt pour l'analyse des signaux a été évoqué au chapitre précédent.

5.1.2 Méthode de caractérisation des textures

Il est aisé d'étendre à la texture (cas 2D), l'interprétation qui vient d'être faite des coefficients d'ondelettes.

L'approche naturelle est de considérer chaque groupe de coefficients de détails (horizontal, vertical et diagonal), comme décrivant suivant une orientation donnée, les variations locales des niveaux de gris de l'image autour de leur moyenne [Lou.95].

La méthode de caractérisation de textures qui découle de cette approche consiste:

- à soumettre la texture au filtrage réalisé par la TOD,

- **puis à caractériser, sur des régions de taille $L \times L$ donnée et suivant différentes orientations, les variations locales des niveaux de gris de la texture autour de leur moyenne; cette caractérisation s'effectuant en réalisant une combinaison des coefficients d'ondelettes des différents niveaux de résolution de la décomposition.**

Cette méthode de caractérisation intègre deux dimensions qui constituent sa spécificité : une dimension dans le plan de l'image (dimension plane), et une autre qui est perpendiculaire à ce plan (dimension en profondeur).

La dimension plane est implicitement prise en compte lors du filtrage induit par la TOD. Ce filtrage permet d'extraire les caractéristiques locales de la texture qui traduisent les relations entre pixels dans un voisinage peu étendu.

La dimension en profondeur apparaît dans la combinaison des informations extraites aux différents niveaux de résolution. Cette dimension permet de réaliser une caractérisation multirésolution de la texture.

Cette dernière dimension constitue un réel avantage de la méthode proposée sur les méthodes classiques qui traitent les différents niveaux de résolution de façon indépendante. Ces méthodes en effet, se trouvent par la suite confrontées à la difficulté du choix du mode de combinaison des informations obtenues séparément à chaque niveau de résolution.

5.1.3 Le voisinage pyramidal

La mise en oeuvre pratique de la TOD par l'algorithme de Mallat, conduit à une représentation pyramidale des données (cf. chapitre 4). Or l'approche envisagée suppose que la texture soit caractérisée à partir d'une combinaison des images de différents niveaux de résolutions. Ce qui impose que les attributs caractéristiques de textures soient déterminés sur un voisinage de type "pyramidal" (*Figure 5.2*).

Ce voisinage constitue une autre particularité de notre méthode. Il se différencie en effet, du voisinage "plan" généralement utilisé dans les méthodes classiques. A un niveau de

résolution j donné, le support du voisinage pyramidal (sur le niveau 0), est de taille plus faible que celui du voisinage plan défini au même niveau j . Nous verrons au chapitre 6 que ce fait constitue un réel avantage dans la résolution d'un problème de segmentation de textures.

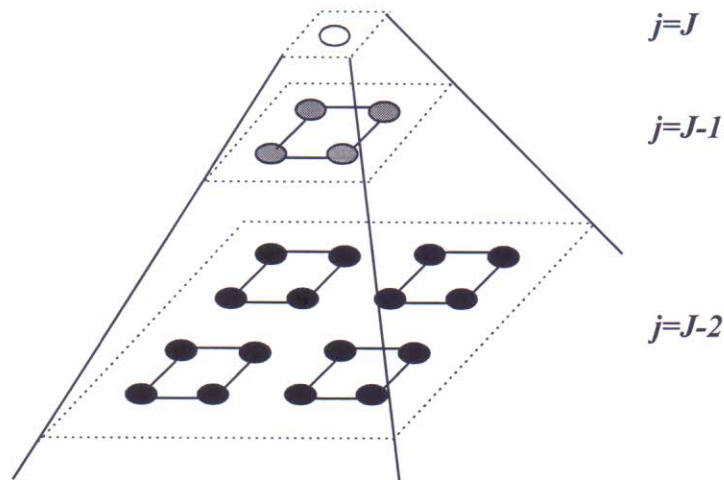


Figure 5.2 : Exemple de voisinage pyramidal

5.1.4 Définition du facteur de forme

Pour mesurer les propriétés caractéristiques de textures sur ce nouveau type de voisinage, il est essentiel de définir un nouvel attribut de texture qui soit conforme à l'approche adoptée pour la caractérisation de la texture.

Ainsi, il est indispensable que l'attribut de texture soit défini de telle sorte :

- qu'il réalise, pour une région de taille définie, la combinaison des coefficients de détails des différents niveaux de résolutions,
- qu'il intègre la notion de moyenne locale (étant entendu que les variations sont caractérisées par rapport à cette valeur), et,
- qu'il se calcule sur un voisinage pyramidal.

Pour prendre en compte tous ces facteurs, nous proposons un nouvel attribut de texture que nous avons dénommé le *facteur de forme*.

Considérons la représentation pyramidale d'une image I_0 , supposée de taille $N_0 \times N_0$ et de résolution $2^0 \times 2^0$, sur une base orthonormale d'ondelettes. Cette représentation au niveau de résolution J , s'écrit :

$$I_0 \leftrightarrow I_{-J} + D_{-J}^1 + D_{-J}^2 + D_{-J}^3 + D_{-J+1}^1 + D_{-J+1}^2 + D_{-J+1}^3 + \dots + D_{-1}^1 + D_{-1}^2 + D_{-1}^3$$

$$\text{avec } I_{-J} = \sum_k i_{-J,k} \phi_{-J,k} \quad \text{et} \quad D_{-j}^l = \sum_k d_{-j,k}^l \psi_{-j,k}^l$$

et pour $l = 1, 2$ et 3 , et $k = (k_x, k_y) \in N_j \times N_j$.

l symbolise une direction particulière (horizontale, verticale, et diagonale respectivement pour $l = 1, 2$ et 3), le signe "+" a le sens de "et", et $N_j \times N_j$ est la taille de l'image de niveau de résolution j .

Le *facteur de forme* est défini pour un pixel k du niveau de décomposition J et pour une direction l donnée, par [Lou.95]:

$$FF_J^l(k) = \frac{\sqrt{(i_{-J,k})^2 + \sum_{j=1}^J \sum_{b \in V_j(k)} (d_{-j,b}^l)^2}}{|i_{-J,k}|}$$

où j est l'indice de résolution et $V_j(k)$ le voisinage plan du pixel k au niveau de résolution j . Pour j donné, $V_j(k)$ correspond à un voisinage de taille $2^{J-j} \times 2^{J-j}$. L'empilement des voisinages $V_j(k)$ (avec $1 \leq j \leq J$) forme le voisinage pyramidal que nous avons précédemment introduit.

La *figure 5.3* montre en exemple pour la direction $l=1$, les images (en gris) dont les pixels interviennent pour le calcul des facteurs de forme au niveau 3.

Tel qu'il a été défini, le numérateur du facteur de forme réalise bien une combinaison des coefficients de détails sur un voisinage pyramidal. Le rapport de cette combinaison à la valeur moyenne locale est effectuée en divisant ce numérateur par le coefficient de l'image

d'approximation $i_{J,k}$. Le facteur de forme peut donc permettre de mesurer le degré des variations locales de niveaux de gris de la texture autour de leur valeur moyenne.

Le terme $i_{J,k}$ délibérément introduit au numérateur fait également du facteur de forme, un indicateur de mesure d'énergie locale (rapportée à la valeur moyenne locale) de texture.

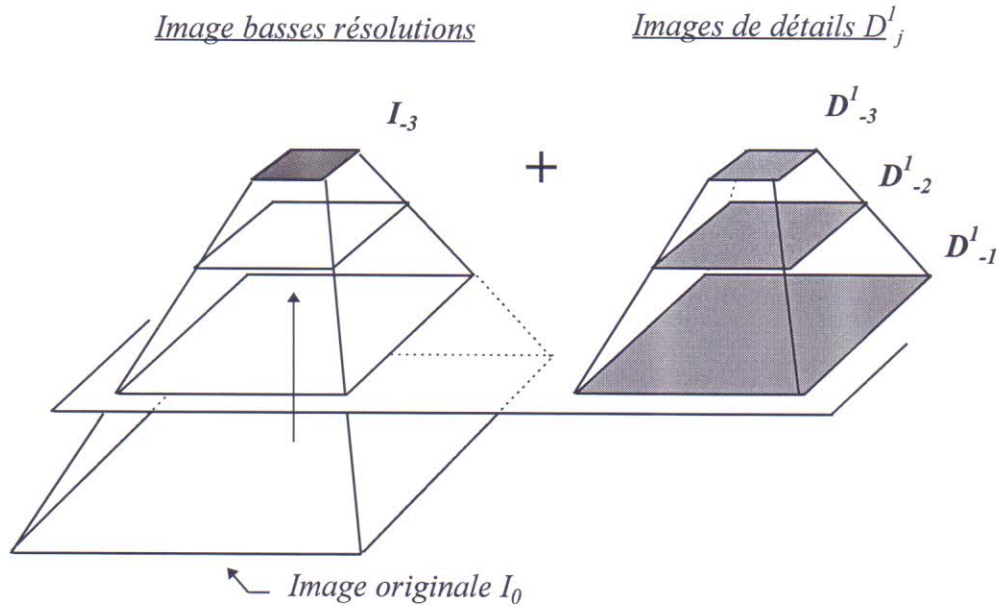


Figure 5.3 : Images considérées pour le calcul du facteur de forme au niveau de résolution 3 dans la direction $l=1$. Les images dont les coefficients interviennent dans ce calcul, sont remplies en gris.

Il est intéressant de remarquer que l'ensemble des coefficients d'ondelettes obtenus en décomposant au niveau de résolution J , une région de taille $2^J \times 2^J$, peut être utilisé pour caractériser la texture. Pour cela, il suffit de considérer comme vecteur d'attributs, le vecteur formé des trois facteurs de forme directionnels suivants : FF^1_J, FF^2_J, FF^3_J .

Il est alors possible d'interpréter ce vecteur d'attributs comme reflétant l'importance des fluctuations directionnelles des niveaux de gris de l'image dans le voisinage spatial immédiat d'un pixel fictif k de l'image originale; la taille de ce voisinage étant $2^J \times 2^J$.

Dans la suite, la texture sera caractérisée au niveau de résolution J , par le vecteur attribut $FF_J = (FF^1_J, FF^2_J, FF^3_J)$.

Les dimensions du voisinage pyramidal sur lequel ce vecteur est calculé, sont automatiquement fixées par le niveau J de la décomposition. Ce voisinage a pour support sur l'image originale, un voisinage plan de taille $2^J \times 2^J$.

5.1.5 Formule simplifiée et récursivité du calcul du facteur de forme

Pour réduire le temps de calcul, le facteur de forme peut être calculé par la formule approchée suivante :

$$FF_J^l(k) = 1 + \frac{\sum_{j=1}^J \sum_{b \in V_j(k)} |d_{-j,b}^l|}{|i_{-J,k}|}$$

Le facteur de forme est donc un réel supérieur à 1.

Pour un niveau de résolution J donné ($J > 1$), sa détermination peut être réalisée de façon récursive. Ainsi, pour une direction donnée, le facteur de forme $FF_J(k)$ d'un pixel k du niveau J , peut être déterminé à partir de la connaissance :

- des quatre facteurs de forme $FF_{J-1}(a)$, $FF_{J-1}(b)$, $FF_{J-1}(c)$, $FF_{J-1}(d)$, des pixels a , b , c et d du niveau $J-1$ correspondant au pixel k (Figure 5.4),
- des coefficients $i_{-(J-1),a}$, $i_{-(J-1),b}$, $i_{-(J-1),c}$, $i_{-(J-1),d}$, et $i_{-J,k}$ des images basse résolution correspondant respectivement aux pixels a , b , c , d et k ,
- et du coefficient de détail $d_{-J,k}$ du pixel k correspondant à la direction choisie.

Pour $J > 1$, il s'exprime par :

$$FF_J(k) = 1 + \frac{d_{-J,k}}{i_{-J,k}} + \frac{i_{-(J-1),a}}{i_{-J,k}} (FF_{J-1}(a) - 1) + \frac{i_{-(J-1),b}}{i_{-J,k}} (FF_{J-1}(b) - 1) \\ + \frac{i_{-(J-1),c}}{i_{-J,k}} (FF_{J-1}(c) - 1) + \frac{i_{-(J-1),d}}{i_{-J,k}} (FF_{J-1}(d) - 1)$$

Cette propriété permet un gain de place mémoire puisqu'il suffit, pour le calcul du facteur de forme au niveau de résolution J , de stocker uniquement les images du niveau inférieur $J-1$.

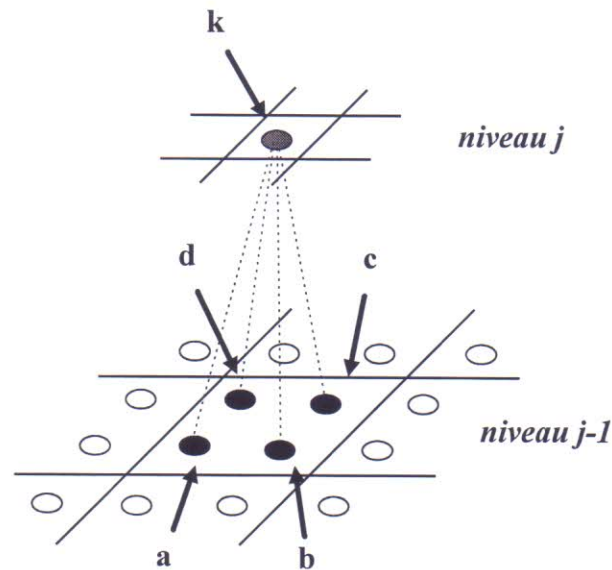


Figure 5.4 : Pixel (k) du niveau j et ses quatre correspondants (a , b , c , d) du niveau inférieur

5.2 Choix de l'ondelette analysante

Le choix de l'ondelette analysante est un aspect important des systèmes d'analyse de textures basées sur la TO. Il est rendu délicat par l'étendue de la gamme de famille d'ondelettes dont on dispose et par l'absence d'une définition précise de la texture qui aurait pu le guider.

Les propriétés les plus importantes souhaitées pour une ondelette sont les suivantes : orthogonalité, support compact, symétrie, régularité, et oscillation.

Les deux dernières propriétés sont liées entre elles. La propriété d'oscillation de l'ondelette se traduit par l'annulation d'un certain nombre de ses moments. Ce nombre permet également de caractériser la régularité du signal.

Pour notre application, nous avons choisi d'utiliser l'ondelette de Haar. Cette ondelette possède les trois premières propriétés précitées, et bénéficie d'une bonne localisation spatiale.

En revanche, elle est mal localisée en fréquence et n'est vraiment pas régulière (elle ne possède qu'un seul moment nul).

Ces faiblesses ne constituent pas en soi un inconvénient pour notre application. En effet :

1. Le principe de notre méthode est de caractériser la texture par les variations de ses niveaux de gris. Il est donc nécessaire d'identifier les fluctuations (même les plus minimes) des niveaux de gris par des coefficients d'ondelette élevés. Cette exigence ne peut pas être satisfaite avec une ondelette régulière qui a plutôt tendance à atténuer les valeurs de ces coefficients lorsque le niveau de résolution j augmente.

2. Ensuite, une étude comparative récente entreprise par Unser [Uns.95], sur trois types d'ondelettes (splines orthogonales de Battle-Lemarié, B-splines et D-splines) de caractéristiques différentes, précise les propriétés les plus pertinentes pour la discrimination des textures. Les conclusions de cette étude expérimentale tiennent en deux points :

- la propriété de régularité ou d'oscillation de l'ondelette ne semble pas constituer un réel avantage pour l'analyse et la discrimination de textures,
- une bonne localisation spatiale des filtres est préférable à celle des ondelettes auxquelles ils sont associés : les expérimentations ont permis de conclure que les filtres qui bénéficient d'une bonne localisation spatiale (comme ceux associés à l'ondelette de Haar), sont plus performants pour la discrimination de textures que ceux qui sont associés à des ondelettes qui sont spatialement bien localisées (comme les B-splines).

Il apparaît donc logique que Unser ait conclu son étude en recommandant l'utilisation de l'ondelette de Haar pour l'analyse de textures.

3. La troisième raison est que l'ondelette de Haar associe des filtres courts. Ce qui constitue un avantage en termes de gains en temps de calcul. Les filtres numériques qui lui sont associés sont : $H=\{1/2;1/2\}$ et $G=\{1/2;-1/2\}$. On peut montrer que la réalisation pratique de la TOD avec cette ondelette, revient à convoluer sans recouvrement, à chaque étape, l'image de basse résolution avec les 4 masques bidimensionnels de taille 2×2 , qui interviennent dans la Transformation Discrète de Hadamard :

$$\frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \frac{1}{4} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \frac{1}{4} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, \frac{1}{4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix};$$

4. Enfin, il a été indiqué précédemment, que le facteur de forme calculé au niveau de résolution J , est représentatif d'une région de taille $2^J \times 2^J$. Seulement, les filtres associés au niveau de résolution J , ont une taille de support généralement plus grande que celle de la dite région. Cela provient du fait que la transformation linéaire locale, réalisée par les filtres associés à l'ondelette choisie, concerne un voisinage de pixels de taille $n \times m$ avec n et m en général supérieurs à 2.

Dans le cas d'une analyse multirésolution dyadique, si l'on désire que la taille de la région et celle des supports des filtres soit strictement identiques, il faut que $n=m=2$. Cette condition n'est remplie que dans un seul cas à savoir, celui de l'ondelette de Haar.

Cette ondelette apparaît donc comme étant celle qui réalise l'analyse dyadique la plus fine, sans recouvrement de régions. Ce qui constitue un atout, notamment dans les problèmes de segmentation de textures.

Ces différentes raisons motivent et expliquent le choix de l'ondelette de Haar. Dans la suite, sauf indications contraires, les développements qui seront présentés sont tous bâtis sur une décomposition avec l'ondelette de Haar.

5.3 Caractérisation de textures à l'aide du facteur de forme

Il est intéressant, avant d'aller plus loin, de s'assurer que le facteur de forme permet de caractériser la texture par ses variations de niveaux de gris suivant différentes directions. Pour cela, nous avons entrepris des expérimentations sur des images de textures synthétiques et naturelles (*Figure 5.5*) tirées de l'album de Brodatz [Bro.66], et de la banque d'images du laboratoire.

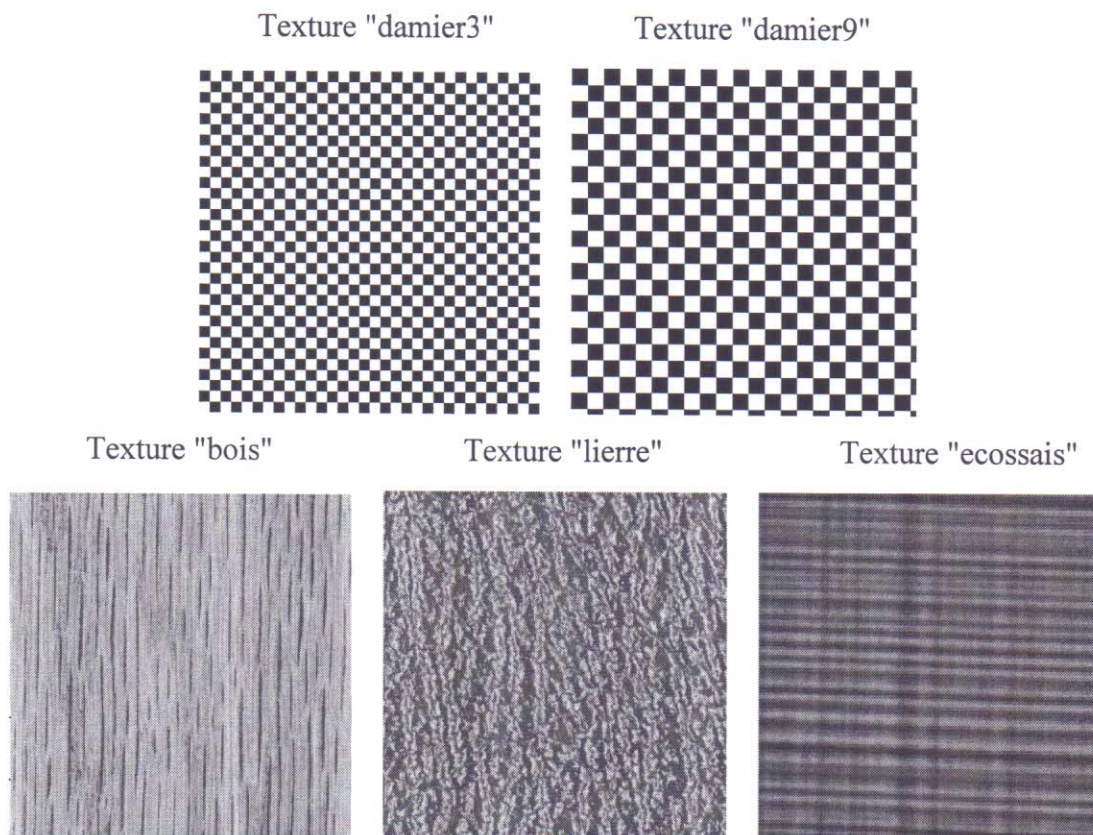


Figure 5.5 : Images test utilisées dans l'expérimentation de caractérisation des textures

Les images utilisées pour ces expérimentations sont de taille 256×256 , avec 256 niveaux de gris. Ce sont les textures synthétiques "damier3" et "damier9" (de même moyenne mais de variations de niveaux de gris différentes), et les textures naturelles de bois (d'orientation verticale), de lierre (qui possède une direction préférentielle oblique), et de tissu écossais (dénommée plus simplement "écossais" et possédant une orientation horizontale).

Pour chacune de ces images, le facteur de forme moyen est déterminé pour les trois différentes directions (H, V, D) et à différents niveaux J (1 à 6) de la décomposition.

Rappelons qu'un niveau de résolution J , correspond sur l'image originale à une région d'intérêt de taille $2^J \times 2^J$.

Les résultats obtenus sont ceux résumés dans le *Tableau 5.1* :

	<i>J=1</i>	<i>J=2</i>	<i>J=3</i>	<i>J=4</i>	<i>J=5</i>	<i>J=6</i>
<i>Damier3</i>						
<i>H</i>	1.216	2.116	5.4	18.61	71.57	281.5
<i>V</i>	1.216	2.116	5.4	18.61	71.57	281.5
<i>D</i>	1.10	1.9	4.58	15.35	58.49	232.3
<i>Damier9</i>						
<i>H</i>	1.094	1.658	3.89	10.74	40.08	156.83
<i>V</i>	1.094	1.658	3.89	10.74	40.08	156.83
<i>D</i>	1.011	1.096	1.61	3.57	11.3	42.8
<i>Lierre</i>						
<i>H</i>	1.06	1.35	2.50	7.10	25.5	105.76
<i>V</i>	1.04	1.21	1.88	4.57	15.33	62.16
<i>D</i>	1.02	1.15	1.67	3.71	11.92	47.6
<i>Bois</i>						
<i>H</i>	1.05	1.25	2.00	5.02	17.05	78.28
<i>V</i>	1.00	1.037	1.15	1.6	3.5	13.19
<i>D</i>	1.00	1.02	1.11	1.45	2.82	9.84
<i>Écossais</i>						
<i>H</i>	1.015	1.081	1.35	2.93	6.58	22.64
<i>V</i>	1.077	1.407	2.67	7.55	27.28	103.05
<i>D</i>	1.004	1.005	1.025	1.096	1.388	2.484

Tableau 5.1 : Facteurs de forme moyen de différents niveaux de résolution et de différentes directions obtenus avec l'ondelette de Haar

Ces résultats suscitent les commentaires suivants :

- Les trois attributs calculés à chaque niveau de la décomposition varient selon la texture et le niveau de la décomposition considéré. Ils peuvent donc permettre de caractériser la texture.
- *Les facteurs de forme rendent compte des variations de niveaux de gris de la texture.* En effet, les facteurs de forme de la texture "damier3" dont les variations de niveaux de gris sont plus importantes, sont supérieurs à ceux de la texture "damier9". Vu sous un autre angle, ces résultats montrent que les facteurs de forme peuvent être utilisés pour caractériser des textures dont la taille des éléments est différente.
- *Les facteurs de forme rendent aussi compte de l'orientation de la texture.* La texture "bois", qui présente une direction verticale dominante, a un facteur de forme horizontal (*H*) nettement supérieur à celui des autres directions. À l'opposé, la texture "écossais" de

direction dominante horizontale, voit son facteur de forme vertical (V) supplanter celui des autres. On peut également noter que la texture "lierre" dont les éléments ont une orientation proche de la verticale, possède des facteurs de forme diagonaux plus élevés que ceux de la texture "bois" qui est franchement verticalement orientée.

- Le caractère discriminant du facteur de forme s'affirme davantage à partir du niveau de résolution $J=3$ c'est-à-dire, pour une région d'intérêt de taille 8×8 sur l'image originale. Cela s'explique par le fait que pour mieux caractériser la texture, il est indispensable de définir un voisinage assez grand afin de se placer dans l'hypothèse de stationnarité qui sous-tend toute analyse de textures. Une analyse sur un voisinage de petite taille n'est pas suffisante pour discriminer la texture. Les variations des niveaux de gris variant considérablement d'un voisinage à un autre.

Les filtres associés à l'ondelette de Haar étant à support compact, il faut monter dans la pyramide de la décomposition pour avoir un voisinage significatif sur lequel la texture peut être correctement caractérisée.

A titre comparatif, nous avons implanté l'ondelette de Daubechies à deux moments nuls [Dau.92]. Les filtres associés à cette ondelette, sont à support compact, et leurs coefficients sont indiqués dans le *Tableau 5.2*.

<i>Coefficients</i>		<i>Daubechies</i>
<i>n</i>	<i>H</i>	<i>G</i>
-2		-0.129409522
-1		-0.224143868
0	0.4829629131	0.8365163073
1	0.8365163073	-0.482962913
2	0.2241438680	
3	-0.129409522	

Tableau 5.2 : Coefficients des filtres associés à l'ondelette de Daubechies à deux moments nuls

L'expérimentation précédente a été reconduite pour les textures "bois", "damier3" et "damier9". Les résultats obtenus sont rapportés dans le *Tableau 5.3*.

On note dans un premier temps, une différence au niveau des valeurs des facteurs de forme. Ces valeurs dépendent naturellement de l'ondelette analysante utilisée.

	<i>J=1</i>	<i>J=2</i>	<i>J=3</i>	<i>J=4</i>	<i>J=5</i>
Bois					
<i>H</i>	1.25	1.75	2.756	4.76	8.772
<i>V</i>	1.083	1.25	1.608	2.272	3.573
<i>D</i>	1.00	1.02	1.052	1.103	1.207
Damier3					
<i>H</i>	1.48	2.125	3.48	6.22	11.7
<i>V</i>	1.59	1.79	2.7	4.41	7.84
<i>D</i>	1.77	1.88	2.84	4.7	8.40
Damier9					
<i>H</i>	1.26	1.43	2.95	4.65	8.55
<i>V</i>	1.122	1.31	3.03	4.89	8.81
<i>D</i>	1.033	1.08	1.32	1.99	2.25

Tableau 5.3 : Facteurs de forme moyen de différents niveaux de résolution et de différentes directions obtenus avec l'ondelette de Daubechies

Il apparaît également que d'un niveau à un autre, l'amplitude des variations des facteurs de forme est assez faible lorsque l'ondelette de Daubechies est utilisée. Ce qui affecte le pouvoir discriminant du facteur de forme.

Ce résultat confirme les conclusions de l'étude de Unser [Uns.95] précédemment énoncées, et nous conforte dans la position que l'ondelette de Haar est la mieux indiquée pour la mise en oeuvre pratique de notre approche.

En conclusion, nous dirons que ces différentes expérimentations ont permis de s'assurer de la capacité du facteur de forme à traduire en termes quantitatifs, les variations de niveaux de gris de la texture suivant différentes orientations. L'étape suivante a pour objectif de tester la pertinence de cet attribut quant à la discrimination des textures.

5.4 Algorithmes de classification de textures

Le problème de classification de textures se pose dans les mêmes termes que celui de la classification classique qui implique un niveau d'extraction de caractéristiques et un niveau d'identification.

Un nombre suffisant d'échantillons représentatifs de chaque classe est nécessaire. Ces échantillons constituent l'ensemble d'apprentissage. Leur analyse permet de définir au cours de la phase d'apprentissage un ensemble de règles qui permettront d'affecter à une classe bien précise, un individu test donné appartenant à une classe a priori inconnue.

Dans le cas de la classification de textures, les individus test peuvent être des pixels ou une sous-image de taille fixe. Ils sont en général associés à une réalisation d'un vecteur attribut qui peut être représenté par un point dans l'espace des attributs. Une mesure de distance effectuée dans cet espace peut être utilisée pour affecter chaque individu à la classe la "plus proche".

Plusieurs types de distances peuvent être utilisées. Les plus courantes sont : la distance de Mahalanobis qui est une mesure de similarité pratique si certaines propriétés statistiques de la texture sont connues, et la distance euclidienne qui conduit à un critère de type géométrique.

La distance euclidienne, facile à mettre en oeuvre, a été utilisée dans les deux algorithmes de classification de textures que nous proposons. Ces algorithmes qui utilisent le facteur de forme comme attribut, diffèrent par leur principe de mise en oeuvre.

5.4.1 Premier Algorithme de classification

- ***Principe***

Les textures diffèrent par la taille et l'orientation de leurs éléments, ainsi que par leurs variations de tons de gris. Leur caractérisation à différents niveaux de résolution peut permettre de les différencier. Le premier algorithme met à profit la représentation multirésolution des données issues de la décomposition en ondelettes, pour caractériser les différentes textures [Lou.95].

La TOD permet de réaliser à la fois, une analyse locale (aux faibles niveaux de résolutions), et une analyse globale (aux niveaux plus élevés). Cette double analyse permet

pour la première, de caractériser les éléments de texture (les textons), et pour la seconde, de décrire le contexte spatial de ces éléments.

Pour caractériser la texture, nous proposons, après la décomposition en ondelettes, de déterminer les trois facteurs de forme à chaque niveau de résolution. L'ensemble d'attributs obtenu, traduit à la fois les propriétés locales et globales de la texture, et sera utilisé pour la discrimination des différentes textures.

La mise en oeuvre pratique de cet algorithme revient à supposer que chaque pixel du niveau J de la décomposition, définit une région d'intérêt (un support) de taille $2^J \times 2^J$ sur l'image originale. Ce qui conduit à subdiviser de façon fictive, l'image originale en sous-images u de taille $2^J \times 2^J$.

Ainsi, on considère les résultats de la décomposition au niveau J de chaque sous-image u , et on calcule aux différents niveaux de la décomposition, les facteurs de forme qui lui correspondent. On détermine ensuite pour chaque niveau de résolution et pour chacune des trois directions, le facteur de forme moyen afin de constituer le vecteur attribut ${}^u\overline{FF}$, caractéristique de la sous-image texturée u :

$${}^u\overline{FF} = ({}^u\overline{FF}_1^1, {}^u\overline{FF}_1^2, {}^u\overline{FF}_1^3, {}^u\overline{FF}_2^1, {}^u\overline{FF}_2^2, {}^u\overline{FF}_2^3, \dots, {}^u\overline{FF}_J^1, {}^u\overline{FF}_J^2, {}^u\overline{FF}_J^3)^T$$

où T désigne la transposée, et ${}^u\overline{FF}_j^i$, le facteur de forme moyen de la sous-image u correspondant au niveau j et à la direction i .

- **Algorithme de classification**

L'algorithme se déroule en deux phases :

Phase d'apprentissage

1. Choisir le niveau J de la décomposition.
2. Soit une région formée par R sous-images adjacentes de taille $2^J \times 2^J$ représentatives d'une même texture t . Réaliser la décomposition en ondelettes des R sous-images.
3. Former les vecteurs attributs ${}^r \overline{FF}$ des R sous-images.
4. Calculer le vecteur attribut moyen ${}^t \overline{FF}$ de la texture t à partir des ${}^r \overline{FF}$.
5. Répéter le processus pour toutes les images de textures t de la base de données.

Phase de classification

1. Décomposer par ondelettes au niveau J , l'image de texture x en sous-images adjacentes de taille $2^J \times 2^J$.
2. Pour chaque sous-image u , calculer le vecteur attribut ${}^u \overline{FF}$.
3. Calculer la fonction discriminante d_t pour chaque vecteur moyen ${}^u \overline{FF}$ déterminé lors de la phase précédente :

$$d_t = d({}^t \overline{FF}, {}^u \overline{FF})$$
4. Affecter la sous-image u , à la même classe que la texture v si :

$$v = \arg \min_t (d_t)$$
5. Répéter le processus pour toutes les textures x de la base de données.

Rappelons que la fonction discriminante utilisée dans nos expérimentations est la distance euclidienne.

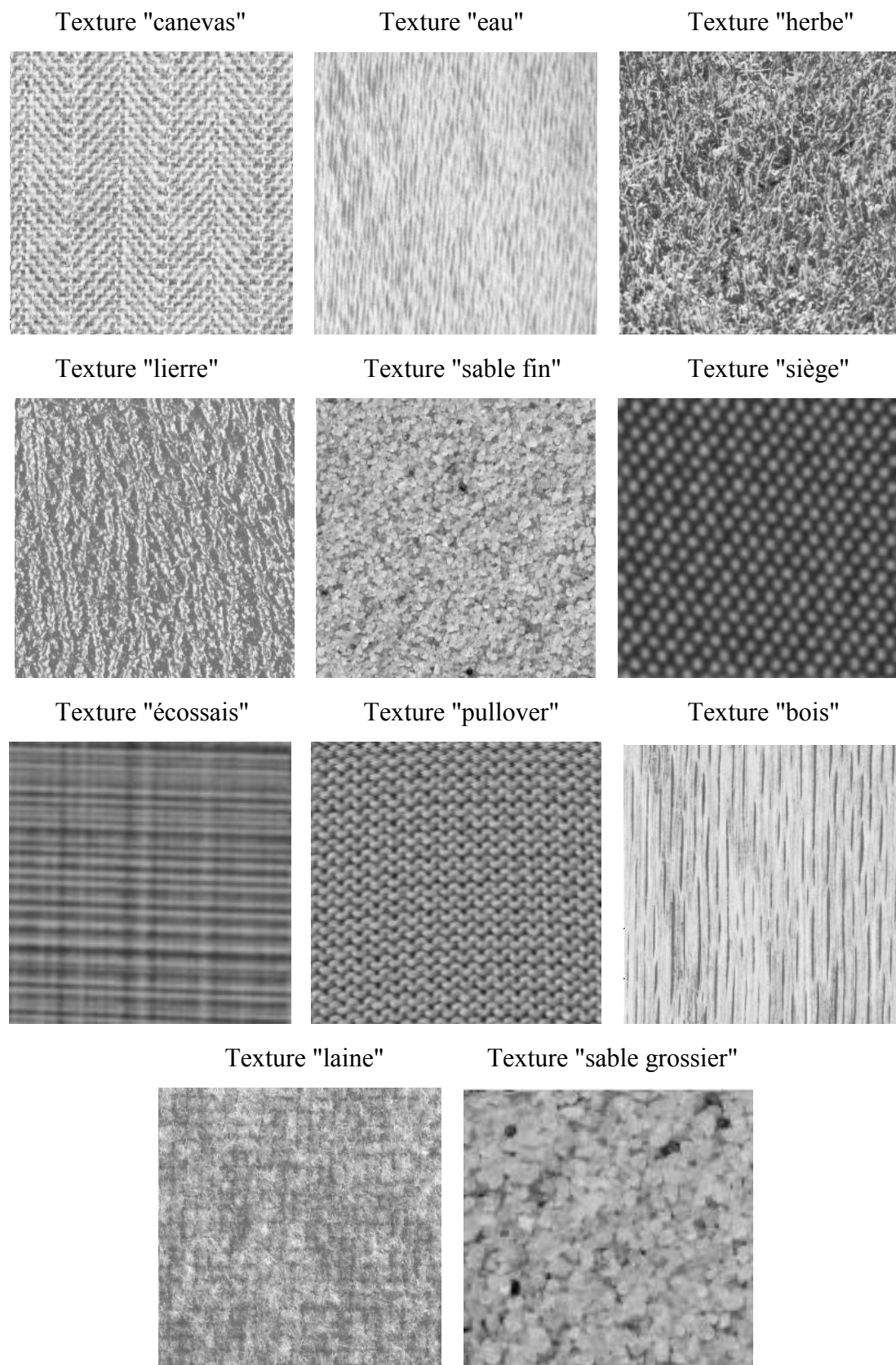


Figure 5.6 : Texture naturelles utilisées dans nos expérimentations

- **Expérimentations**

Les expérimentations ont été conduites sur onze textures naturelles extraites de l'album de Brodatz et de la banque d'images du laboratoire. Ces images sont présentées à la *figure 5.6*. Elles sont toutes de taille 256×256 , avec une quantification de 256 niveaux de gris.

Les paramètres caractéristiques de chaque texture ont été extraits au cours de la phase d'apprentissage, sur une fenêtre d'observation de taille $2^{J+1} \times 2^{J+1}$ (J étant le niveau de résolution où la classification est conduite). Ce qui fixe R à 4. La fenêtre a été positionnée arbitrairement sur l'image originale de façon à avoir une bonne représentativité de la texture.

La méthode utilisant le facteur de forme (*MFF*), a été expérimentée pour trois valeurs de J ($J=3, 4$ et 5). Les résultats obtenus et exprimés en %, sont résumés dans le *Tableau 5.4*.

Textures	J=3	J=4	J=5
<i>Canevas</i>	54.8	68.3	82.8
<i>Eau</i>	86.8	100	100
<i>Herbe</i>	60.9	75.0	86.0
<i>Lierre</i>	74.3	92.9	100
<i>Sable fin</i>	30.5	61.7	73.4
<i>Siège</i>	61.5	92.4	96.4
<i>Ecossais</i>	73.6	90.4	98.3
<i>Pullover</i>	61.8	96.1	100
<i>Bois</i>	54.5	78.1	89.0
<i>Laine</i>	54.0	81.2	93.8
<i>Sable grossier</i>	58.2	83.8	93.2

Tableau 5.4 : Résultats de la classification pour différents niveaux de résolutions

- Il apparaît clairement que les résultats vont en s'améliorant lorsque J croît. Ils sont peu satisfaisants pour $J=3$, mais sont par contre excellents pour $J=5$. Ces résultats confirment le fait que la texture est mieux caractérisée sur un voisinage de pixels de taille importante. La détermination du voisinage adapté pour la caractérisation des texture pose le problème du choix de l'échelle d'analyse pertinente. Le second algorithme de classification qui sera présenté plus loin tente d'y apporter une solution.

- Les résultats les moins bons sont obtenus pour les textures "sable fin" et "canevas".

Le résultat de la texture "sable fin", ne peut s'expliquer que par la distribution statistique aléatoire de ses niveaux de gris. Ce type de textures est difficilement caractérisable par la mesure des variations de ses niveaux de gris puisqu'il présente une grande variabilité. On peut noter que la texture "sable grossier", dont la variabilité est moins importante, obtient un score nettement meilleur.

La matrice de confusion du niveau $J=4$ de la texture "sable fin", montre que les pourcentages d'erreur de classification sont essentiellement réparties entre les textures "laine" et "canevas". Le *Tableau 5.5* donne la matrice de confusion de ces trois textures. Les valeurs sont exprimées en % et la classe assignée est indiquée sur la première ligne du tableau. Par exemple, le taux d'erreur de classification des régions de la texture "sable fin" à la texture "laine", est de 20.8%.

	Sable fin	laine	Canevas
Sable fin	61.7	20.8	11.4
laine	11.9	81.2	04.1
canevas	18.3	9.2	68.3

Tableau 5.5 : Matrice de confusion des textures "sable fin", "laine", et "canevas". La première ligne est celle de la classe assignée.

La lecture de ce tableau montre que la grande variabilité qui caractérise la texture "sable fin", a pour effet de dégrader fortement les résultats des deux autres textures. La texture "laine" présente somme toute, un aspect visuel assez proche du "sable fin". Ce qui peut expliquer leur confusion jusqu'à un certain niveau. En effet au niveau 5, cette confusion est beaucoup moins nette (les résultats étant meilleurs). Ce fait démontre à nouveau, l'intérêt de réaliser l'analyse sur plusieurs niveaux de résolution.

Quant à la texture "canevas", sa confusion avec la texture "sable fin" provient probablement du nombre insuffisant de directions préférentielles considérées dans la méthode proposée. Un regard sur cette texture montre que la prise en compte de la direction d'angle 135° peut aider à mieux la caractériser.

Il est d'ailleurs intéressant de noter que la méthode proposée fournit de bons résultats pour des textures ("eau", pullover", "bois", par exemple) dont les orientations dominantes (verticale ou horizontale) sont privilégiées par le calcul de la TOD.

Pour éprouver l'efficacité de notre algorithme par rapport à l'information extraite, nous avons choisi de comparer les résultats de la *MFF*, avec ceux émanant de deux méthodes différentes à savoir :

- la transformation linéaire locale réalisée par la Transformée en Cosinus Discrète (DCT) qui est une méthode monorésolution,
- et, la méthode classique fondée sur la TOD, qui utilise un voisinage plan pour la détermination des caractéristiques. Nous la noterons TODC.

• ***Mise en oeuvre pratique de la DCT***

Les filtres 2D associés à la DCT, sont déterminés à partir des vecteurs de base suivants: $U1=\{1,1,1\}$, $U2=\{1,0,-1\}$, $U3=\{1,-2,1\}$. Ces filtres sont de taille 3×3 et peuvent être considérés comme des filtres passe-bande couvrant le plan de Fourier.

Pour rendre les résultats comparables, nous avons implanté la DCT uniquement avec les filtres $U1U2$, $U2U1$, et $U2U2$ de la *Figure 5.7*. Ces trois filtres permettent d'extraire les caractéristiques de la texture suivant les mêmes orientations préférentielles que la TOD à savoir, verticale, horizontale et diagonale.

$$U1U2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad U2U1 = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad U2U2 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Figure 5.7 : Les 3 filtres de la DCT utilisés

Le vecteur d'attributs associé à chaque pixel sera donc de dimension 3. Pour discriminer les textures, une mesure d'énergie locale est réalisée sur une région d'intérêt W de taille 15×15 . Cette taille correspond à peu près, aux dimensions (16×16) du support du

voisinage pyramidal défini pour le niveau de résolution $J=4$. Les résultats de la *MFF* correspondant à $J=4$ serviront donc de bases à la comparaison.

L'énergie locale est obtenue par :

$$E = \frac{1}{\#W} \sum_{(k,l) \in W} |x(i,j)|$$

où $\#W$ désigne le nombre de pixels de la région d'intérêt W .

- ***Mise en oeuvre pratique de la TODC***

La TODC que nous avons implantée, utilise également l'énergie locale comme attribut de textures. Cet attribut n'est calculé que sur les images de détails, afin d'obtenir un vecteur d'attributs de même dimension que celui utilisé dans notre méthode. Ainsi, si la décomposition est itérée jusqu'au niveau J , la région sera associée à un vecteur attribut de dimension $3 \times J$.

Le procédé de calcul de cet attribut est celui utilisé par Unser [Uns.95].:

On définit une région d'intérêt sur l'image originale. Cette région induit à chaque niveau de résolution des voisinages "plan" de tailles différentes. L'énergie locale est calculée à chaque niveau de résolution sur le voisinage correspondant à ce niveau.

Dans notre application, la région d'intérêt a été choisie de taille 16×16 pour pouvoir réaliser une comparaison avec les autres méthodes. Seuls les trois premiers niveaux sont considérés. Ce qui nous donne des voisinages de tailles 8×8 , 4×4 et 2×2 respectivement aux niveaux 1, 2 et 3.

Tous les paramètres des différentes expérimentations effectuées, sont rassemblés dans le *Tableau 5.6*.

<i>Méthodes</i>	<i>Nature de la décomposition</i>	<i>Taille des filtres</i>	<i>Type du voisinage</i>	<i>Région d'intérêt sur l'image d'entrée</i>
MFF	multirésolution	2×2	pyramidal	16×16
TODC	multirésolution	3×3	plan	16×16
DCT	monorésolution	3×3	plan	15×15

(J : Niveau maximal de la décomposition, j : indice de résolution).

Tableau 5.6 : Tableau récapitulatif des paramètres des trois méthodes implantées

- **Résultats des trois méthodes**

Les résultats obtenus pour les trois méthodes sont résumés dans le *Tableau 5.7*. Ils sont exprimés en %.

<i>Textures</i>	MFF	DCT	TODC
<i>Canevas</i>	68.3	78.8	62.5
<i>Eau</i>	100	93.8	93.7
<i>Herbe</i>	75.0	73.2	72.5
<i>Lierre</i>	92.9	73.3	80.0
<i>Sable fin</i>	61.7	70.8	62.5
<i>Siège</i>	92.4	77.7	90.8
<i>Ecossais</i>	90.4	93.8	87.5
<i>Pullover</i>	96.1	92.3	98.4
<i>Bois</i>	78.1	72.1	79.5
<i>Laine</i>	81.2	81.2	84.4
<i>Sable grossier</i>	83.8	81.8	82.2

Tableau 5.7 : Résultats de la classification fournis par les trois méthodes

Le taux moyen de classification correcte des trois méthodes est de 83,63% pour la MFF, 80,79% pour la DCT, et 81,29% pour la TODC.

- On constate que les deux méthodes multirésolution (MFF et TODC) fournissent dans l'ensemble, de meilleurs résultats que la méthode monorésolution (DCT). Cette remarque est particulièrement vraie pour les textures dont les éléments sont de tailles relativement grandes ("siège", "pullover",...). Ces résultats démontrent l'intérêt d'une caractérisation multirésolution de la texture. L'analyse multirésolution en effet, apporte des informations sur la structure locale et le contexte de la texture, qui sont utiles et complémentaires pour sa caractérisation.

- La DCT est par contre plus performante sur la texture "sable fin" qui présente un aspect aléatoire (gradient plus prononcé). Cela s'explique par le fait que le calcul des attributs de la DCT est réalisé sur un voisinage qui prend en compte au même niveau de résolution, un nombre plus important de pixels. La considération d'un nombre plus important de pixels pour la détermination des attributs, a pour effet de réduire la variabilité de ces attributs.
- Notons pour achever cette série d'expérimentations que pour une même région d'intérêt, la MFF fournit des résultats légèrement supérieurs à la DCT.

En guise de conclusion, nous dirons que les résultats obtenus, démontrent la pertinence du facteur de forme comme attribut de textures. Cette pertinence croît avec le niveau de résolution de la décomposition.

Les performances, déjà très appréciables de la méthode proposée, peuvent être améliorés en augmentant la sélectivité angulaire des filtres de la TOD.

Cette augmentation peut être obtenue en itérant la décomposition également sur les images de détails (décomposition par paquets d'ondelettes). Ce qui conduit à une composition des différentes directions préférentielles et à l'obtention d'informations sensibles à différentes orientations. Un critère doit alors être défini pour s'assurer de la nécessité ou non de poursuivre la décomposition.

Une autre possibilité plus appropriée à notre avis, est d'avoir recours à la solution proposée par Vautrot et Bonnet [Vau.95] qui consiste à filtrer avant la décomposition, l'image originale avec un ensemble de filtres gaussiens de différentes orientations.

5.4.2 Deuxième algorithme de classification

- ***Principe***

Le second algorithme peut être vu comme une simplification du premier. Il essaie de remédier à deux faiblesses de l'algorithme précédent à savoir : la redondance d'informations des

composantes du vecteur attribut de texture utilisé (le facteur de forme s'obtenant de façon récursive), et le choix arbitraire du niveau d'arrêt de la décomposition. Son principe de base est d'exploiter les propriétés d'unicité et de complétude de la décomposition en ondelettes orthonormales pour discriminer les textures. La mise en oeuvre de cet algorithme tire son inspiration de l'analyse de Fourier.

On sait que la décomposition en séries de Fourier d'un signal périodique quelconque conduit à une somme infinie d'ondes sinusoïdales de fréquences multiples de la fréquence fondamentale. Les coefficients de Fourier décroissent progressivement pour des rangs d'harmoniques de plus en plus élevés. Ce qui, dans la pratique, justifie le choix d'un rang d'arrêt de la décomposition à partir duquel les coefficients de rangs supérieurs sont négligés.

Ce rang (qui n'est pas le même pour tous les signaux), et les coefficients jugés significatifs, peuvent aider à caractériser le signal. Toutefois, cette caractérisation n'est pas unique. Il est en effet possible de trouver deux signaux périodiques, de caractéristiques différentes pour lesquels on a défini un même rang d'arrêt, et dont les décompositions en séries de Fourier ne diffèrent malheureusement, que par les coefficients jugés non significatifs. Cette possibilité pose le problème du choix du rang d'arrêt si l'on désire caractériser correctement les signaux à l'aide des deux paramètres sus-cités.

Ce problème ne se présente pas dans le cas d'une analyse par ondelettes qui réalise une analyse de Fourier à toutes les échelles [Mey.90]. Cela, parce que l'analyse par ondelettes est complète à chaque niveau de résolution. C'est dire que la décomposition peut être arrêtée à l'un des niveaux sans qu'aucune information soit perdue. Ainsi, deux signaux de caractéristiques différentes, considérés à n'importe quel niveau de la décomposition, diffèrent obligatoirement par leurs coefficients d'ondelettes.

Comme de plus, l'analyse par ondelettes, permet de s'affranchir du problème de périodicité propre à l'analyse de Fourier, la caractérisation de la texture peut être effectuée en utilisant comme attributs, le niveau d'arrêt de la décomposition en ondelettes et les trois facteurs de forme de ce niveau qui sont calculés à partir de l'ensemble des coefficients

d'ondelettes de la décomposition [Lou.96a]. Le niveau d'arrêt fixe d'emblée, la taille du support du voisinage sur lequel la texture est analysée.

- **Définition d'un critère d'arrêt de la décomposition**

Le niveau maximum de la décomposition est implicitement fixé par la taille de l'image. La définition d'un critère d'arrêt de la décomposition est rendue nécessaire si l'on désire ne pas atteindre la pleine décomposition de l'image.

Le critère d'arrêt que nous avons choisi, est un critère de variance appliqué sur les images basses résolutions.

Le choix de ce critère est suggéré par deux éléments :

- Le premier est que si la variance de l'image basse résolution courante est relativement faible (ce qui est le cas si les niveaux de gris de l'image tendent vers leurs valeurs moyennes), la poursuite de la décomposition est vaine puisque les images de détails n'apportent plus d'informations significatives.
- Le second est l'assurance de la convergence vers zéro de ce critère. En effet, l'application itérée du filtre passe-bas H associé à l'analyse multirésolution, tend à ne conserver que la fréquence nulle puisque :

$$\lim_{n \rightarrow \infty} \prod_{j=0}^n H(2^j \omega) = 0 \quad \text{si } \omega \in [-\pi; \pi] / \{0\}$$

Les niveaux de gris des images basses résolutions convergent donc vers une valeur constante qui est la moyenne de ceux de l'image originale. Cette convergence implique de fait, celle des variances des images de basse résolution vers zéro.

Nous proposons de définir un seuil C par rapport à la variance de la première image basse résolution I_1 , en considérant qu'en deçà de ce seuil, l'itération de la décomposition doit être arrêtée. Cette condition conduit à la détermination d'un niveau d'arrêt J_a . Ce dernier fixe les limites de la décomposition et correspond au niveau où le vecteur d'attributs de texture devra être calculé.

- **Algorithme de classification**

Les décompositions en ondelettes orthonormales de deux régions textuellement distinctes, sont différentes. L'idée de base de notre algorithme, est de répartir en différentes classes, les textures de la base de données suivant leur valeur du niveau d'arrêt J_a de la décomposition, puis de réaliser la classification à l'intérieur de chaque classe avec comme attribut, le facteur de forme déterminé à ce niveau [Lou.96a].

L'intérêt d'une telle approche est :

- la réduction de la dimension du vecteur d'attributs (3 au lieu de $3 \times J$ pour le premier algorithme), et
- la diminution du taux d'erreur de classification puisque l'algorithme met finalement en compétition dans chaque sous-classe, un nombre réduit de textures.

La mise en oeuvre de cet algorithme se déroule comme suit : les images de la base de données sont décomposées par ondelettes. L'itération est poursuivie si le critère d'arrêt qui a été défini n'est pas respecté. A terme, le niveau d'arrêt J_a est déterminé et les trois facteurs de forme sont calculés à ce niveau afin de former le vecteur attribut de texture. La discrimination des textures est réalisée par la suite en effectuant une mesure de distance dans l'espace d'attributs de texture.

L'algorithme comporte donc deux phases : une phase de préclassification qui conduit à l'obtention de Q classes, et une phase de classification qui se déroule dans chaque classe constituée. L'évolution de ces deux phases est décrite ci-dessous :

Phase de préclassification

1. Réaliser la décomposition en ondelettes de la texture t de la banque d'images, en calculant à chaque niveau j , la variance $Var(j)$ de l'image basse résolution.
2. Arrêter la décomposition si : $Var(j) < C \hat{Var}(1)$ et en déduire la valeur du niveau d'arrêt J_a par : $J_a = j - 1$.
3. Affecter la texture t à la classe correspondant à la valeur de J_a . Répéter le processus pour toutes les images de texture de la base de données.

Phase de classification finale

Pour une classe q donnée de niveau de résolution d'arrêt J_a ,

Phase d'apprentissage

1. Soit une région formée par R sous-images adjacentes de taille $M \times M$ représentatives d'une texture t donnée de la classe q .
 - 1.1 Réaliser la décomposition en ondelettes au niveau de résolution J_a des R sous-images.
 - 1.2 Former les vecteurs attributs ${}^rFF_{Ja} = (FF^1_{Ja}, FF^2_{Ja}, FF^3_{Ja})$ des R sous-images.
2. Calculer le vecteur attribut moyen ${}^tFF_{Ja}$ de la texture t à partir des ${}^rFF_{Ja}$.
3. Répéter le processus pour toutes les textures de la classe q .

Phase de classification

1. Soit une texture inconnue x de la classe q
 Pour chaque sous-image u de taille $2^{J_a} \times 2^{J_a}$ de la texture x ,
 - 1.1 Calculer le vecteur attribut ${}^uFF_{Ja}$.
 - 1.2 Calculer la fonction discriminante pour chaque texture t de la classe q :

$$d_t = ({}^uFF_{Ja}, {}^tFF_{Ja})$$
2. Affecter la sous-image u , à la même sous-classe que la texture v de la classe q si :

$$v = \arg \min_t (d_t)$$

Répéter le processus pour les Q classes à partir de la phase d'apprentissage.

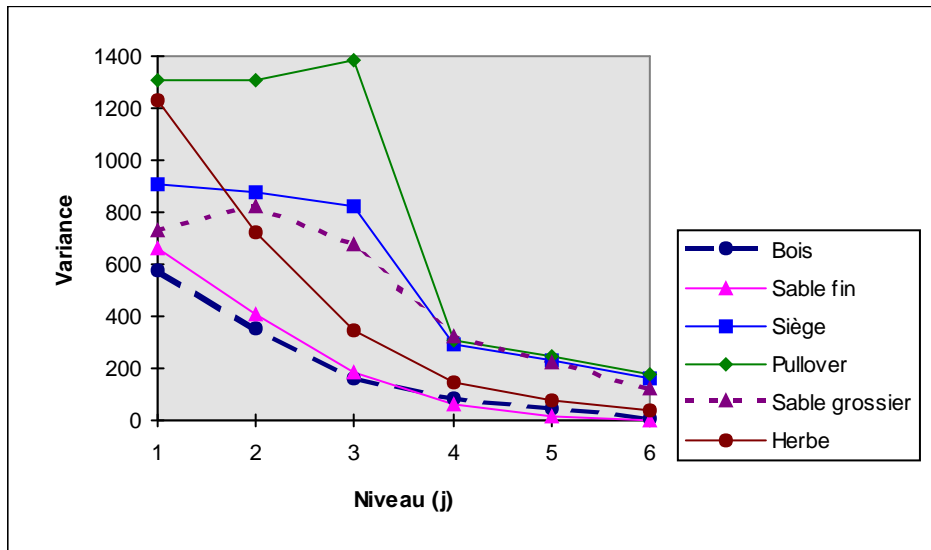
• *Expérimentations et résultats*

Les expérimentations ont été conduites sur les onze images précédentes de textures naturelles, pour lesquelles l'évolution de la variance des images basses résolutions de la décomposition a été observée. Les valeurs obtenues sont résumées dans le *Tableau 5.8*.

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
Bois	577.6	352.4	158.3	81.7	42.9	8.5
Canevas	802.5	395.4	85.9	25.2	10.4	5.2
Eau	296.	170.8	96.1	50.9	26.9	16.4
Herbe	1233	719.3	344.1	145.5	76.4	35.9
Lierre	897.7	466.6	175.7	97.9	14.3	3.5
Laine	499.1	358.3	218.2	122.4	58.2	29
Sable fin	657.8	410.2	183	59.6	16	1.2
Siège	905.3	877.8	821.6	289.6	231.5	163.5
Pullover	1310	1306	1386	307.5	247.3	174.1
Sable grossier	732	823	675.4	321.1	224.6	119.4
Ecossais	1369	1296	1054.3	314.7	154.51	108.68

Tableau 5.8 : Variance des images basses résolutions des textures test

Seulement quelques unes ont été représentées sur le graphique 5.1.



Graphique 5.1 : Courbes d'évolution de la variance des images basses résolutions de quelques textures.

Au regard de ce graphique, il ressort que la courbe d'évolution des variances de chaque texture lui est propre. Chaque courbe connaît à un niveau quelconque, une décroissance vers 0 dont la rapidité semble être liée aux variations de niveaux de gris de la texture. Cette décroissance est nettement marquée dans le cas de la texture "herbe" quand on la compare à la texture "sable", ou mieux à la texture "bois".

On note également une différence dans l'évolution des variances des deux textures de sable : les variances de la texture la plus grossière évoluent plus lentement que celles de la texture fine. D'autre part, les valeurs obtenues pour $j=1$, dépendent du contraste de l'image

texturée. Plus le contraste est fort, (cas de la texture "siège"), plus la variance à ce niveau est élevée.

Comme nous l'avions espéré, toutes les courbes convergent vers 0, et pour chaque texture, il existe un niveau au-delà duquel la variance de l'image de basse résolution devient relativement faible. Ce niveau varie d'une texture à une autre. Il peut de ce fait, être effectivement utilisé pour caractériser les différentes textures.

Il est également intéressant d'attirer l'attention du lecteur sur la variation brutale que connaît, à un niveau donné de la décomposition, l'évolution de la courbe des textures "siège", "pullover", et "écossais", et à un degré moindre, celle de la texture "sable grossier". A l'observation, cette variation ne se produit que pour les textures qui ont un aspect assez régulier, voire périodique.

Cela donne à penser que la descente brutale des courbes, s'effectue au niveau de résolution pour lequel la taille du voisinage considéré correspond à celle des éléments de la texture ou à celle de sa structure. Si tel est le cas, ce niveau particulier devrait constituer le niveau minimum pour une analyse convenable de la texture.

Pour s'en assurer, nous avons entrepris des expérimentations complémentaires sur deux images de synthèse (*Texture 1* et *Texture 2* de la *Figure 5.8*).

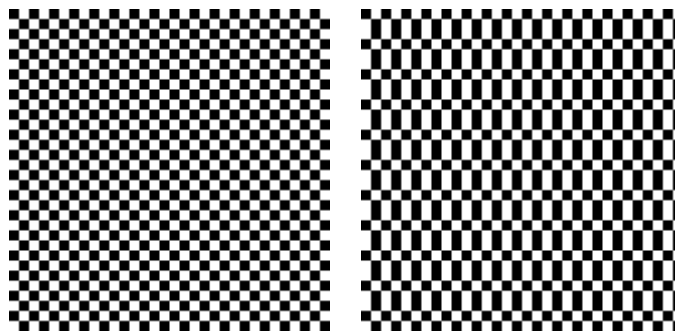


Figure 5.8 : Images de textures synthétiques

En faisant varier la taille des éléments de chaque texture, nous avons observé l'évolution des variances de leurs images basses résolutions. Les résultats obtenus sont résumés dans le *Tableau 5.9*.

Taille des éléments de textures	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
<i>Texture 1</i>						
$8\hat{1} 8$	16256	16256	16256	0	0	0
$9\hat{1} 9$	16894	8591	2750	3	1.9	1.2
$10\hat{1} 10$	16259	10322	4204	42.2	14.1	0.18
$12\hat{1} 12$	16256	16256	7001	484	0	0
$16\hat{1} 16$	16256	16256	16256	16256	0	0
<i>Texture 2</i>						
$4\hat{1} 4+4\hat{1} 8$	16256	16256	16256	0	0	0
$4\hat{1} 4+4\hat{1} 12$	16254	16254	10666	878.4	98.4	19.4

Tableau 5.9 : *Variance des images basses résolutions des textures synthétiques*

Ces résultats confirment notre supposition initiale. Lorsque la taille des éléments de la "Texture 1" correspond exactement à celle du voisinage induit par le niveau, on observe une chute brutale de la variance lors du passage au niveau immédiatement supérieur. C'est le cas par exemple, lorsque les éléments sont de taille $8\hat{1} 8$ et que l'on se trouve au niveau $j=3$. Par contre la variation, même si elle demeure perceptible, est moins marquée lorsque l'écart entre la taille du motif et celle du voisinage s'agrandit.

Un constat identique peut être fait en ce qui concerne la "Texture 2", sauf que la brusque variation se produit cette fois-ci, lorsque la taille de la structure de la texture et celle du voisinage coïncident : La "Texture 2" est constituée pour le premier cas, de deux types d'éléments de tailles respectives $4\hat{1} 4$ et $4\hat{1} 8$. Mais sa structure élémentaire est de taille $8\hat{1} 8$ (elle est répétée en ligne, et d'une ligne à une autre, elle subit une rotation de $\pi/2$). On peut remarquer que le passage brutal de la variance à 0 se produit après le niveau $j=3$, qui induit un voisinage de taille $8\hat{1} 8$. Cette évolution est moins marquée dans le second cas où la taille de la structure ($8\hat{1} 10$) a été fixée de sorte qu'elle ne corresponde à aucun voisinage défini par les différentes valeurs de j .

En définitive, dans le cas idéal où la variance passe brutalement à zéro, le niveau de résolution qui précède son annulation, correspond au niveau optimal d'analyse de la texture. Il est évident qu'une analyse à un niveau inférieur à ce dernier, s'effectuerait sur un voisinage trop petit pour une description correcte de la texture, et qu'une analyse entreprise à un niveau supérieur, n'apporterait aucune information supplémentaire.

Ces expérimentations montrent qu'il est possible de déterminer la taille de la structure de la texture, ou au minimum, celle de ses éléments. Elles suggèrent l'élaboration d'une méthode structurale d'analyse de texture par ondelettes. On peut en effet, envisager de déterminer la structure de ces textures par une action sur les deux paramètres que nous avons évoqués lors de la description de notre approche, à savoir, le niveau de résolution et l'ondelette analysante. L'idée est de régler graduellement la taille du voisinage induit par l'analyse, par une action conjointe sur l'ondelette analysante et le niveau de la décomposition, afin d'ajuster cette taille à celle de la structure de la texture. Cela nécessite l'utilisation d'un jeu d'ondelettes qu'il reste à déterminer.

La classification des onze textures de l'expérience précédente a été réalisée. La valeur de C a été prise égale à $0,1$. En pratique, on est limité par la taille des images qui dans notre cas, est de 256×256 . Par conséquent, la décomposition a été volontairement arrêtée au niveau 5, lorsque le critère défini précédemment n'avait pas été satisfait.

Les résultats obtenus sont rapportés dans le *tableau 5.10* où P_c (exprimé en %) correspond au pourcentage de classification correcte.

On peut se rendre compte qu'au bénéfice de la classification primaire des éléments de la base de données, le résultat de la classification en général est satisfaisant. Par exemple, le score réalisé au niveau 3 pour des textures comme "canevas" et "sable fin", est nettement meilleur par rapport à celui réalisé avec le premier algorithme.

On peut également noter qu'au niveau 5, les résultats de la classification sont satisfaisants, bien que la décomposition ait été délibérément arrêtée. Il est donc possible, lorsque le voisinage induit par le niveau de la décomposition est de taille suffisamment importante, de caractériser correctement la texture par ses coefficients d'ondelette.

Textures	Niveau d'arrêt J_a	Méthode proposée P_c (%)
<i>Bois</i>	4	79.8
<i>Canevas</i>	3	94.1
<i>Eau</i>	4	93.3
<i>Herbe</i>	4	94
<i>Lierre</i>	4	97.2
<i>Laine</i>	5	100
<i>Sable fin</i>	3	90.8
<i>Siège</i>	5	99.1
<i>Pullover</i>	5	100
<i>Sable grossier</i>	5	98.5
<i>Ecossais</i>	5	100

Tableau 5.10: Résultats du second algorithme de la classification

Il faut dire pour terminer que par son principe de mise en oeuvre, cet algorithme ne peut donner sa pleine mesure que pour des images de tailles relativement grandes (images satellitaires par exemple). Notons d'autre part que cet algorithme peut constituer un point de départ pour l'élaboration d'une analyse structurale de la texture.

5.5 Conclusion

Nous avons présenté une nouvelle approche de caractérisation de la texture fondée sur une analyse par ondelettes. L'ensemble des coefficients d'ondelettes de la décomposition est vu comme décrivant, suivant différentes directions, les variations locales des niveaux de gris de la texture autour de leur valeur moyenne. Cette interprétation nous a permis d'introduire la notion de voisinage pyramidal et de définir un nouvel attribut de texture, appelé *facteur de forme*.

Cet attribut qui permet de caractériser les fluctuations des niveaux de gris de la texture, présente la particularité d'être défini sur le voisinage pyramidal introduit, dont les dimensions varient automatiquement avec le niveau de résolution considéré.

L'emploi du facteur de forme dans deux algorithmes de classification de textures différents, a permis de démontrer qu'il est suffisamment pertinent pour discriminer différentes textures.

Le premier algorithme présente un intérêt pratique lorsque les images sont de tailles relativement petites. Le second, qui utilise un vecteur d'attributs de dimension plus faible que le premier, est par contre indiqué lorsque les images sont de plus grandes dimensions. Dans ce cas, la décomposition en ondelettes peut être conduite à des niveaux élevés, permettant ainsi de déterminer avec une plus grande précision, le niveau d'arrêt de chaque texture.

L'étape logique suivante est d'envisager à présent, l'utilisation du facteur de forme dans un algorithme de segmentation de textures. Les performances d'un tel algorithme sont fortement conditionnées par la pertinence des attributs utilisés, mais aussi par la taille des voisinages sur lesquels ils sont calculés.

Dans le chapitre suivant, la pertinence du facteur de forme et la taille variable de son voisinage de définition, sont avantageusement mises à profit dans un processus de segmentation pyramidale de textures.

6

Segmentation pyramidale de textures

La segmentation d'une image consiste en son partitionnement en différentes régions qui présentent au sens d'un critère donné (visuel, mesure d'un attribut de texture) une certaine homogénéité. Cette opération est sans aucun doute, l'une des phases les plus délicates de la procédure d'analyse d'images. Elle conditionne en effet, les performances des processus de traitement des niveaux supérieurs (interprétation des scènes, reconnaissance des formes,...).

La segmentation par analyse de textures est un processus qui repose essentiellement sur la détermination d'un ensemble d'attributs qui peuvent efficacement caractériser les différentes textures présentes dans l'image à segmenter.

De nombreux algorithmes de segmentation ont été élaborés en employant les techniques de caractérisation de textures présentées au chapitre 3.

Certains auteurs [Bov.91, Jai.91] ont exploité les résultats des travaux de recherche récents sur la vision humaine, pour proposer des méthodes de segmentation fondées sur une approche multicanaux. La décomposition multicanaux met en oeuvre un filtrage spatio-fréquentiel et angulaire, qui permet d'appréhender les caractéristiques de la texture suivant différentes orientations et pour différentes fréquences.

D'autres auteurs [Pie.81, Uns.89] avant eux, ont suggéré que la discrimination de textures variait selon l'échelle à laquelle l'image est représentée, et ont de ce fait recouru à une représentation multirésolution des données pour réaliser la segmentation.

Les ondelettes réalisent non seulement une décomposition multicanaux, mais permettent également d'accéder à une représentation multirésolution. Ce double avantage suggère l'élaboration d'une méthode de segmentation de textures fondée sur une analyse par ondelettes.

Nous présentons dans ce chapitre la méthode que nous avons développée. Elle exploite la représentation pyramidale que génère la décomposition en ondelettes pour réaliser une segmentation de l'image suivant une stratégie du type du "plus grossier au plus fin" ("coarse to fine"). Elle nécessite la connaissance du nombre de régions distinctes présentes dans l'image, et tire son originalité du voisinage pyramidal utilisé pour la détermination des attributs. Ce voisinage semble être plus approprié lorsque les données sont représentées à différents niveaux de résolution.

6.1 Techniques de segmentation fondées sur une approche par ondelettes

Les méthodes de segmentation qui se fondent sur une approche par ondelettes, mettent à profit les résultats de la décomposition pour procéder à la segmentation de l'image.

La *figure 6.1* schématise le principe de ces méthodes. Une opération non linéaire est nécessaire à la sortie des filtres mis en oeuvre par la TOD, afin de discriminer efficacement les différentes textures. Elle précède une phase de calcul du vecteur d'attributs de texture. Les attributs obtenus, permettent d'achever le processus de segmentation en regroupant dans une même classe, les pixels jugés similaires.

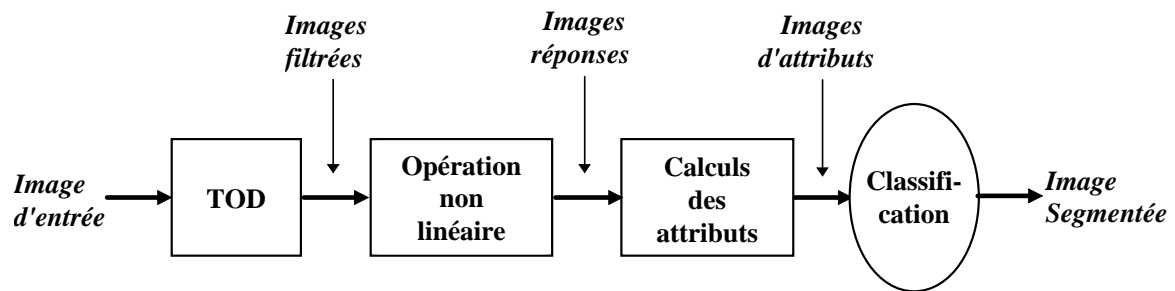


Figure 6.1 : Schéma de principe d'une segmentation de textures fondée sur une analyse par ondelettes.

Les diverses méthodes élaborées sur ce modèle, diffèrent par :

- les bancs de filtres utilisés (filtres de Daubechies, de Battle-Lemarié, ou filtres associés aux ondelettes B-splines),
- le type de décomposition en ondelettes (classique [Sal.95], biorthogonale [Vau.95], paquets d'ondelettes [Cha.92]),
- les attributs choisis,
- le type d'algorithme de classification choisi pour regrouper les pixels d'une même région,
- la stratégie de segmentation adoptée.

Certaines de ces méthodes [Cha.92, Sal.95] adoptent une analyse graduelle de l'image suivant une stratégie du type "du plus grossier au plus fin". Cette stratégie est suggérée par la représentation pyramidale des données issues de la décomposition. Elle conduit à une segmentation pyramidale de l'image : la segmentation est réalisée à chaque niveau de résolution, et est conduite du sommet vers la base de la pyramide. Les résultats sont progressivement corrigés, affinés et propagés au niveau inférieur, au fur et à mesure de la descente de la pyramide.

L'intérêt d'une telle approche peut se résumer en trois points essentiels.

- Premièrement, l'analyse réalisée à différents niveaux de résolution permet d'exploiter à chaque niveau, les différences en tailles et en orientations des éléments de textures, pour une discrimination correcte des régions de l'image;
- Deuxièmement, l'initialisation du processus de segmentation à partir des basses résolutions, permet de prendre en compte l'information contextuelle de l'image analysée, et de ce fait, assure une meilleure classification des différentes textures. Sa conduite progressive vers les

hautes résolutions, permet en général d'affiner la détermination des contours des différentes régions.

- Enfin, la représentation pyramidale des données améliore sensiblement les performances en temps de calcul du processus de segmentation.

L'algorithme de segmentation que nous proposons, procède suivant la stratégie "du plus grossier au plus fin". Il diffère des algorithmes déjà proposés qui lui sont similaires [Cha.92, Sal.95], par le procédé de classification des pixels, et surtout par le mode de caractérisation des images de la décomposition.

L'attribut utilisé à savoir le facteur de forme, est déterminé sur un voisinage pyramidal, dont l'intérêt apparaît nettement dans les problèmes de segmentation. La segmentation est une opération délicate car elle nécessite, à la différence de la classification, une localisation précise des contours des différentes régions.

6.2 Intérêt d'un voisinage pyramidal pour une segmentation multirésolution d'une image texturée

Les techniques de segmentation de textures analysent l'espace des attributs dans le but d'extraire les régions homogènes de l'image. Comme dans le cas monorésolution classique, les techniques qui utilisent une représentation multirésolution des données, déterminent les attributs de textures à chaque niveau de résolution sur un voisinage "plan" de taille $n \times n$ (avec n fixe ou variable).

Le regard critique que l'on peut jeter sur cette démarche, concerne la taille que ce voisinage représente sur l'image originale. Si l'on suppose en effet, qu'un tel voisinage est défini au niveau j de la décomposition, il correspond alors sur l'image d'entrée, à un voisinage de taille $2^j n \times 2^j n$. Cette taille prend des dimensions impressionnantes pour des valeurs de j et de n élevées. Par exemple, si $j=4$ et $n=5$, la taille du voisinage correspondante est de 80×80 pixels sur l'image originale!!!

Il est certain qu'un voisinage de cette dimension devient gênant quand il s'agit de localiser exactement les frontières entre les différentes textures, comme c'est le cas dans les problèmes de segmentation d'images. Il introduit forcément une ambiguïté sur la localisation exacte des frontières, et affecte les performances en temps de calcul du système d'analyse.

Pour preuve, Chang et Kuo [Cha.92], ainsi que Salari et Ling [Sal.95] qui se sont trouvés confrontés à ce problème, ont été contraints de remettre en cause les résultats de la classification effectuée à un niveau donné, en reconsidérant au niveau inférieur, la classification de la totalité ou celle d'un grand nombre de pixels. Cette opération redondante, est rendue en effet nécessaire, pour délimiter avec précision les contours des différentes régions.

Sans remettre en cause le bien-fondé de la prise en compte d'un voisinage pour la détermination des attributs, on peut se demander s'il est opportun d'utiliser dans le cas multirésolution, le même voisinage plan que celui du cas monorésolutionnel. Rappelons que la notion de voisinage a été introduite dans ce dernier cas, pour prendre en compte les informations contextuelles véhiculées par l'image. Or l'analyse multirésolution que réalise entre autres la TOD, intègre déjà ces informations à chaque niveau de résolution.

Il nous semble plus convenable de calculer les attributs sur un voisinage adapté à ce type de représentation. Le voisinage "pyramidal" que nous avons introduit au chapitre précédent, peut constituer une alternative intéressante au voisinage "plan", d'autant plus que les attributs calculés sur ce type de voisinage se sont révélés être aussi pertinents que ceux déterminés sur le voisinage "plan".

Le voisinage pyramidal tel qu'il a été défini, présente un grand intérêt pratique pour une opération de segmentation d'images. Le calcul de l'attribut sur un voisinage pyramidal de support de taille $2^j \times 2^j$, ne fait intervenir qu'un et un seul pixel du niveau de résolution j (les autres pixels étant pris sur des niveaux inférieurs). Il n'en est pas de même lorsque l'attribut est déterminé sur un voisinage plan. Dans ce cas, indépendamment du niveau considéré un groupe de pixels adjacents du même niveau de résolution (j par exemple) interviennent dans la détermination de l'attribut à ce niveau (j).

Cette différence fondamentale, se traduit, dans le cas d'un voisinage pyramidal, par l'induction sur l'image originale, d'un voisinage de taille beaucoup plus réduite (*figure 6.2 [b]*) que celle induite par l'utilisation du voisinage plan (*figure 6.2 [a]*).

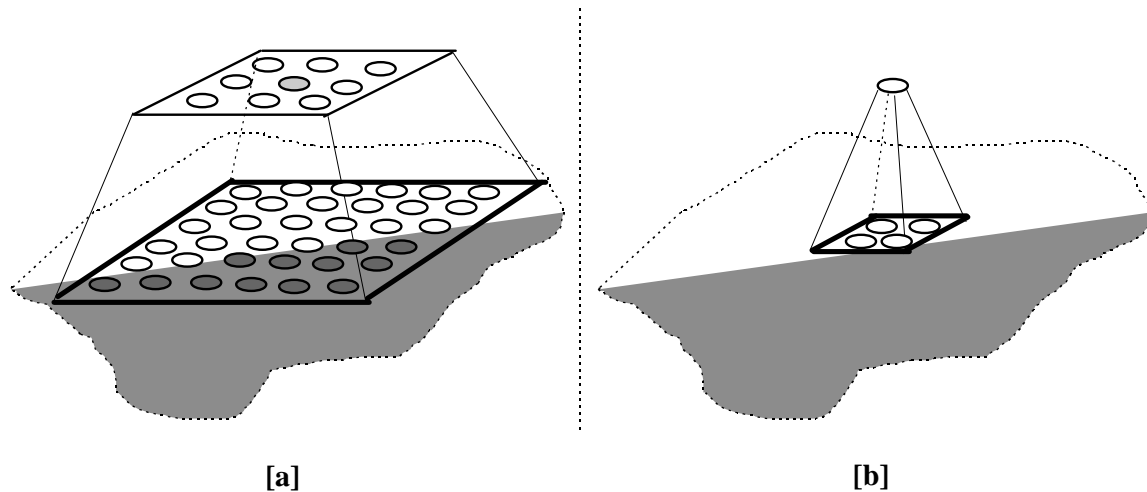


Figure 6.2 : Exemples de support d'un voisinage plan [a] et d'un voisinage pyramidal [b]

La conséquence immédiate de l'utilisation d'un voisinage pyramidal est donc la réduction du nombre de pixels représentant de zones inter-régions c'est-à-dire de zones frontières. Cette réduction permet d'améliorer les performances en temps de calcul du processus de segmentation.

D'autre part, le passage d'un niveau de résolution donné, à celui qui lui est immédiatement inférieur, se traduit par une focalisation de l'analyse sur une région de plus petite taille. Cette focalisation graduelle de l'analyse qui s'opère lors de la descente de la pyramide, permet de localiser avec une précision grandissante, les contours des différentes régions.

L'algorithme de segmentation pyramidale que nous proposons, met à profit les différents avantages rattachés à l'analyse par ondelettes et à l'utilisation d'un voisinage pyramidal.

6.3 Segmentation pyramidale de textures

6.3.1 Principe de mise en oeuvre

En remarquant que les attributs calculés sur un voisinage pyramidal, sont d'autant pertinents qu'ils sont obtenus à des niveaux de résolution élevés, il est alors naturel de démarrer le processus de segmentation au sommet de la pyramide d'images que génère la décomposition en ondelettes.

Ce processus est fondé sur la stratégie du "plus grossier au plus fin". L'approche que nous avons employée pour la segmentation, nécessite la connaissance a priori du nombre de textures présentes dans l'image à segmenter [Lou.96b].

Les trois facteurs de forme directionnnels sont déterminés à chaque niveau de la décomposition. Ceux qui correspondent au niveau le plus élevé de la décomposition, vont constituer les points d'entrée d'un algorithme de classification floue qui permet de déterminer les "meilleurs" prototypes de chaque classe.

Ces prototypes sont utilisés pour réaliser une classification des pixels du niveau courant par une mesure de distance dans l'espace des attributs de textures. La distance employée est la distance euclidienne. Au terme de cette étape, une image partitionnée en un nombre fini de classes est obtenue.

Une analyse de voisinage effectuée sur cette image, permet de déterminer un ensemble de pixels isolés ou frontaliers que nous regroupons sous l'expression de pixels ambigus. La classification de ces pixels est reportée à la résolution supérieure, où de nouveaux prototypes sont déterminés à partir des précédents.

Le résultat intermédiaire de la segmentation est interpolé pour être ramenée à la nouvelle résolution, et le processus est itéré jusqu'à ce que la résolution initiale soit atteinte.

6.3.2 Description de l'algorithme de segmentation

L'algorithme de segmentation que nous proposons [Lou.96b], s'articule autour de trois points essentiels :

- la détermination des prototypes,
- la segmentation primaire,
- et la localisation précise des contours.

- **Détermination des prototypes**

Cette phase du traitement consiste à déterminer les "meilleurs" prototypes de chaque classe. Elle se déroule au niveau le plus élevé de la pyramide de décomposition et requiert la connaissance du nombre c de classes (nombre de textures présentes dans l'image).

C'est une phase importante de l'algorithme puisque le choix de ces prototypes conditionne les performances de la segmentation. Ce choix est d'autant important que ni les prototypes déterminés, ni la grande majorité des résultats intermédiaires de la segmentation, ne sont remis en cause tout au long du déroulement du processus. Par conséquent, il est impératif de déterminer les "meilleurs" prototypes de chaque classe, afin d'éviter une propagation des erreurs éventuelles.

Pour pouvoir quantifier ce critère subjectif, nous avons recouru à une pré-classification de type flou. Ce type de classification associe à chaque pixel, un degré d'appartenance par rapport à l'ensemble des classes en présence. Cela permet de distinguer des pixels dont l'appartenance à une classe est considérée comme sûre, et ceux pour lesquels cette affirmation n'est pas vérifiée (pixels non sûrs ou pixels flous). Les "meilleurs prototypes de chaque classe peuvent par conséquent être obtenus en ne retenant dans chaque classe constituée, que les pixels dont le degré d'appartenance est élevé.

L'algorithme de classification floue que nous avons utilisé est celui des c -moyennes floues (FCM) [Bez.81]. Cet algorithme procède de façon itérative, avec le nombre c de classes définies dans l'image comme information fixée a priori.

A chaque pixel x_i de l'image de taille $n=N \times N$, est associé un vecteur d'attribut X_i formé des trois facteurs de forme et un vecteur d'appartenance $U_i=(u_i^{<1>}, u_i^{<2>}, \dots, u_i^{<c>})$.

Le vecteur X_i est dans notre cas, de dimension $p=3$ et $u_i^{<k>}$, désigne le degré d'appartenance du pixel x_i à la classe C_k , tel que :

$$u_i^{<k>} \in M=[0,1] \quad \text{et} \quad \sum_{k=1}^c u_i^{<k>} = 1$$

Soit V_k le vecteur d'attribut type de la classe C_k , appelé également le centre de la classe C_k . On définit une distance à partir d'une forme bilinéaire définie positive telle que :

$$(d_i^{<k>})^2 = \|X_i - V_k\|^2 = (X_i - V_k)^T A (X_i - V_k)$$

La partition floue optimum est obtenue en minimisant sous contraintes par rapport à \underline{U} et \underline{V} , le critère suivant :

$$J_m(\underline{U}, \underline{V}) = \sum_{k=1}^c \sum_{i=1}^n (u_i^{<k>})^m (d_i^{<k>})^2$$

où $\underline{U} \in M^{nc}$ et $\underline{V} \in \mathbf{R}^{cp}$, et où m est un facteur de flou supérieur ou égal à 1 (généralement compris entre 1,5 et 2).

Ce problème de minimisation sous contraintes peut se résoudre par la méthode du Lagrangien en recherchant itérativement le minimum de la fonction par rapport à \underline{V} pour \underline{U} fixé, puis par rapport à \underline{U} pour \underline{V} fixé.

A chaque itération, le centre V_k de la classe C_k , est calculée à partir des vecteurs d'attributs X_i des pixels x_i (avec $i=1, \dots, n$) de l'image par :

$$V_k = \frac{\sum_{i=1}^n (u_i^{<k>})^m X_i}{\sum_{i=1}^n (u_i^{<k>})^m}$$

La détermination des centres des c classes permet d'actualiser le vecteur de degré d'appartenance associé au pixel x_i . Les nouvelles composantes de ce vecteur s'obtiennent par :

$$u_i^{<k>} = \frac{1}{\sum_{q=1}^c \left(\frac{d_i^{<k>}}{d_i^{<q>}} \right)^{\frac{2}{m-1}}}$$

L'itération est poursuivie jusqu'à ce que l'on converge vers un minimum.

Dans notre application, la réalisation de la classification sur un nombre réduit de pixels (ceux du sommet de la pyramide) et l'utilisation d'un vecteur attribut de dimension trois, contribuent à une convergence rapide de l'algorithme.

Au terme de la procédure itérative, une règle de décision par coupe de niveaux permet de déterminer les "meilleurs" prototypes de chaque classe.

La règle de décision par coupe de niveaux r s'exprime comme suit :

$$x_i \text{ est prototype de } C_k \text{ si } u_i^{<k>} \geq r \text{ avec } 0.5 < r \leq 1$$

- **Segmentation primaire**

L'algorithme de classification floue, intègre les attributs associés aux pixels jugés non prototypes dans le calcul des centres de classe. Pour éviter tout risque de propagation d'erreurs, on débute la phase de segmentation primaire, en calculant le nouveau centre de chaque classe à partir uniquement des prototypes de la classe. Ce nouveau centre est le seul à être considéré dans la suite.

Dans le seul souci d'initialiser l'algorithme itératif de segmentation, les pixels non prototypes du niveau de résolution courant sont repartis dans les différentes classes en réalisant une mesure de distance dans l'espace des attributs de texture. Les pixels prototypes ne sont pas concernés par cette classification. Ils conservent leurs classes de départ. La

fonction discriminante utilisée pour cette classification est la distance euclidienne. On obtient finalement une image partitionnée en c classes.

La segmentation obtenue est assez grossière et la délimitation des contours imprécise. De plus, les propriétés de connexité des régions, sont ignorées dans la segmentation par classification. Ce qui peut se traduire par une grande sensibilité au bruit c'est-à-dire qu'un pixel d'une classe k , peut se retrouver isolé au milieu d'un ensemble de pixels de même classe q (avec $q \neq k$).

Ainsi, dans le but d'affiner cette segmentation primaire, il est intéressant de rechercher les pixels ambigus, c'est-à-dire les pixels frontaliers ou isolés. Pour cela, une analyse du contexte spatial de chaque pixel de l'image segmentée est nécessaire.

Cette analyse conduit à terme, à la constitution d'une nouvelle classe de pixels. Un pixel est considéré comme ambigu, s'il n'a pas la même étiquette que la majorité des pixels de son voisinage. La classification de ces pixels ambigus est reportée à un niveau de résolution inférieur.

Ce report permet en réalité, d'analyser beaucoup plus finement les régions représentées par les pixels ambigus. En effet, le passage d'un niveau de résolution à celui qui lui est immédiatement inférieur, revient selon la construction pyramidale, à représenter un pixel (un noeud) par ses quatre pixels correspondants (ses parents) du niveau de résolution inférieur. On scinde ainsi une région ambiguë en quatre sous-régions ambiguës.

- **Localisation précise des contours**

Au début de cette phase, on dispose :

- des prototypes de chaque classe,
- des pixels ambigus,
- et d'une image intermédiaire segmentée en $(c+1)$ classes dont une classe de pixels ambigus.

L'image segmentée est interpolée pour être ramenée au niveau de résolution inférieur où seuls les pixels ambigus font l'objet d'une nouvelle classification. Pour réaliser cette classification, les parents des pixels prototypes sont repérés et leurs *facteurs de forme* sont utilisés pour déterminer les nouveaux centres des c classes. La classification des pixels ambigus qui s'ensuit, se traduit par un affinement des contours grossièrement délimités à la résolution inférieure. Cette classification utilise la distance euclidienne comme fonction discriminante.

Le processus est ensuite itéré jusqu'au niveau de résolution $j=1$ sans que l'on ait à nouveau recours à l'algorithme de classification floue pour la détermination des prototypes. A ce dernier niveau, tous les pixels sont systématiquement attribués à une classe. La notion de texture en effet, n'a aucun sens lorsque l'analyse est réalisée sur un seul pixel. En conséquence, pour $j=1$, aucun pixel n'est classé ambigu.

Afin de lisser les frontières obtenues après la phase d'interpolation qui permet de passer du niveau $j=1$ au niveau $j=0$, l'image interpolée est soumise à un filtrage non linéaire. Le principe de ce filtre est d'attribuer au pixel central, le même label que celui de la classe majoritaire dans son voisinage immédiat.

- **Algorithme de segmentation**

Le squelette de l'algorithme est le suivant :

Etape 1. Réalisation la décomposition en ondelette de l'image jusqu'au niveau de résolution J .

Etape 2. Calcul des trois facteurs de forme pour chaque niveau de résolution j ($1 \leq j \leq J$).

Etape 3 Au niveau de résolution J :

Classification floue (vecteur attribut : les trois facteurs de forme).

Si le degré d'appartenance d'un pixel à une classe $q > \alpha$,
alors marquer ce pixel comme prototype de la classe q ,
 donner à ce pixel le label de la classe q .

Initialisation de j (variable du niveau de résolution courant) : $j = J$

Etape 4. Calcul du centre des prototypes de chaque classe au niveau j .

Etape 5. Affectation des pixels du niveau $j = J$ ou des pixels ambigus de niveau $j \neq J$,
Mesure de distance d'un pixel au centre d'une classe dans l'espace (tridimensionnel) des attributs de texture.

Si $j = J$, alors pixels à considérer = pixels du niveau de résolution j
sinon pixels à considérer = pixels ambigus.

Pour chaque pixel à considérer,

Si la distance d'un pixel au centre d'une classe k est inférieure
 à la distance de ce pixel à tous les autres centres de classe,
alors donner à ce pixel le label de la classe k .

Etape 6. Détermination des pixels ambigus de l'image de résolution courante.

Analyse du voisinage de chaque pixel de l'image de niveau de résolution j .

Si j est différent de 1, et que la classe du pixel considéré n'est pas
 majoritaire dans le voisinage de ce pixel,
alors marquer ce pixel comme ambigu.

Etape 7. Extension de l'image de niveau de résolution j , au niveau de résolution $j-1$.

Pour chaque pixel de l'image de résolution j ,
 donner à ses quatre parents la même étiquette (classe, prototype,
 ambigus) que lui.

Décrémenter j

Etape 8. Amélioration de la segmentation

Si $j \neq 0$, retourner à l'Etape 4.

sinon filtrer l'image du niveau 0 par le filtre non linéaire pour obtenir
 l'image segmentée finale.

6.3.3 Expérimentations et résultats

- *Données des expérimentations*

Les trois images test (*figure 6.3*) de nos expérimentations sont de taille 256×256 avec 256 niveaux de gris. Elles ont été composées à partir de quelques images de textures extraites de la banque d'images de Brodatz et de celles du laboratoire.

Image1 est la composition de deux textures. La classe 1.1 désignera dans la suite, la classe de la texture de sable de gauche et la classe 1.2, celle de la texture de droite. Image2 est composée de gauche à droite et de bas en haut des textures "pullover" (classe 2.1), "siège" (classe 2.2), "écossais" (classe 2.3) et "bois" (classe 2.4). Enfin, Image3 est une mosaïque des textures "canevas" (classe 3.1) et "lierre" (classe 3.2).

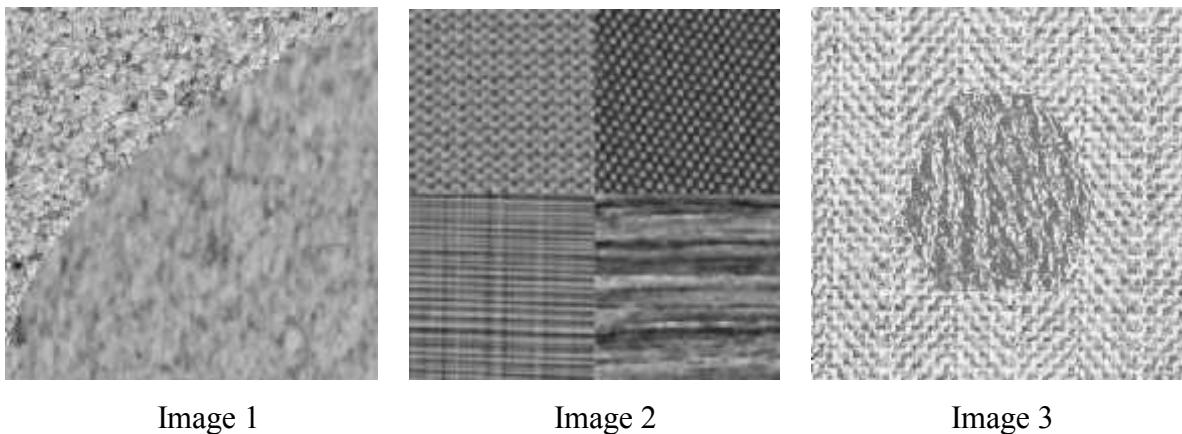


Figure 6.3 : Images test de la segmentation

Tous les algorithmes ont été implantés en langage C sur un micro-ordinateur de type PC, 486DX4, 50 MHz.

La décomposition en ondelettes a été effectuée jusqu'à un niveau J donné. Ce niveau induit une analyse sur des voisinages pyramidaux de support de taille $2^J \times 2^J$. Le niveau J est déterminé suivant le critère défini dans le deuxième algorithme de classification du chapitre précédent. Le niveau de résolution maximum est fixé à 5.

L'ondelette utilisée (à l'étape 1 de la méthode) est celle de Haar. Comme nous l'avons indiqué au chapitre précédent, cette ondelette est très pratique pour la segmentation de textures. La taille des filtres qui lui sont associés, permet en effet de calculer les attributs (étape 2) sur des voisinages pyramidaux, dont les supports sont parfaitement adjacents. De plus, la mise en oeuvre pratique de cette ondelette est d'un coût faible en temps de calcul.

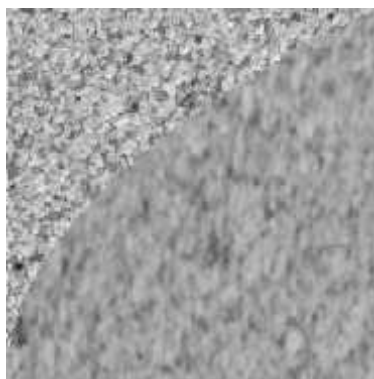
L'initialisation des centres de l'algorithme de classification floue se fait de façon aléatoire en veillant tout de même à ce qu'ils ne soient pas choisis trop proches l'un de l'autre. La valeur de r qui permet de déterminer les "meilleurs" prototypes de chaque classe (étape 3), est fixé à 0.96 pour l'ensemble des expérimentations. Le facteur flou m a été pris égal à 1,5.

Pour une classe donnée, les vecteurs d'attributs (composés des trois facteurs de forme) des meilleurs prototypes permettent de déterminer le centre de la classe (étape 4). Ce centre est le barycentre des attributs associés aux prototypes.

Une mesure de distance à ces centres permet d'affecter tout au long du processus, chaque pixel non classé ou ambigu à une classe (étape 5). La distance employée est la distance euclidienne.

L'analyse du voisinage (étape 6) qui permet de déterminer les pixels ambigus, est effectuée sur une fenêtre de taille 3×3 pour les niveaux de résolution $j \geq 3$, et de taille 5×5 pour les niveaux inférieurs. Ce choix est dicté par la taille du support du voisinage induit et par le nombre plus ou moins important de pixels des images des différentes résolutions.

Le passage d'un niveau de résolution à celui qui lui est immédiatement inférieur (étape 7) se fait par interpolation, en attribuant aux parents, le même label que leur noeud. Le filtre non linéaire final (de l'étape 8) est de taille 7×7 .

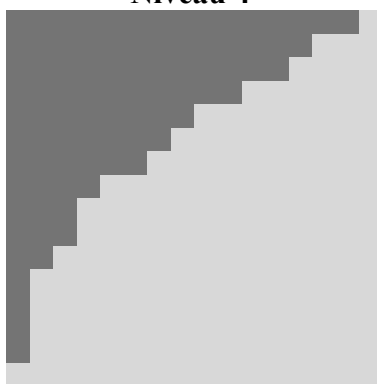


[a]

Niveau 4



[b]

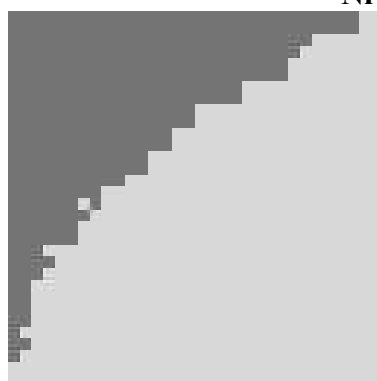


[c]

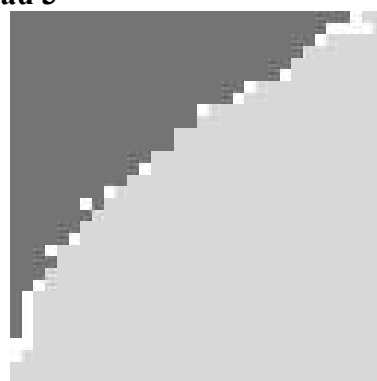


[d]

Niveau 3



[e]



[f]

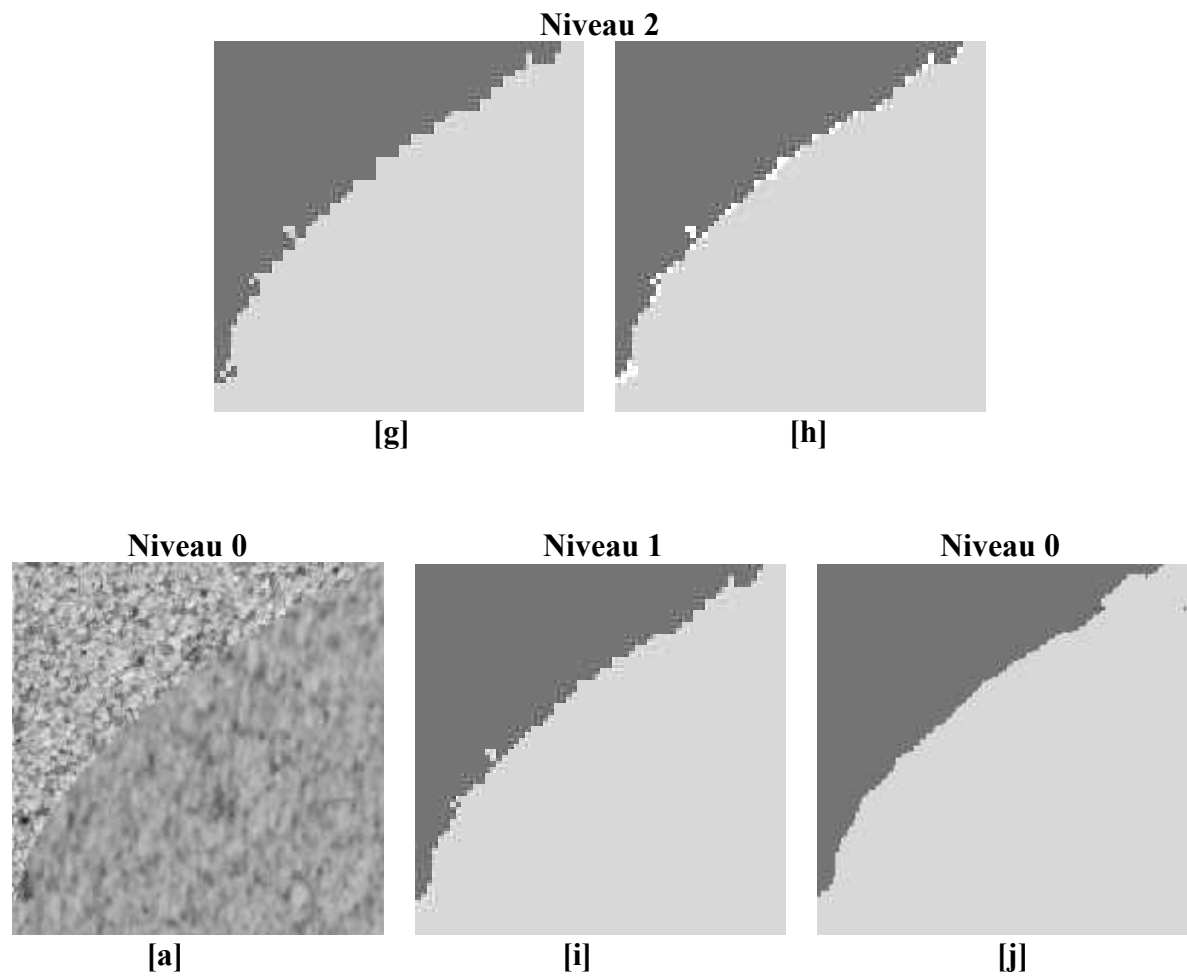


Figure 6.4 : Résultats de la segmentation de Image1. Les pixels flous apparaissent en noir sur l'image [b], et les pixels ambigus en blanc. L'image [j] est le résultat de la segmentation de l'image [a].

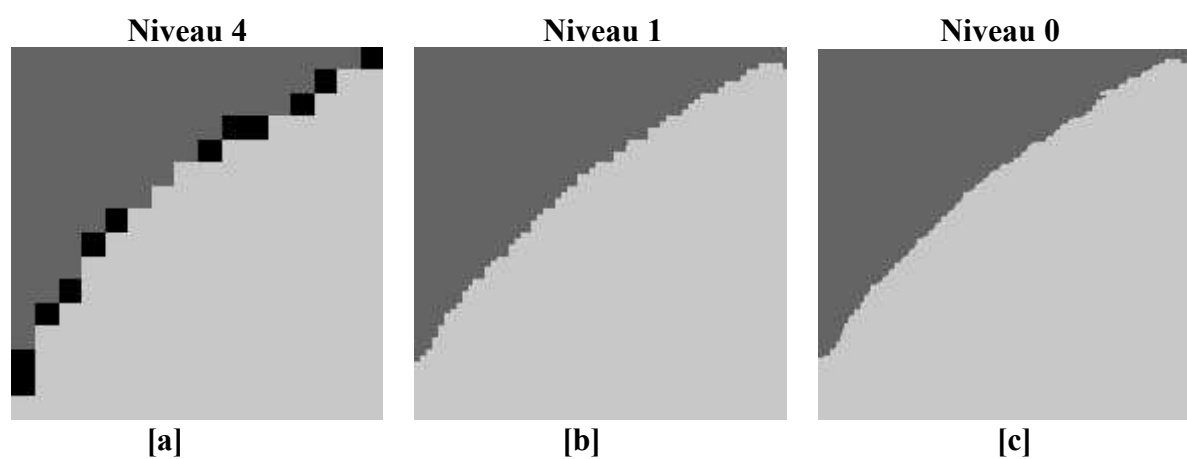


Figure 6.5 : Résultat de la segmentation de Image1 moyennée. L'image [a] est obtenue après le processus de classification floue (les pixels flous sont en noir). Les images [b] et [c] correspondent respectivement aux images résultats avant et après application du filtre non linéaire.

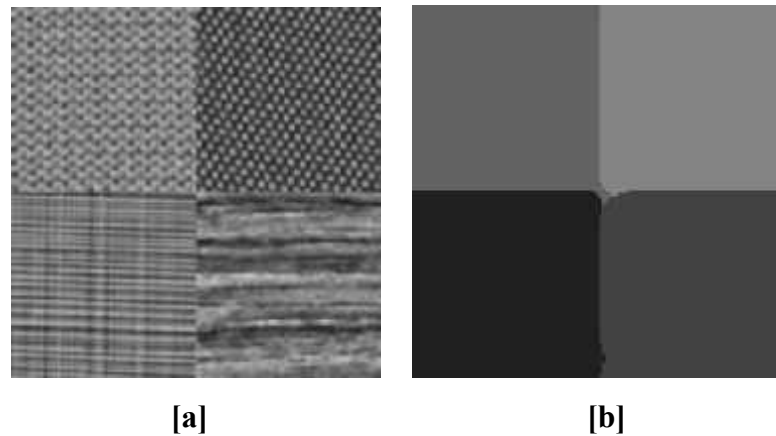


Figure 6.6 : Résultat ([b]) de la segmentation de Image2 ([a])

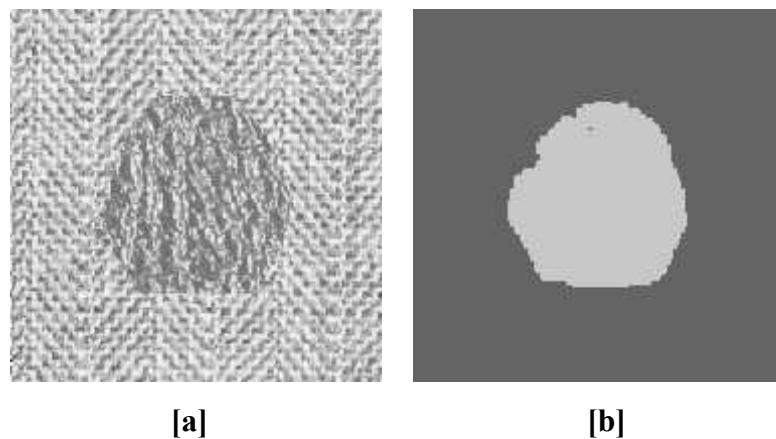


Figure 6.7 : Résultat ([b]) de la segmentation de Image3 ([a])

- **Résultats et commentaires**

Les images de la *figure 6.4 [b]* à *6.4 [j]*, illustrent les différentes phases de l'algorithme de segmentation proposé. L'image d'origine est celle de la *figure 6.4 [a]*. Le processus de segmentation débute au niveau 4.

Toutes les images intermédiaires de la segmentation ont été interpolées pour être ramenées au niveau de résolution 0 afin d'en faciliter la lecture.

La *figure 6.4 [b]* montre le résultat de la classification floue. Huit itérations ont été nécessaires pour la détermination des centres des classes de l'algorithme du FCM. Les pixels flous dont le degré d'appartenance est inférieur au seuil r choisi, sont représentés en noir. On note que seuls 7 pixels sur 256 se trouvent dans ce cas. La mesure de distance de ces pixels aux centres des deux classes de prototypes, permet de les affecter à une classe. Le résultat de cette opération est donné à la *figure 6.4 [c]*.

Les résultats montrés par les *figures 6.4 [b], [c], [e] et [g]*, confirment une fois de plus, le caractère discriminant du facteur de forme. Ce caractère se traduit par une bonne localisation des pixels prototypes et de celle des pixels appartenant aux zones intérieures des régions. Remettre en cause leur classification aux niveaux inférieurs s'avère donc inutile.

En revanche, la mauvaise délimitation des frontières au niveau de résolution 4, impose que les pixels frontières soient reclassés. Ces pixels sont en couleur blanche sur les images des *figures 6.4 [d], [f] et [h]*. Le *Tableau 6.1* indique pour chaque itération j , le nombre de ces pixels et le pourcentage qu'ils représentent par rapport à la taille de l'image de niveau de résolution. Rappelons que la méthode proposée ne prévoit pas de détermination de pixels ambigus au niveau 1.

<i>Niveaux</i>	<i>Pixels ambigus</i>	<i>Pourcentage ($P_c\%$)</i>
4	11	04.3
3	43	04.2
2	40	00.98

Tableau 6.1 : Nombre et pourcentage des pixels ambigus de l'Image1 pour différents niveaux de résolution.

On peut constater que le pourcentage de pixels à reconsidérer à chaque itération est faible.

Le résultat final de la segmentation qui est obtenu après interpolation et application du filtre non linéaire de l'image segmentée du niveau de résolution 1 (*figure 6.4 [i]*), est donné par la *figure 6.4 [j]*. Le filtre non linéaire permet visiblement de lisser les contours.

Comme nous l'avions indiqué au chapitre précédent, le caractère aléatoire de la texture "sable" influence fortement les résultats de la classification. Le résultat de la segmentation peut être nettement amélioré si l'on fait précéder la TOD d'un filtrage moyennneur (*figure 6.5 [b et c]*). On peut constater que les pixels flous (en noirs sur la *figure 6.5 [a]*) sont alors uniquement localisés à la frontière des deux régions.

Les *figures 6.6 [b]* et *6.7[b]* présentent les résultats de la segmentation de Image2 et de Image3 (*figure 6.6 [a]* et *figure 6.7 [a]*). Pour ces deux images test, le niveau de départ du processus de segmentation est également le niveau $J=4$.

Tous les résultats chiffrés de la segmentation des trois images test, sont résumés dans le *Tableau 6.3*.

	<i>Image 1</i>		<i>Image 2</i>		<i>Image 3</i>	
<i>Niveau de départ de la segmentation</i>	4		4		4	
<i>Nombre de pixels flous</i>	7		14		7	
<i>Taux correct de classification de la classe (x.x) en (%)</i>	(1.1)	97.3	(2.1)	98.6	(3.1)	96.5
	(1.2)	97.8	(2.2)	98.9	(3.2)	95.7
			(2.3)	99.9		
			(2.4)	98.8		
<i>Durée de l'opération (s)</i>	33.5		37		35	

Tableau 6.2 : Résultats de la segmentation des images test . La classe est indiquée entre parenthèses.

- **Influence du niveau de départ de la segmentation**

Cette expérimentation ne concerne que la *figure 6.4 [a]*. Elle a pour objet d'apprécier l'influence du niveau de la décomposition sur le résultat de la segmentation.

Les images de la *figure 6.8* montrent les résultats obtenus respectivement pour les valeurs 3, 4 et 5 de J . Sur cette figure, l'indice [a] correspond au résultat de la classification floue, l'indice [b] à celui de la segmentation obtenue avant le filtrage non linéaire, et l'indice [c] au résultat final.

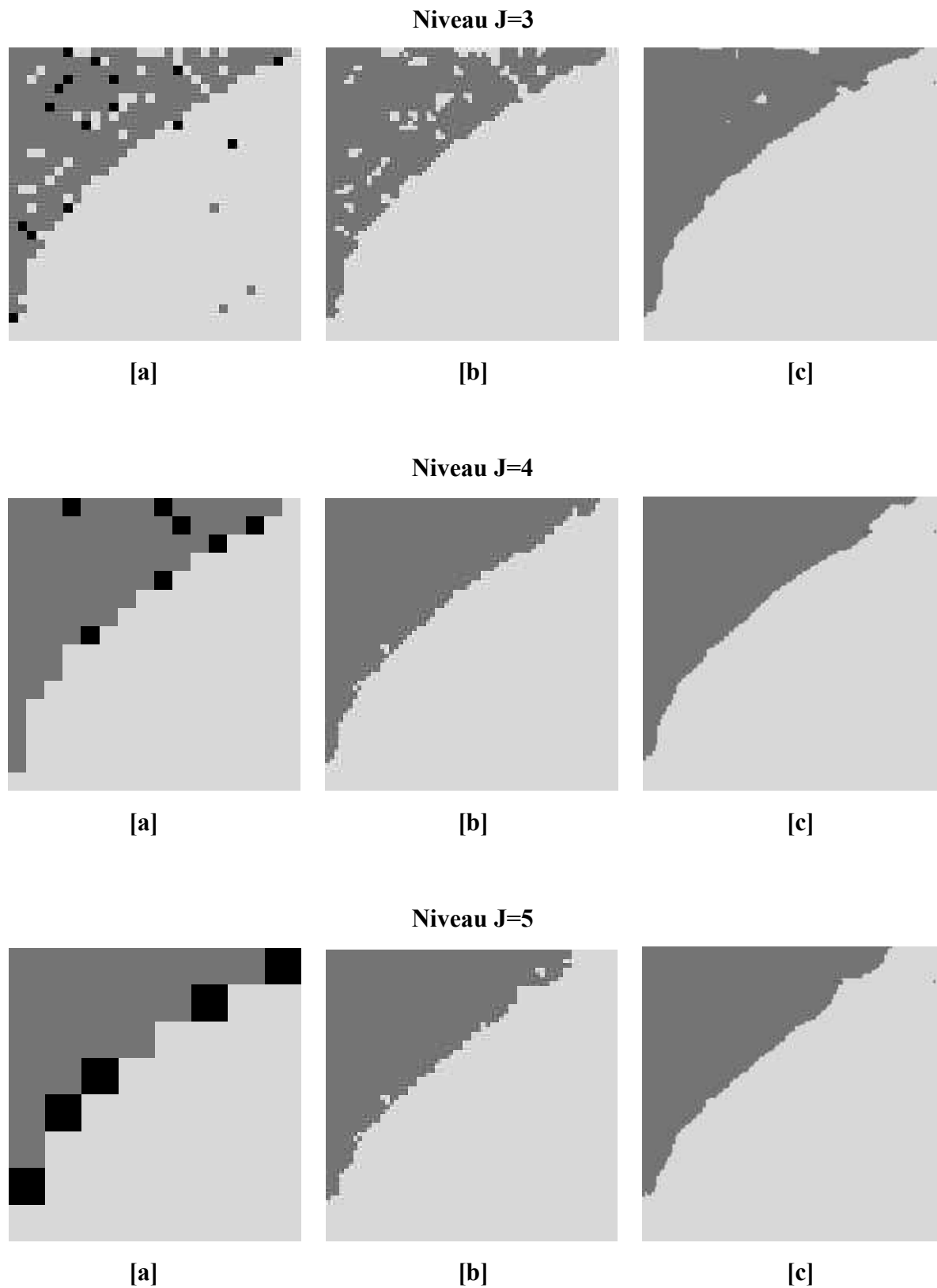


Figure 6.8 : Résultats de la segmentation à différents niveaux de résolutions de Image1. L'image [a] est obtenue après le processus de classification floue (les pixels flous sont en noir). Les images [b] et [c] correspondent respectivement aux images résultats avant et après application du filtre non linéaire

On constate que le classifieur flou fournit pour $J=3$, un taux d'erreur de classification élevé : plusieurs pixels (41) sont classés à l'intérieur d'une région à laquelle ils n'appartiennent pas. Ce qui laisse supposer que les attributs sont moins pertinents à ce niveau. Par contre, pour $J=4$ et $J=5$, le fort pouvoir discriminant du facteur de forme aux niveaux élevés est confirmé.

L'analyse du voisinage effectuée à chaque itération de la décomposition permet progressivement de corriger les erreurs de classification des niveaux élevés. Les images résultats obtenues pour $J=3$ et $J=5$, le démontrent bien. On peut cependant noter que la segmentation grossière obtenue pour $J=5$, se caractérise par de gros pavés qui influencent le résultat de la segmentation finale.

Ces images illustrent particulièrement la délicatesse du choix du niveau d'arrêt de la décomposition par ondelettes. Un niveau d'arrêt élevé assure une bonne discrimination des textures en présence, mais induit une taille importante du support de voisinage qui affecte les performances de la classification. Les erreurs de classification dans le cas $J=5$ par exemple, sont essentiellement localisées dans la zone frontière. En revanche, un niveau de départ de la segmentation trop bas, entraîne trop bas le phénomène inverse. Les erreurs de classification pour le niveau de départ $J=3$, le prouvent. Elles sont également localisées dans la zone intérieure de la texture "sable".

Le choix du niveau de départ doit donc relever d'un compromis qui doit prendre en compte la taille des éléments de texture et le nombre de textures de l'image. Le critère sur les variances basses résolutions que nous avons défini au chapitre précédent, peut contribuer au choix de ce niveau notamment lorsque le nombre de régions texturées de l'image est relativement limité. Pour la présente application, le Tableau 6.3 montre que le niveau 4 déterminé par ce critère, fournit les meilleurs résultats.

		$J=3$	$J=4$	$J=5$
<i>Pixels flous</i>		16	12	5
<i>Pixels ambigus</i>	$J=5$			8
	$J=4$		11	11
	$J=3$	53	43	25
	$J=2$	152	40	43
<i>Taux correct de classification (%)</i>				
(1.1)		92.6	97.3	93.2
(1.2)		93.8	97.8	95.2
<i>Durée totale de l'opération (s)</i>		37	33.5	35

Tableau 6.3 Résultats de la segmentation réalisée à différents niveaux de résolution

- ***Intérêt du voisinage pyramidal***

Afin de démontrer l'intérêt du voisinage pyramidal dans le problème de segmentation, nous avons reconduit le processus de segmentation de l'image de la *figure 6.4 [a]*, en définissant cette fois-ci, un voisinage plan de taille 5×5 sur lequel les attributs sont calculés. L'attribut utilisé est l'énergie locale de texture qui a été définie dans le premier algorithme de classification du chapitre précédent.

Les *figures 6.9 [a]* et *[b]* montrent les résultats de la classification floue et de la présegmentation au niveau de résolution $J=4$. Le voisinage utilisé pour le calcul des attributs est de type plan. L'inconvénient de ce voisinage plan est nettement visible sur la *figure 6.9 [a]*. En effet, contrairement aux cas précédents (*figures 6.4 [b]* et *6.5 [a]*), on peut noter que la grande ambiguïté qui affecte la localisation exacte des contours, se traduit par une bande épaisse de pixels flous au niveau de la frontière des deux régions texturées. Ce qui conduit finalement à une importante délocalisation des contours des régions texturées de l'image d'origine (*figures 6.9 [c]*).

Il apparaît nettement sur la *figure 6.9 [b]* (qui montre le résultat de la segmentation grossière du niveau 4), qu'il est nécessaire de reporter la classification d'un grand nombre de

pixels au niveau inférieur pour éviter d'obtenir une segmentation avec un taux d'erreur élevé (figure 6.9 [c]).

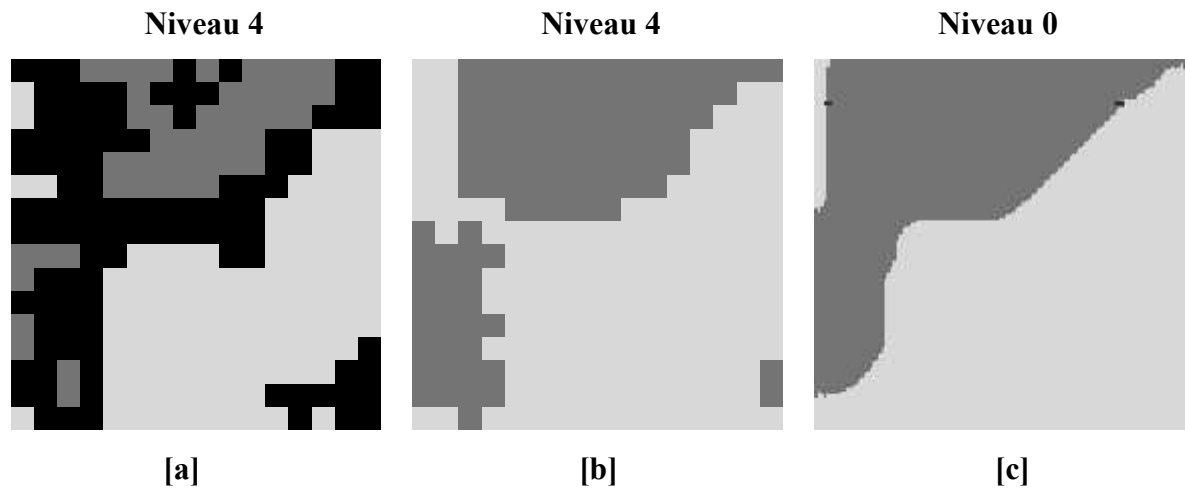


Figure 6.9 : Résultats de la segmentation obtenus en utilisant un voisinage plan. L'image [a] est obtenue après le processus de classification floue (les pixels flous sont en noir). Les images [b] et [c] correspondent respectivement aux images résultats de la présegmentation grossière et de la segmentation finale.

Nous avons également observé un temps de calcul plus important (46 s). Cela provient du nombre plus élevé de pixels ambigus (dû au voisinage utilisé) et de la prise en compte à chaque niveau de résolution, d'un nombre constant de pixels (nombre fixé par la taille du voisinage utilisé) pour le calcul des attributs.

Ces résultats illustrent parfaitement l'intérêt du voisinage pyramidal en segmentation d'images.

- **Influence de l'ondelette analysante**

Nous avons également regardé l'influence de l'ondelette choisie sur les résultats de notre méthode de segmentation. Pour cela, l'ondelette de Daubechies à deux moments nuls (cf. Chapitre 5) a été implanté. Le facteur de forme est déterminé sur un voisinage pyramidal.

Les figures 6.10 [a], [b], et [c], montrent respectivement les résultats de la classification floue, de la segmentation avant et après application du filtre non linéaire.

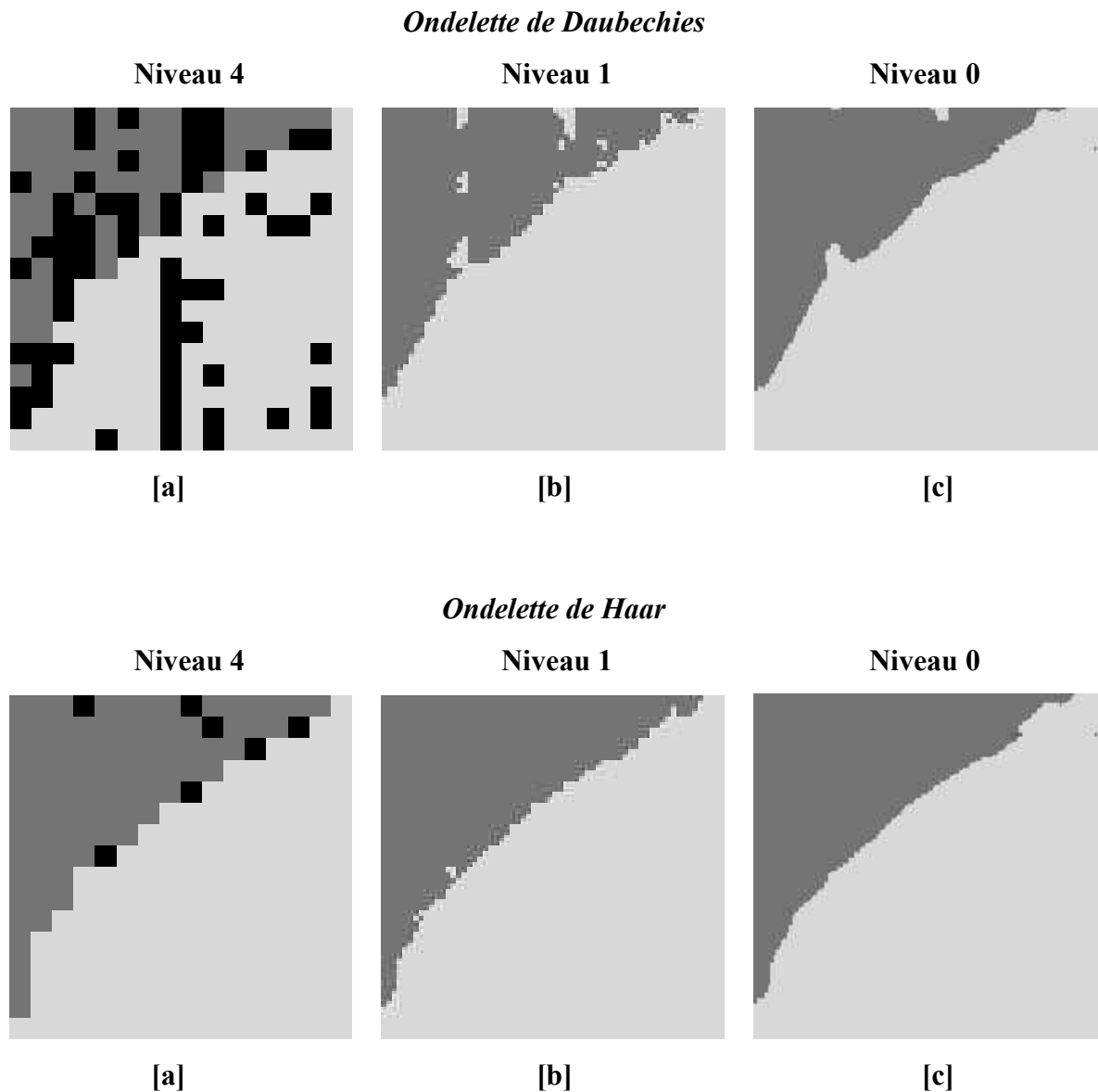


Figure 6.10 : Résultats de la segmentation obtenus avec deux ondelettes différentes. L'image [a] est obtenue après le processus de classification floue (les pixels flous sont en noir). Les images [b] et [c] correspondent respectivement aux images résultats avant et après application du filtre non linéaire.

Les résultats obtenus avec l'ondelette de Daubechies sont nettement moins bons que ceux fournis par l'ondelette de Haar. La délimitation des contours est en effet moins précise dès le départ; ce qui a pour conséquence d'affecter le résultat final de la classification. Cette délocalisation de la frontière, qui est plus marquée dans le cas de l'ondelette de Daubechies que dans celui de Haar, s'explique par la taille relativement grande des supports des filtres associés à l'ondelette de Daubechies. Dans les zones frontières, cette grande taille fait

intervenir pour le calcul des attributs, des pixels appartenant à différentes régions texturées. Par conséquent, les attributs associés aux pixels appartenant à ces zones sont erronés et ne caractérisent en réalité aucune région texturée.

Il est intéressant de noter que l'algorithme de segmentation qui utilise l'ondelette de Daubechies et un voisinage pyramidal, fournit de meilleurs résultats que celui qui met en oeuvre l'ondelette de Haar et un voisinage plan. C'est dire de nouveau, tout l'intérêt du voisinage pyramidal.

Les résultats chiffrés des expérimentations, mettant en oeuvre les ondelettes de Haar et de Daubechies, sont récapitulés dans le *Tableau 6.4*.

	<i>Ondelette de Haar</i>		<i>Ondelette de Daub.</i>	
<i>Nombre d'itérations</i>	8		17	
<i>Pixels flous</i>	7		65	
<i>Taux correct</i>	(1.1)	97.3	(1.1)	93.6
<i>de classification (%)</i>	(1.2)	97.8	(1.2)	94.6
<i>Durée de l'opération (s)</i>	33.5		46	

Tableau 6.5 : Tableau récapitulatif des résultats obtenus avec deux ondelettes différentes.

La lecture de ce tableau laisse apparaître le fait que l'ondelette de Haar conduit dans notre approche, à de meilleurs résultats tant en temps de calcul qu'en taux de classification correcte.

Indiquons, pour finir cette série d'expérimentations, que la valeur du paramètre α de l'algorithme de classification floue. Les valeurs associées aux centres des classes varient en effet très peu en fonction de α . Comme on peut d'ailleurs le constater sur les images présentées (*figures 6.10[a]* par exemple), une grande majorité des pixels est déterminée "prototype", même pour des valeurs élevées de α (0.96 dans notre cas).

- *Critiques de la méthode de segmentation proposée*

Deux points principaux peuvent limiter les performances de l'algorithme proposé.

Le premier point découle de l'utilisation du filtre non linéaire, qui bien que lissant les contours, les délocalise également. Cette propriété inhérente à ce type de filtrage, pose le problème du passage du niveau de résolution 1 au niveau 0. Ce problème est difficile à résoudre étant donné d'une part, que notre approche ne permet pas de calculer les attributs correspondants au niveau 0, et que d'autre part, une segmentation de textures ne peut pas être effectuée en considérant individuellement chaque pixel. Une solution serait de réaliser uniquement pour le niveau 1, une TOD sans sous-échantillonnage, et de se servir des informations des images filtrées pour affiner le contour. Cette solution revient en quelque sorte à envisager une coopération contour-régions dans la dernière étape de la segmentation.

Le second point limitatif est rattaché aux propriétés de l'attribut de textures que nous avons défini, et à la relative simplicité de notre algorithme.

Nous avons montré dans ces deux derniers chapitres, que la pertinence du facteur de forme croît avec le niveau de résolution. Or un niveau de résolution élevé, implique une taille relativement grande du voisinage pyramidal sur lequel l'attribut est calculé. Notre algorithme qui est initialisé à un niveau de résolution relativement élevé, est pris à défaut lorsque l'image à segmenter est constituée de plusieurs régions de petites tailles. L'intérieur des régions dans ce cas est alors difficilement caractérisable. L'algorithme de classification floue fournit alors généralement des prototypes qui sont plutôt représentatifs de zones frontières que de zones situées à l'intérieur des régions.

6.4 Conclusion

Nous avons décrit dans ce chapitre, un algorithme de segmentation pyramidale d'images texturées fondée sur une décomposition en ondelettes. Le voisinage pyramidal introduit au chapitre précédent s'est révélé être véritablement avantageux pour une opération de segmentation qui opère sur des données représentées à différents niveaux de résolution.

L'algorithme proposé évolue suivant une stratégie "du plus grossier au plus fin" qui se traduit par une analyse graduelle des zones singulières (zones frontières et zones isolées). Son déroulement rappelle certains mécanismes d'analyse (décomposition multicanaux) et de perception (préattentive et focale) du système visuel humain.

Les performances en temps de calcul de cet algorithme sont réelles. Les résultats obtenus sont excellents. Ils établissent la pertinence de l'attribut facteur de forme et la démontrent la justesse de l'utilisation d'un voisinage pyramidal pour le calcul des attributs lorsqu'une approche multirésolution est adoptée.

Il est envisageable de faire évoluer cet algorithme vers un algorithme de segmentation non supervisée afin de s'affranchir de la connaissance a priori du nombre de classes.

CONCLUSION

CONCLUSION

Nous avons montré qu'à partir des résultats de la décomposition en ondelettes d'une texture, il est possible de la caractériser par ses variations locales de niveaux de gris, en combinant les informations des différents niveaux de résolution.

Pour réaliser pratiquement cette caractérisation, nous avons introduit un nouveau type de voisinage, le voisinage pyramidal, sur lequel se calcule le nouvel attribut de textures, le facteur de forme, que nous avons défini.

Ces deux éléments essentiels de notre méthode, possèdent des propriétés intéressantes qui ont été exploitées dans les algorithmes proposés.

La taille adaptative du voisinage pyramidal permet de caractériser selon le niveau de résolution considéré, les propriétés locales ou globales de la texture. Il nous a ainsi été possible dans le premier algorithme de classification, de réaliser une discrimination multirésolution efficace des différentes textures.

De même, dans le second algorithme qui met à profit les propriétés d'unicité et de complétude de la décomposition en ondelettes orthonormales, le facteur de forme a été utilisé pour conduire la classification des textures.

L'étude comparative réalisée avec des méthodes classiques et les résultats satisfaisants obtenus, ont démontré l'efficacité de la méthode proposée, et établi la pertinence du facteur de forme.

Cette pertinence a été exploitée dans l'algorithme de segmentation notamment pour réaliser au niveau de résolution le plus élevé de la décomposition, une segmentation grossière

de l'image de textures. La délimitation exacte des contours des régions a été ensuite obtenue de façon graduelle en passant aux niveaux de résolution inférieurs. L'utilisation du voisinage pyramidal en lieu et place du voisinage plan classique, a permis de réduire sensiblement l'ambiguïté sur la localisation exacte des frontières entre textures; cette ambiguïté étant minimale lorsque l'analyse par ondelettes est réalisée avec l'ondelette de Haar.

Ce travail démontre de nouveau l'intérêt de l'outil "ondelettes". Le voisinage pyramidal qui a été introduit, constitue pour les méthodes d'analyse multirésolution, une alternative intéressante au voisinage plan.

*La connaissance enfle, mais l'amour édifie. Si
quelqu'un croit savoir quelque chose, il n'a pas
encore connu comme il faut connaître.*

(2 Cor. 8 v. 1-2, LA BIBLE)

BIBLIOGRAPHIE

REFERENCES BIBLIOGRAPHIQUES

- [Ade.83] F. Ade,
"Characterization of Textures by Eigenfilters",
 Signal Processing, Vol.5, p. 451-457, (1983).
- [Ale.92] C. D'Alessandro, C. Demars,
"Représentation Temps-Fréquence du Signal de Parole",
 Traitement du signal, Vol. 9, No. 2, p.153-173, 2ème Trimestre, (1992)
- [Ant.93] R.M.J.P. Antoine, P. Carette, B. Piette,
"Image Analysis with Two-Dimensionnal Continous Wavelet Transform",
 Signal Processing, Vol. 31, p.241-272, (1993)
- [Baj.73] R. Bajcsy,
"Computer Identification of Visual Surfaces",
 Computer Graphics and Image Processing, Vol. 2, p. 118-130, (1973).
- [Bal.82] D.H. Ballard, C.M. Brown,
"Computer Vision",
 Ed. Prentice-Hall, inc., New Jersey, (1982).
- [Ber.91] J.P. Berroir,
"Application des Mesures Multifractales à l'Analyse et à la Segmentation d'Images Médicales"
 Congrès RFIA, Lyon-Villeurbanne, p. 979-983, (1991).
- [Bou.91] C. Bouman, B. Liu,
"Multiple Resolution Segmentation of Textured Images",
 IEEE Trans. on Pattern Anal. Machine Intell., Vol. 13(2), p. 99-113, February (1991).
- [Bro.66] P. Brodatz,
"Textures - A Photographic Album for Artists and Designers",
 New York : Dover, (1966).
- [Bov.90] A.C. Bovik, M. Clark, W.S. Geisler,
"Multichannel Texture Analysis using Localized Spatial Filters",
 IEEE Trans. on Pattern Anal. Machine Intell. Vol. 12(1), p. 55-72, January (1990).
- [Bov.91] A.C. Bovik,
"Analysis of Multichannel Narrow-Band Filters for Image Texture Segmentation",
 IEEE Trans. Signal Processing. Vol. 39, p. 2025-2043, (1991).
- [Buf.90] J.M. Du Buf,
"Gabor Phase in Texture Discrimination",
 Signal Processing, 21, p. 221-240, (1990).
- [Bur.81] P.J. Burt,
"Fast Filter Transforms for Image Processing",
 Computer Graphics and Image Processing, Vol. 16, p. 20-51, (1981).

[Ade.83-Bur.81]

[Bur.83] P.J. Burt, A.E. Adelson,

"The Laplacian Pyramid as a Compact Image Code"
IEEE Trans. Communications, Vol. 31, p. 532-540, April. (1983).

[Bus.87] P. Buser, M. Imbert,
"Vision",
Hermann, Paris, Collection Méthodes, (1987).

[Cam.68] F.W. Campbell, J.G. Robson,
"Application of Fourier Analysis to the Visibility of Gratings",
J. Physiol., vol. 197, p. 551-566, (1968).

[Cha.92] T.C. Chang, C.-C.J. Kuo,
"A Wavelet Transform Approach for Hierarchical Texture Segmentation",
Conference Rec. of The 26th Asilomar, Conf. on Signals, Systems and Comp. Vol. 2, p. 816-820, (1992).

[Cha.93] T.C. Chang, C.-C.J. Kuo,
"Texture Analysis and Classification with Tree-Structured Wavelet Transform",
IEEE Trans. on Image Processing, Vol. 2, No.4, p. 429-441, October (1993).

[Che.85] R. Chellapa, S. Chatterjee,
"Classification of Textures using Gaussian Markov Random Fields",
IEEE ASSP-33(4), p. 959-963, August (1985).

[Che.95] J.L. Chen, A. Kundu,
"Unsupervised Texture Décomposition and Hidden Markov Models ",
IEEE Trans. on Image Proc., Vol. 4(5), p. 603-619, May (1995).

[Chu.92] C.K. Chui
"Vol. 1 : An Introduction to Wavelets",
"Vol. 2 : Wavelet: A Tutorial in Theory and Applications"
Editions Academic Press, New-York, (1992)

[Coc.95] J.-P. Cocquerez, S. Philip, et al,
"Analyse d'Images : Filtrage et Segmentation",
Editions Masson, Paris, (1995).

[Coh.92a] A. Cohen
"Ondelettes et Traitement Numérique du Signal"
Recherches en Mathématiques Appliquées, RMA 25, Editions Masson, Paris, (1992)

[Coh.92b] A. Cohen, I. Daubechies, J.C. Feauveau
"Biorthogonal Basis of Compactly Supported Wavelets"
Comm. Pure Appl. Math., Vol. 45, p. 485-560, (1992)

[Con.80a] R.W. Connors, C.A. Harlow,
"Towards a Structural Textural Analyser based on Statistical Methods",
Comput. Graphics Image Process., Vol. 12, p. 224-256, (1980).

[Con.80b] R.W. Connors, C.A. Harlow,
"A Theoretical Comparison of Texture Algorithms",
IEEE Trans. Pattern Analysis Mach. Intell., PAMI-8(1), p. 118-125, (1980).

[Cro.83] G.C. Cross, A.K. Jain,
"Markov Random Field Texture Models",
IEEE Trans. Pattern Analysis Mach. Intell., PAMI-5, p. 25-39, (1983).

[Bur.83-Cro.83]

[Dau.80] J.G. Daugman,
"Two-Dimensionnal Spectral Analysis of Cortical Receptive Field Profile",

Vision Research, vol. 20, p. 847-856, (1980).

[Dau.85] J.G. Daugman,

"Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation, Optimised by a Two-Dimensional Visual Cortical Filters",

J. Opt. Soc. Am., vol. A2, p. 1160-1169, (1985).

[Dau.88] I. Daubechies,

"Orthonormal Bases of Compactly Supported Wavelets,"

Comm. in Pure and Applied Math., Vol. 41, No. 7, p. 909-996. (1988).

[Dau.90] I. Daubechies,

"The Wavelet Transform. Time-Frequency Localization and Signal Analysis,"

IEEE Trans. on Info. Theory, Vol. 36, No. 5, pp. 961-1005. Sept. (1990).

[Dau.92] I. Daubechies,

"Ten lectures on Wavelets,"

SIAM Publications, No. 61, (1992).

[Dun.95] D. Dunn, W.E. Higgins,

"Optimal Gabor Filters for Texture Segmentation",

IEEE Trans. on Image Processing, Vol. 4, No. 7, p. 947-964, (1995).

[Fea.90] J.C. Feauveau

Analyse Multirésolution pour les Images avec un Facteur de Résolution de 2.

Traitement du Signal, vol. 7, No. 2, pp. 117-128, (1990)

[Fla.84] P. Flandrin,

"An Interpretation of the Pseudo Wigner-Ville Distribution ",

Signal Processing, 6(1), p. 27-36, (1984).

[Gag.83] A. Gagalowicz,

"Vers un modèle de textures",

Thèse d'Etat, Paris VI, (1983).

[Gal.75] M;M. Galloway,

" Textural Analysis using Gray Level Run Lenghts",

Comput. Graphics Image Process., Vol.4, p. 172-179, (1975).

[Gau.96] L. Gaudart, J. Crebassa, J.P. Petrakian,

"Fonction Ondelette et Champs Récepteurs Chromatiques",

Innovation et Technologie en Biologie et Médecine, ITBM, Vol.17, n°1, (1996).

[Goo.85] L. Van Gool, P. Dewaele, A. Oosterlink,

"SURVEY : Texture Analysis anno 1983",

Comput. Vision Graphics Image Process., 29, p. 336-357, (1985).

[Gro.84] A. Grossman, J. Morlet,

"Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape,"

SIAM J. Math., Vol. 15, p. 723-736, (1984).

[Har.73] R.M. Haralick, K. Shanmugam, I. Dinstein,

"Textural Features for Image Classification,"

IEEE Trans. Syst. Man Cybern., SMC-3(6), p. 610-621, (1973).

[Dau.80-Har.73]

[Har.79] R.M. Haralick,

"Statistical and Structural Approaches to Texture",

Proc. IEEE 67, p. 786-804, (1979).

-
- [He.88] D.C. He, L. Wang, J. Guibert,
"Texture Discrimination Based on an Optimal Utilization of Texture Features",
 Pattern Recognition 21, No.2, p. 141-146, (1988).
- [He.92] D.C. He, L. Wang,
"Unsupervised Textural Classification of Images using the Texture Spectrum",
 Pattern Recognition 25, p. 247-255, (1992).
- [Her.93] C. Herley, J. Kovacevic, K. Ramchandran, M. Vetterli,
"Time Varying Orthonormal Tilings of the Time-Frequency Plane",
 IEEE Proc. ICASSP, Vol. III, p.1-4, (1993).
- [Hub.62] D.H. Hubel, T.N. Wiesel,
"Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's visual Cortex",
 Journal of Physiology, 160, p.106-154, (1962).
- [Jac.88] L. Jacobson, H. Wechsler,
"Joint Spatial/ Spatial-frequency representation",
 Signal Processing, 14(1), p. 37-68, January (1988).
- [Jai.91] A. K. Jain, F. Farrokhnia,
"Unsupervised Texture Segmentation using Gabor Filters",
 Pattern Recognition, vol. 34, no. 12, p. 1167-1186, (1991).
- [Jul.62] B. Julesz,
"Visual Pattern Discrimination",
 IRE Trans. on Information Theory, Vol. IT-8, p. 84-92, February, (1962).
- [Jul.81] B. Julesz,
"Textons, the Elements of Texture Perception and their Interactions",
 Nature, 290, p. 91-97, (1981).
- [Kas.86] R.L. Kashyap, A. Khotanzad,
"A Model-based Method for Rotation Invariant Texture Classification",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI-8(4), July (1986).
- [Ker.94] C. Kervrann, F. Heitz,
"Segmentation non Supervisée des Images Naturelles Texturées : une approche statistique",
 Traitement du Signal, Vol. 11- No. 1, p. 31-41, (1994).
- [Kim.92] H-B. Kim, R-H. Park,
"Extracting Spatial Arrangement of Structural Textures using Projection Information",
 Pattern Recognition 25, No.3, p. 237-245, (1992).
- [Kir.83] S. Kirkpatrick, C. Gelatt, M. Vecchi,
"Optimization by simulated annealing ",
 Science, 220, p. 671-680, (1983).
- [Kis.94] C. Kiss, G. Nemeth, P. Csillag, L. Boroczky
"Texture analysis based on Wavelet Decomposition"
 Int. Workshop on Image Proc. : Theory, Methodology, Systems and Applications, Budapest, Hungary, Vol. 45,
 p.47-49, June (1994).
-
- [Kpa.91] K. Kpalma, V. Haese-Coat, A. Bruno,
"Analyse Pyramidale de Textures Naturelles à travers la Dimension Fractale",
 13ème Colloque GRETSI, Juan-Les Pins, p. 1053-1056, 16 au 20 Sept. (1991)
- [Har.79-Kis.94]

-
- [Lai.93] **A. Laine, J. Fan**
"Texture Classification by Wavelet Packet Signatures"
 IEEE PAMI, Vol. 15, No. 11, p.1186-1191, November (1993).
- [Law.79] **K.I. Laws,**
"Texture Energy Measures",
 Proc. Image Understanding Workshop, p. 47-51, (1979).
- [Lev.92] **J. Levy-Vehel, P. Mignot, J.P. Berroir,**
"Multifractals, Texture, and Image Processing",
 In Proc. IEEE Comp. Vis. Pat. Recog., CVPR, p. 661-664, (1992).
- [Lev.95] **Jacques Levy-Vehel,**
"Approches fractales en Traitement d'Images",
 15ème Colloque GRETSI, Juan-Les Pins, p. 1377-1382, 18 au 21 Sept. (1995)
- [Lou.95] **G. Loum, J. Lemoine, E. Petit, P. Provent,**
"An Application of Wavelet Transform to Texture Analysis",
 9th Conf. on Image Analysis SCIA, Uppsala, Sweden, p. 583-590, 6-9 Juin (1995).
- [Lou.96a] **G. Loum, P. Provent, J. Lemoine, E. Petit,**
"A New Method for Texture Classification based on Wavelet Transform",
 3rd IEEE SP Int. Symp. on Time-Frequency and Time-Scale Analysis, TSTS 96, Paris, France,
 p. 29-32, 18-21 Juin (1996).
- [Lou.96b] **G. Loum, S. M'hiri, J. Lemoine, E. Petit, P. Provent,**
"Segmentation Hierarchique de textures à l'aide d'une Décomposition en Ondelettes",
 3ème Congrès Africain sur la Recherche en Informatique CARI 96, accepté, 9-19 Octobre (1996).
- [Maf.76] **L. Maffei, A. Fiorentini,**
"The Unresponsive Regions of Visual Cortical Receptive Fields",
 Vision Res., Vol. 16, p. 1131-1139, (1976).
- [Mal.89a] **S.G. Mallat,**
"A Theory for Multiresolution Signal Decomposition : the Wavelet Représentation",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI, Vol. 2, No. 7, July, (1989).
- [Mal.89b] **S.G. Mallat,**
"Multifrequency Channel Decompositions of Images and Wavelets Models",
 IEEE Trans. Acoust., Speech, and Signal Proc., vol. 37, no. 12, p. 2091-2110, Dec., (1989).
- [Mal.90] **J. Malik, P. Perona,**
"Preattentive Texture Discrimination with early Vision Mechanisms",
 J. Opt. Soc. Amer., Vol. A 7, p. 923-932, (1990).
- [Man.91] **B.S. Manjunath, R. Chellapa,**
"Unsupervised Texture Segmentation using Markov Random Field Models",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI Vol. 13, pp. 478-482, (1991).
- [Mao.92] **J. Mao, A.K. Jain,**
"Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models",
 Pattern Recognition 25, pp. 173-188, (1992).
-

[Kpa.91-Mao.92]

- [Mar.80] **S. Marcelja,**
"Mathematical Description of the Responses of Simple Cortical Cells",
 J. Opt. Soc. Amer., Vol. 70, p. 1297-1300, (1980).

- [Mar.82] **D. Marr,**

"Vision : a Computational Investigation into the Human Representation and Processing of Visual Information", W.H. Freeman and Company, San Francisco, (1982).

[Mat.82] **T. Matsuyama, K. Saburi, M. Nagao**,
"A Structural Analyzer for Regularly Arranged Textures",
 Comput. Vision Graphics Image Process., Vol. 18, pp. 259-278, (1982).

[Mat.83] **T. Matsuyama, S.I. Muria, M. Nagao**,
"Structural Analysis of Natural Textures by Fourier Transformation ",
 Comput. Vision Graphics Image Process., CVGIP, Vol. 24, p. 347-362, (1983).

[Mat.90] **P. Mathieu, M. Barlaud, M. Antonini**
"Compression d'image par Transformée en Ondelette et Quantification Vectorielle",
 Traitement du signal, Vol. 7, No. 2, p. 101-115, (1990)

[Mey.87] **Y. Meyer, S. Jaffard, O. Rioul**,
"L'analyse par Ondelettes",
 Pour la Science, p. 28-37, Septembre (1987).

[Mey.90] **Y. Meyer**
"Ondelettes et Opérateurs 1 : Ondelettes "
 Ed. Hermann, Paris, (1990).

[Ng.92] **I. Ng, T. Tan, J. Kittler**,
"On Local Linear Transform and Gabor Filter Representation of Texture "
 Proc. IEEE IAPR, p. 627-631, (1992).

[Olz.92] **L.A. Olzak, J.P. Thomas**,
"When Orthogonal Orientations are not Processed Independently",
 Vision Research, 31, p. 51-57, (1992).

[Pel.84] **S. Peleg, J. Naor, R. Hartley, D. Avnir**,
"Multiple Resolution Texture Analysis and Classification",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI-6, p. 518-523, July, (1984).

[Pen.84] **A.P. Pentland**,
"Fractal-based Description of Natural Scenes",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI-6(6), p. 661-674, (1984).

[Per.92] **P. Perez, F. Heitz**,
"Une Approche Multiéchelle à l'Analyse d'Images par Champs Markoviens ",
 Traitement du Signal, Vol. 9, No 6, p. 459-466, (1992).

[Pie.81] **M. Pietikäinen, A. Rosenfeld**,
"Image Segmentation by Texture Using Pyramid Node Linking ",
 IEEE Trans. on Syst., Man and Cybern. Vol. SMC-11, No.12, p. 822-825, December (1981).

[Pie.83] **M. Pietikäinen, A. Rosenfeld**,
"Experiments with Texture Classification using Averages of Local Pattern Matches ",
 IEEE Trans. Syst. Man Cybern., SMC-13, p. (1983).

[Mar.80-Pie.83]

[Pol.83] **D.A. Pollen, S.F. Ronner**,
"Visual Cortical Neurons as Local Spatial Frequency filter",
 IEEE Trans. on Syst. Man and Cybern., vol. 13, Sept (1983).

[Pos.87] **J.G. Postaire**,
"De L'image à la Décision : Analyse des Images Numériques et Théorie de la Décision"

Ed. Dunod Informatique, Paris, (1987).

[Pra.78] W.K. Pratt, O.D. Faugeras, A. Gagalowicz,
"Discrimination of Stochastic Texture Fields",
 IEEE Trans. on System Man and Cybernetic, Vol. SMC-8, n° 11, p. 796-804, (1978).

[Pra.91] W.K. Pratt,
"Digital Image Processing",
 Second Edition, John Wiley & Sons, New-York, (1991).

[Pro.91] P. Provent,
"Segmentation d'Images par Analyse Statistique de Textures : Application aux Images Echocardiographiques",
 Thèse de Doctorat, Université Paris XII, 24 (1991).

[Ree.90] T.R. Reed, H. Wechsler,
"Segmentation of Textured Images and Gestalt Organisation using Spatial/ Spatial-frequency Représentations",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI-12(1), p. 71-12, (1990).

[Ree.93] T.R. Reed, J. M. Hans Du Buf,
"A Review of Recent Texture Segmentation and Feature Extraction Techniques",
 CVGIP: Image Understanding, Vol. 57, No.3, p. 359-372, May (1993).

[Reu.95] P. Reuzé, A. Bruno, J. Bezy-Wendling, R. Collorec,
"Caractérisation d'Images Texturées basées sur les Statistiques d'Ordre Trois",
 15ème Colloque GRETSI, Juan-Les Pins, p. 645-648, 18 au 21 Sept. (1995)

[Rio.91] O. Rioul, M. Vetterli,
"Wavelets and Signal Processing,"
 IEEE Signal Processing Magazine, p.14-38, (1991).

[Rio.93a] O. Rioul,
"Ondelettes Régulières : Application à la Compression d'Images Fixes",
 Thèse de Doctorat, Telecom Paris, Mars (1993)

[Rio.93b] O. Rioul,
"A Discrete-time Multiresolution Theory,"
 IEEE Trans. Signal Processing, Vol. 41, No. 8, p.2591-2606, (1993).

[Ros.71] A. Rosenfeld, M. Thurston,
"Edge and Curve Detection for Visual Scene Analysis",
 IEEE Trans. on Computers, Vol. C-20, p. 562-569, (1971).

[Sal.95] E. Salari, Z. Ling,
"Texture Segmentation using Hierarchical Wavelet Decomposition",
 Pattern Recog. Vol. 28, No.12, p. 1819-1824, December (1995).

[Seb.89] A. Sebbar, A. Smolarz,
"Adaptative Markov Random Field for Texture Synthesis and Segmentation",
 In Proc. 6th Scand. Conf. on Image Anal., Oulu, Finland, pp. 1010-1017, June 9-22, (1989).

[Pol.83-Seb.89]

[Sen.93] L. Senhadji, G. Carrault, J.J. Bellanger, G. Passariello
Quelques Nouvelles Applications de la Transformée en Ondelettes
 Innovation Tech. Biol. Med., Vol. 14, No. 4, p. 389-403, (1993).

[Smi.86] M.J. Smith, D.P. Barnwell,
"Exact Reconstruction Techniques for Tree-Structured Subband Coders",
 IEEE Trans. ASSP 34, p. 434-441, (1986).

- [Smi.95] **J.R. Smith, S-F. Chang,**
"Frequency and Spatially Adaptative Wavelet Packets",
 IEEE Proc. ICASSP, Detroit, USA, p. 2233-2236, (1995).
- [Sto.90] **D.G. Stork, H.R. Wilson,**
"Do Gabor Functions Provide Appropriate Descriptions of Visual Cortical Receptive Fields?",
 J. Opt. Soc. Am. A, vol. 7(8), p. 1362-1373, (1990).
- [Tan.75] **S. Tanimoto, T. Pavlidis**
"A Hierarchical Data Structure for Picture Processing",
 Computer Graphics and Image Processing, Vol. 4, n°2, p. 104-119, Juin, (1975).
- [Tan.90] **T.N. Tan, A.G. Constantinides,**
"Texture Analysis based on Human Visual"
 Proc. IEEE Acoust. Speech Signal Proc. ICASSP, Albuquerque, p. 2137-2140, (1990).
- [Tan.95] **T.N. Tan,**
"Texture Edge Detection by Modelling Visual Cortical"
 Pattern Recognition Vol.28, No.9, p. 1283-1298, September (1995).
- [Tsa.92] **M.K. Tsatsanis, G.B. Giannalis,**
"Object and Texture Classification using Higher Orders Statistics",
 IEEE Trans. Pattern Analysis Mach. Intell., PAMI-14(7), p. 733-750, July, (1992).
- [Tur.86] **M.R. Turner,**
"Texture Discrimination by Gabor Functions"
 Biol. Cybern. Vol.55, p.71-82, (1986).
- [Uns.84] **M. Unser,**
"Description Statistique de Texture : Application à l'Inspection Automatique",
 Thèse de Docteur ès Sciences Techniques, EPFL, Lausanne, (1984).
- [Uns.86a] **M. Unser,**
"Sum and Difference Histograms for Texture Classification",
 IEEE Trans. on Pattern Analys. and Image Proc., Vol. 1(8), p. (1986).
- [Uns.86b] **M. Unser,**
"Local Linear Transforms for Texture Measurements ",
 Signal Processing, 11, p. 61-79, (1986).
- [Uns.89] **M. Unser, M. Eden**
"Multiresolution Feature Extraction and Selection for Texture Segmentation"
 IEEE PAMI Vol. 11, n°7, p. 717-728, July (1989)
- [Uns.95] **M. Unser,**
"Texture Classification and Segmentation using Wavelet Frames",
 IEEE Trans. on Image Processing, Vol.4, No.11, p. 1549-1560, November (1995).

[Sen.93-Uns.95]

- [Val.82] **R.L. De Valois, D.G. Albrectht, L.G. Thorell,**
"Spatial Frequency Selectivity of Cells in Macaque Visual Cortex",
 Vision Research, 22, p. 545-559, (1982).
- [Vau.95] **P. Vautrot, N. Bonnet,**
Application des Ondelettes Splines à la Segmentation d'Images Texturées. Comparaison avec les Filtres de Gabor,
 15ème Colloque GRETSI, Juan-Les-Pins, 18 au 21 Septembre, (1995).

-
- [Vet.92] **M. Vetterli, C. Herley,**
"Wavelets and Filters Banks : Theory and Design",
IEEE Trans. on Signal Processing, Vol.40, No.9, p. 2207-2232, (1992).
- [Vil.86] **F.M. Vilnrotter, R. Nevatia, F.E. Price,**
"Structural Analysis of Natural Textures",
IEEE Trans. Pattern Analysis Mach. Intell., PAMI-8, p. 79-89, (1986).
- [Wan.86] **R. Wang, A.R. Hanson, E.M. Riseman,**
"Texture Analysis Based on Local Standard Deviation of Intensity "
Proc. IEEE Comp. Vis. Pattern Rec CVPR, Miami Beach, June 22-26, p. 627-631, (1986).
- [Wat.87] **A.B. Watson,**
"Efficiency of a Model Human Image Code",
J. Opt. Soc. Am. A, vol. 4, p. 2401-2417, (1987).
- [Web.85] **M.A. Webster, R.L. De Valois,**
"Relationship between Spatial-frequency and Orientation Tuning of Striate-Cortex Cells",
J. Opt. Soc. Am. A, vol. 2, p. 1124-1132, July (1985).
- [Wic.92] **M.V. Wickerhauser,**
"Acoustic Signal Compression with Wavelet Packets",
in Wavelets : A Tutorial in Theory and Applicationc, C.K. Chui, Ed Academic Press, New-York, p.679-700, (1992)
- [Wil.91] **H.R. Wilson, R. Blake, D.L. Halpern**
"Coarse Spatial Scales Constrain the Range of Binocular Fusion on Fine Scales",
J. Opt. Soc. Am. A, vol. 8, no. 1, p. 229-236, January (1991).
- [Yum.94] **S.O. Yum, A.M. Landraud-Lamole, G.S. Stamon,**
"Une Méthode de Classification des Textures utilisant l'Information Dérivée de la Phase dans un Système de Filtrage Multicanaux"
Congrès AGI 94, Poitiers, p. 349-352, (1994)
-

Segmentation pyramidale de textures par décomposition en ondelettes

Résumé

Cette thèse présente une méthode de segmentation pyramidale d'images d'aspect texturé par décomposition en ondelettes. D'une façon générale, tout processus de segmentation de textures comporte une étape de caractérisation qui précède l'étape de segmentation proprement dite.

Notre méthode de caractérisation de textures est fondée sur l'interprétation des coefficients de détails de la décomposition en ondelettes, comme décrivant les variations locales des niveaux de gris de la texture autour de leur valeur moyenne. Cette interprétation conduit à la définition d'un nouvel attribut, le "facteur de forme", qui se calcule sur un nouveau voisinage de type pyramidal.

Pour éprouver la pertinence de ce nouvel attribut, deux algorithmes de classification supervisée de textures sont proposés. Le premier exploite la représentation multirésolution de l'analyse par ondelettes pour réaliser une caractérisation efficace des textures à plusieurs niveaux de résolution. Le second détermine pour chaque texture, un niveau maximal de la décomposition en ondelettes, en définissant un seuil sur les variances des images d'approximation. Ce niveau est utilisé pour réaliser une préclassification, avant que le processus de classification ne soit conduit à son terme dans chaque classe constituée. Les résultats satisfaisants obtenus par ces deux algorithmes ont validé la méthode de caractérisation proposée et démontré la pertinence de l'attribut facteur de forme.

La méthode de segmentation présentée, tire profit de la forme pyramidale du voisinage sur lequel le facteur de forme est calculé. Cette forme constitue un véritable atout pour un processus de segmentation qui se fonde sur une représentation multirésolution des données. Elle permet de réduire sensiblement le nombre de pixels de l'image d'attributs représentant de zones inter-régions et minimise de ce fait, l'ambiguïté sur la localisation précise des frontières. De plus, les dimensions variables du voisinage pyramidal en fonction du niveau de résolution, suggèrent l'élaboration d'un processus de segmentation évoluant suivant une stratégie du plus grossier au plus fin. Les meilleurs prototypes de chaque classe de texture sont déterminés à l'aide d'un algorithme qui met en oeuvre un classificateur flou. Ces prototypes permettent de réaliser une segmentation grossière au niveau de résolution le plus élevé. Cette segmentation primaire est progressivement affinée lors de la descente de la pyramide en dirigeant le processus de segmentation vers les zones de singularités (zones frontalières ou bruitées). La méthode proposée fournit les meilleurs résultats et réalise une segmentation sans recouvrement des régions lorsque la décomposition de l'image est effectuée avec l'ondelette de Haar.

Mots clés :

Segmentation, Textures, Ondelettes, Voisinage pyramidal, Facteur de forme, Classification floue, Analyse multirésolution.