

# Méthodes d'analyse de textures 2D et solides

L'analyse des textures est un sujet ancien de l'analyse d'image et de la vision par ordinateur. Les techniques existantes sont utilisées pour identifier une texture quelconque ou diviser une image texturée en plusieurs régions comportant des caractéristiques identiques. Nous les retrouvons dans plusieurs domaines d'application comme l'analyse de scènes et de vidéos, l'imagerie médicale, l'industrie etc. Les recherches en analyse de textures se divisent en trois grandes catégories : la classification, permettant de reconnaître une texture donnée ; la segmentation, ayant pour objectif d'identifier les différentes classes de textures dans une même image ; et la synthèse de textures qui peut être également utilisée pour la compression d'images. Différents types d'images comprenant les images binaires, en niveaux de gris et en couleurs sont concernés par ces outils d'analyse. Selon les données à traiter, les techniques d'analyse de textures peuvent varier. Pour analyser une texture couleur, il est possible d'utiliser une méthode standard d'analyse de textures en niveaux de gris sur chacun des canaux RGB, mais certains auteurs préfèrent la combinaison d'informations afin de traiter des entités simples [Shi et Funt, 2007]. Cette thèse traite plus particulièrement des méthodes d'analyse de textures en niveaux de gris. Dans cet état de l'art, nous présentons tout d'abord les notions et la terminologie essentielles concernant les textures et leur analyse, avec notamment une discussion sur les propriétés perceptuelles concernant les textures. Par la suite, nous passons en revue les principaux descripteurs de textures utilisés sur les images texturées 2D et 3D.

## 1.1 Notion de textures

### 1.1.1 Les familles de textures

Dans la littérature, il est souvent question de plusieurs familles de textures avec notamment les textures déterministes, les textures stochastiques et les textures observables [Gool *et al.*, 1985] (figure 1.1) . Les textures déterministes se caractérisent par la répétition d'un motif géométrique aussi appelé "texton". Ce genre de textures peut donc facilement être synthétisé à partir d'un motif si on connaît son orientation et sa taille.

Au contraire, les textures stochastiques sont irrégulières. Elles obéissent à certaines lois statistiques et peuvent être considérées comme une réalisation d'un champ aléatoire bidimensionnel.

Une texture observable se décrit comme un mélange entre une texture stochastique et une texture déterministe. En effet, à l'intérieur de ce type de textures, nous pouvons trouver une répétition de motifs très similaires mais pas identiques.

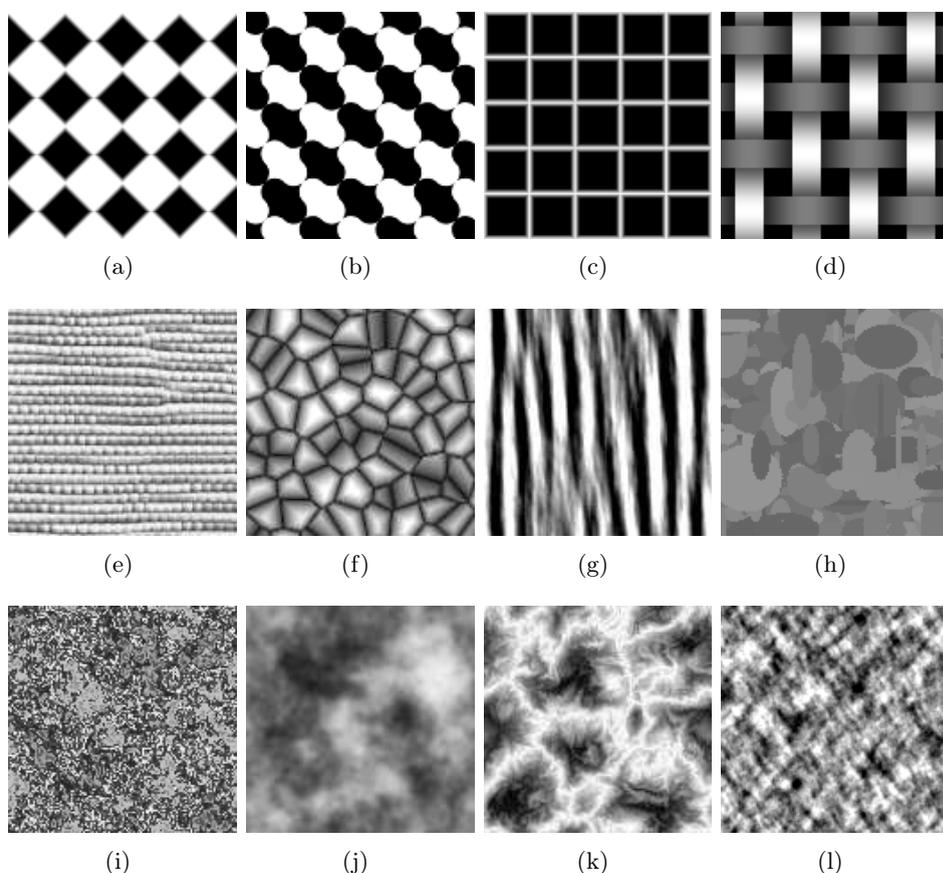


FIG. 1.1 – Exemple de familles de textures : [a-d] textures déterministes, [e-h] textures observables, [i-l] textures stochastiques

Certains auteurs n'utilisent pas le vocabulaire précédent et préfèrent parler de textures structurées (macro-textures) ou de textures aléatoires (micro-textures) [Unser, 1995]. Les textures structurées ont pour caractéristique la répétition de primitives à intervalles réguliers. Parmi elles, nous pouvons différencier les textures parfaitement périodiques (textures déterministes) et les textures dont les primitives ont subi quelques transformations (textures observables). Les textures aléatoires (textures stochastiques) se distinguent par un aspect plus fin (sable, herbe, etc.). Il est difficile d'en isoler une primitive de base et d'identifier une fréquence de répétition.

Il existe une grande diversité de textures. Certains ouvrages mettent à disposition des textures avec différentes propriétés, très utiles pour tester les algorithmes et les mé-

thodes. Parmi ceux-ci, la base présentée dans [Brodatz, 1966] et la base Meastex (<http://www.texturesynthesis.com/meastex/meastex.html>) proposent des collections de textures naturelles 2D en niveaux de gris souvent utilisées pour tester les algorithmes d'analyse et de classification. Ces bases pourront nous servir d'exemples pour la génération de nouvelles images ou pour tester certaines de nos méthodes.

### 1.1.2 Comment définir une texture ?

Donner une définition précise du terme "texture" est quelque chose de difficile. Il en existe un grand nombre et [Coggins, 1983] nous en donne un panel dont voici quelques exemples :

- "Une région dans une image contient une texture constante si un ensemble de statistiques locales ou autres propriétés de la fonction image sont constantes, varient faiblement ou sont approximativement périodiques.[Sklansky, 1978]"
- "Une image texturée est décrite par le nombre et les types de ses primitives tonales ainsi que leurs orientations spatiales. Elle ne peut pas être analysée sans une vue de la primitive tonale de référence. Pour certaines surfaces ayant un ton de gris flouté, il existe une échelle telle que la texture est inexistante. Au fur et à mesure que la résolution augmente, on observe une texture fine puis une texture grossière.[Haralick, 1979]"

Dans sa thèse, [Unser, 1984] donne une définition perceptuelle de la texture :

- "Une texture est une région d'une image pour laquelle il est possible de 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."

Certaines de ces définitions décrivent les textures de façon perceptuelle, d'autres se focalisent plus sur une stabilité de mesures statistiques à l'intérieur d'une région ou sur la méthode de génération utilisée. La plupart d'entre elles ne sont pas satisfaisantes dans le sens où elles proposent une description soit trop générale et donc imprécise, soit trop restrictive pour s'adapter à la diversité des cas rencontrés. Une texture est donc quelque chose d'abstrait, difficile à définir quantitativement de façon précise. Cependant une texture peut presque toujours être décrite par des termes, des adjectifs souvent dépendants du domaine concerné (images médicales, images naturelles, matériaux, etc.) mais plus faciles à comprendre et donc plus faciles à manipuler.

### 1.1.3 Perception visuelle des textures

Julesz a étudié de façon poussée la perception visuelle des textures [Julesz, 1962, Julesz, 1971]. Afin de pouvoir différencier un ensemble de textures, il se concentre sur l'analyse des statistiques spatiales des niveaux de gris d'une image. Ses travaux se sont donc beaucoup orientés sur la comparaison de textures selon leur ordre statistique. Rappelons que les statistiques de premier ordre mesurent la probabilité d'observer un niveau de gris à un emplacement aléatoire d'une image alors que les statistiques du deuxième ordre déterminent la probabilité d'observer une paire de niveaux de gris selon une orientation et

une distance.

[Julesz, 1962] donne une description de la perception visuelle humaine des textures en affirmant qu'elle utilise les statistiques d'ordre 2 des niveaux de gris. Cependant des contre-exemples ont été trouvés par [Gagalowicz, 1983] et Julesz lui-même. En effet, il existe des textures ayant des moments d'ordres 1 et 2 identiques. De plus, les statistiques d'ordre 2 décrivent bien souvent des processus à distribution gaussienne ; or [Coroyer, 1996] montre l'aspect non-gaussien des textures naturelles. Malgré ces restrictions, les approches comme les matrices de cooccurrences [Haralick *et al.*, 1973, Haralick, 1979] sont encore très utilisées pour leurs simplicité de mise en oeuvre et leur bons résultats. Julesz affirme également qu'il est difficile de différencier deux textures si leurs statistiques de second ordre sont identiques. Dans la figure 1.2(a), nous avons deux régions texturées ayant un second ordre statistique identique et nous pouvons voir qu'il est très difficile de les discerner. Cependant la figure 1.2(b) présente deux textures de second ordre statistique identiques et les deux régions sont bien visibles.

Pour expliquer cela, Julesz propose la théorie des textons, ceux-ci étant des symboles de base pouvant être décrits par leur colinéarité, leur terminaison (points de fond d'un segment), etc. Dans la figure 1.2(a), le nombre de terminaisons entre les deux régions est identique alors que dans la figure 1.2(b), il y a trois terminaisons pour la région supérieure et quatre pour la région inférieure, ce qui expliquerait pourquoi elles sont plus facilement discernables. Récemment, [Tyler, 2004] suggère que les statistiques globales ne devraient pas être le seul domaine de recherche dans la perception des textures. En effet, il attache plus d'importance aux mécanismes utilisés par les humains pour comparer différents exemples de textures. Concernant le processus de visualisation, des recherches ont été menées sur l'étude des cerveaux de macaques qui réagissent de façon similaire à celle des cerveaux humains. [Valois *et al.*, 1982] enregistrent la réponse des cellules du cortex visuel du macaque à un signal sinusoïdal selon plusieurs fréquences et orientations et en concluent que ces cellules sont réglées pour de petites variations de fréquences et d'orientations. Ces recherches ont motivé l'utilisation de filtres comme par exemple les filtres de Gabor [Turner, 1986, Jain et Farrokhnia, 1990, Higgins et Wakeley, 1994, Shen et Bai, 2008] qui permettent une bonne approximation des caractéristiques de cellules dans le cortex visuel de certains mammifères.

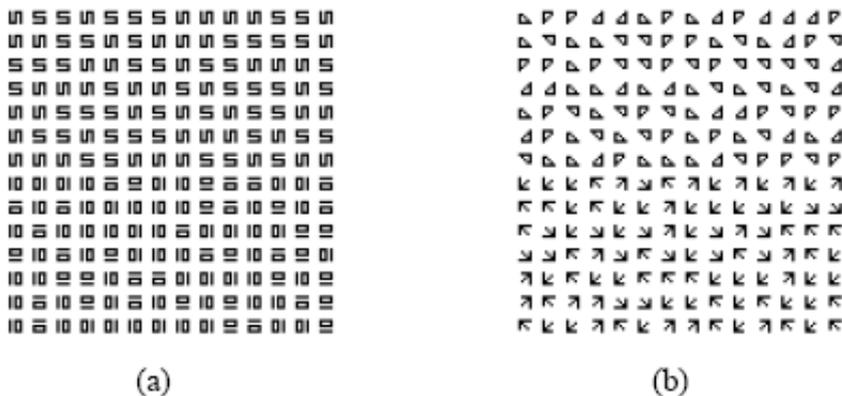


FIG. 1.2 – Images contenant des textures de second ordre statistique identiques

De nombreuses études ont été et sont encore menées pour essayer d'expliquer les capacités du système visuel humain à discerner les textures. La vision humaine offre des performances remarquables même en l'absence d'un contexte favorable. En effet, nous sommes capables d'identifier, de reconnaître une texture ayant subi de fortes distorsions géométriques. Le système visuel interprète la déformation de la texture non comme une perturbation des propriétés géométriques du champ texturé, mais comme une variation de la géométrie qui supporte la texture. Toute la question réside alors dans la possibilité ou non de produire un système informatique ayant une interprétation proche de celle du système visuel humain.

### 1.1.4 Propriétés de perception des textures

Dans [Brady, 2003], l'auteur décrit deux propriétés qui sont essentielles dans la façon de percevoir une texture : la régionalité et la résolution.

La régionalité correspond au voisinage autour d'un pixel/voxel. La figure 1.3 présente plusieurs régionalités d'une même texture avec tout d'abord en (a) une régionalité de 1. Dans cet exemple, nous ne pouvons pas encore parler de textures puisque la région est décrite par un unique pixel. En effet, la notion de textures apparaît dès lors où une région est composée d'un pixel/voxel et d'un voisinage non nul qui lui est associée. La figure 1.3(b) présente un exemple de régionalité égale à  $3^2$ . Si nous la comparons avec la figure 1.3(c) (régionalité égal à  $9^2$ ), celle-ci ne contient qu'une partie des motifs présents dans la texture. L'information que l'on peut obtenir sur une texture dépend donc très fortement de la taille du voisinage que l'on considère. Il est donc difficile de caractériser correctement une texture si la régionalité n'est pas choisie de façon adéquate.

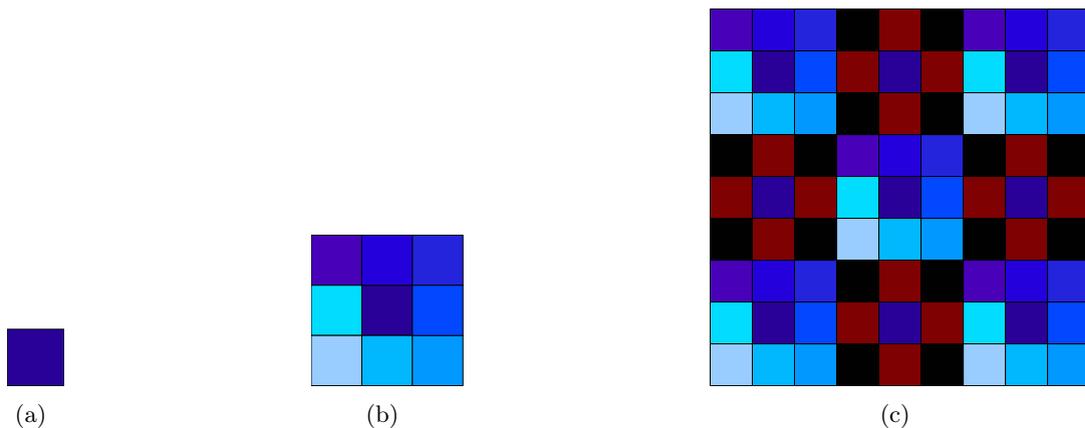


FIG. 1.3 – Information sur une texture selon plusieurs régionalités

La résolution est également primordiale en analyse de textures dans le sens où une texture peut apparaître comme différente selon le niveau de précision auquel celle-ci est observée. A forte résolution, nous caractérisons ce qui constitue les primitives de bases d'une texture (perle, granulé de sable, ...). A faible résolution, la texture peut apparaître comme une juxtaposition des motifs de base plus grossiers. Selon l'échelle d'observation, on aura des caractéristiques différentes mais complémentaires. Ce principe est notamment

utilisé par les ondelettes ou les fractales multirésolutions. La figure 1.4 illustre bien ce phénomène. En nous éloignant, l'information concernant la texture aux résolutions hautes est perdue au profit des informations textures basses résolutions.

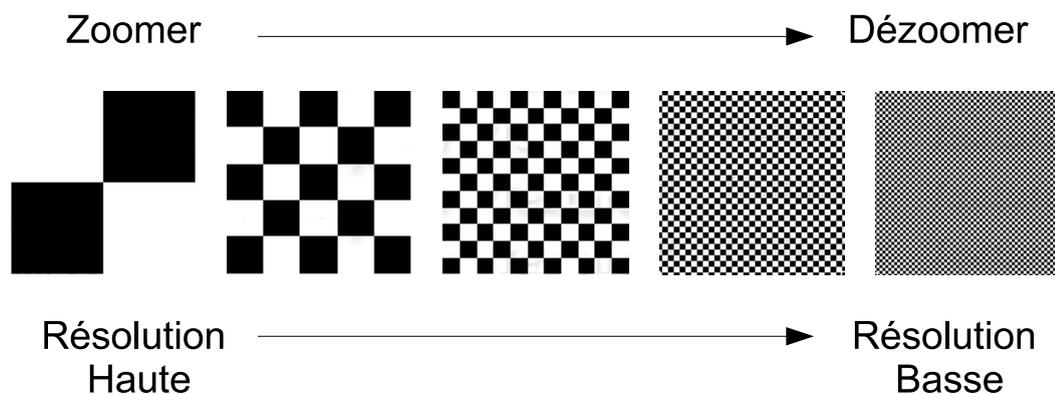


FIG. 1.4 – Représentation d'une texture selon plusieurs résolutions

Nous pouvons conclure qu'une bonne méthode d'analyse de textures doit être capable de donner des informations sur une texture à plusieurs niveaux de précision. Dans [Unser, 1995], l'auteur affirme qu'une méthode d'analyse multi-niveaux est plus performante qu'une méthode d'analyse limitée à une seule résolution. Les techniques de multirésolution permettent donc d'extraire des caractéristiques à des échelles variées mais nous verrons aussi que leur utilisation peut se justifier pour d'autres raisons. Par exemple, elles peuvent être utilisées pour réduire la complexité des traitements, en analysant une image à une résolution plus faible ou bien pour initialiser une segmentation [Jung, 2006, Aldasoro et Bhalerao, 2007]. Parmi les méthodes utilisant les propriétés de multirésolution, nous retrouvons les méthodes d'analyse fréquentielle avec notamment les filtres de gabor ou ondelettes [Knutsson et Granlund, 1983, Jain et Farrokhnia, 1990, Mallat, 1989], mais aussi les méthodes pyramidales [Samet, 1984, Spann et Wilson, 1985, Wilson et Spann, 1988].

## 1.2 Méthodes d'extraction de caractéristiques de textures

Dans [Tuceryan et Jain, 1998], les auteurs présentent un état de l'art sur les méthodes d'analyse de textures 2D à l'intérieur duquel ils proposent une classification de ces techniques selon quatre familles principales :

- les méthodes statistiques
- les méthodes géométriques
- les méthodes basées sur un modèle
- les méthodes empruntées au traitement du signal

Afin d'introduire dès à présent la philosophie des méthodes que nous présentons dans la suite, nous avons choisi de produire une classification différente des méthodes existantes. Nous présentons dans un premier temps les méthodes proposant une analyse basée sur

l'extraction des motifs d'une texture. Les méthodes basées sur une analyse fréquentielle regroupant principalement la transformée de Fourier, les filtres de Gabor et la transformée en ondelette sont ensuite présentées. Celles-ci permettent notamment de mettre en évidence l'importance de la multirésolution. Dans une troisième sous-section, nous abordons les méthodes basées sur un modèle, qui permettent l'introduction de connaissances *a priori*. Une quatrième partie nous permet de présenter des méthodes qui proposent une analyse spatiale de l'intensité des niveaux de gris. Enfin, nous passons en revue les propositions de certains auteurs qui tentent de déployer des méthodes calquées sur le système visuel humain afin de décrire les textures de manière plus efficace.

## 1.2.1 Les méthodes basées sur une analyse des motifs

### 1.2.1.1 Le spectre de textures

Le spectre de textures est une approche présentée dans [Wang et He, 1990]. Cette méthode propose de décomposer une image en un ensemble d'unités appelées "unités de textures". Chacune d'elles se compose de  $n$  éléments ayant pour valeur 0, 1 ou 2. Pour un voisinage de pixels  $3 \times 3$  (8 voisins), il est possible d'obtenir jusqu'à  $3^8$  soit 6561 unités de textures différentes.

Étant donné une matrice de pixels de taille  $3 \times 3$ , cette méthode considère un ensemble d'éléments  $V = V_0, V_1, \dots, V_8$  avec  $V_0$  l'intensité du pixel central et  $V_1, \dots, V_8$  l'intensité des pixels voisins. Les unités de textures correspondantes sont représentées par l'ensemble  $TU = E_1, E_2, \dots, E_8$ .

$$E_i = \begin{cases} 0 & \text{si } V_i < V_0 \\ 1 & \text{si } V_i = V_0 \\ 2 & \text{si } V_i > V_0 \end{cases} \quad (1.1)$$

avec  $i = 1, 2, \dots, 8$

A cette unité de textures est alors associé un label calculé à partir de la formule suivante :

$$N_{TU} = \sum_{i=1}^8 E_i 3^{i-1}$$

avec  $N_{TU}$  le label de l'unité de textures et  $E_i$  l'élément  $i$  de l'unité de textures.

A partir du spectre de textures, [He et Wang, 1991] proposent de calculer un certain nombre de caractéristiques comme la symétrie noir et blanc, la symétrie géométrique, le degré de direction, l'orientation, la symétrie centrale. Les formules permettant d'obtenir ces valeurs sont les suivantes :

#### Symétrie noir et blanc

$$BWS = \left[ 1 - \frac{\sum_{i=0}^{3279} |S(i) - S(6560 - i)|}{\sum_{i=0}^{6560} S(i)} \right] * 100 \quad (1.2)$$

où  $S(i)$  est la fréquence d'apparition de l'unité de textures  $i$  dans le spectre de textures.

### Symétrie géométrique

$$GS = \left[ 1 - \frac{1}{4} * \sum_{j=1}^4 \frac{\sum_{i=0}^{6560} |S_j(i) - S_{j+3}(i)|}{2 * \sum_{i=0}^{6560} S_j(i)} \right] * 100 \quad (1.3)$$

où  $S_j(i)$  est la fréquence d'apparition de l'unité de textures  $i$  dans le spectre de textures prenant en compte l'ordre de direction  $j$ .

Illustration de l'ordre de direction :

$E_1$	$E_2$	$E_3$	.....	$E_6$	$E_7$	$E_8$
$E_8$		$E_4$		$E_5$		$E_1$
$E_7$	$E_6$	$E_5$		$E_4$	$E_3$	$E_2$
Ordre 1.				Ordre 4.		

### Degré de direction

$$DD = \left[ 1 - \frac{1}{6} * \sum_{m=1}^3 \sum_{n=m+1}^4 \frac{\sum_{i=0}^{6560} |S_m(i) - S_n(i)|}{2 * \sum_{i=0}^{6560} S_m(i)} \right] * 100 \quad (1.4)$$

où  $m$  et  $n$  représentent un ordre de direction.

### Caractérisation de l'orientation

#### 1. Micro - structure horizontale

$$MHS = \sum_{i=0}^{6560} S(i) * HM(i) \quad (1.5)$$

où  $HM(i) = P(a, b, c) * P(e, f, g)$ , la mesure horizontale de l'unité de textures  $i$ .

#### 2. Micro - structure verticale

$$MVS = \sum_{i=0}^{6560} S(i) * VM(i) \quad (1.6)$$

où  $VM(i) = P(a, h, g) * P(c, d, e)$ , la mesure verticale de l'unité de textures  $i$ .

#### 3. Micro - 1<sup>ère</sup> diagonale de la structure

$$MDS1 = \sum_{i=0}^{6560} S(i) * DM1(i) \quad (1.7)$$

où  $DM1(i) = P(h, a, b) * P(d, e, f)$ , la mesure de la première diagonale de l'unité de textures  $i$ .

4. Micro - 2<sup>nde</sup> diagonale de la structure

$$MDS2 = \sum_{i=0}^{6560} S(i) * DM2(i) \quad (1.8)$$

où  $DM1(i) = P(b, c, d) * P(f, g, h)$ , la mesure de la seconde diagonale de l'unité de textures  $i$ .

où  $P(a, b, c), P(e, f, g), P(a, h, g), P(c, d, e), P(h, a, b), P(d, e, f), P(b, c, d)$  et  $P(f, g, h)$  est le nombre d'éléments ayant la même valeur parmi  $E_a, E_b, E_c$ , etc. Ci-dessous, une illustration de cette correspondance :

$E_a$	$E_b$	$E_c$
$E_h$		$E_d$
$E_g$	$E_f$	$E_e$

### Symétrie centrale

$$CS = \sum_{i=0}^{6560} S(i) * K(i)^2 \quad (1.9)$$

où  $K(i)$  est le nombre de paires possédant la même valeur entre les éléments  $(E_a, E_e), (E_b, E_f), (E_c, E_g)$  et  $(E_d, E_h)$ .

Afin d'analyser la répartition des motifs, [Al-Janobi, 2001] propose une nouvelle méthode d'analyse de textures, appelée *Cross diagonal texture Matrix (CDTM)*, utilisant les propriétés de la GLCM et des spectres de textures. Les huit éléments de l'unité de textures obtenus à partir d'un voisinage de  $3 \times 3$  pixels sont divisés en deux groupes contenant chacun quatre éléments comme illustré par la figure 1.5. Ces deux groupes sont deux unités de textures appelées "unité de textures en croix"(CTU) et "unité de textures en diagonale" (DTU). Comme dans la méthode du spectre de textures chacun des éléments des deux groupes se voit affecter une des trois valeurs 0, 1, ou 2. Il est alors possible de calculer un label de textures pour chacune des unités de textures CTU et DTU. Voici comment les calculer :

$$N_{CTU} = \sum_{i=1}^4 E_{Ci} \cdot 3^{i-1} \quad (1.10)$$

$$N_{DTU} = \sum_{i=1}^4 E_{Di} \cdot 3^{i-1} \quad (1.11)$$

où  $N_{CTU}$  est le label de l'unité de textures CTU,  $N_{DTU}$  est le label de l'unité de textures DTU,  $E_{Ci}$  est le  $i^{eme}$  élément appartenant à l'unité de textures CTU,  $E_{Di}$  est le  $i^{eme}$  élément de l'unité de textures DTU.

Une matrice de cooccurrence est alors construite à partir des paires de label ( $N_{CTU}, N_{DTU}$ ). Afin de caractériser une texture, des caractéristiques sont calculées sur cette matrice tout comme dans la méthode des matrices de cooccurrence des niveaux de gris d'Haralick (voir section 1.2.4.3).

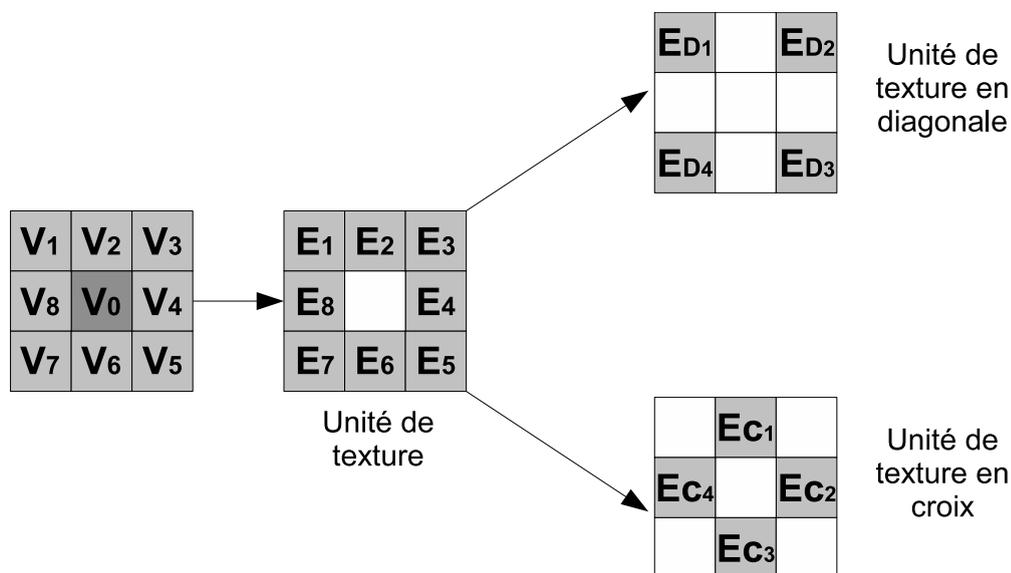


FIG. 1.5 – Construction des unités de textures en croix et en diagonale avec un voisinage de  $3^2$  pixels.

### 1.2.1.2 Méthode des motifs binaires (LBP)

Cette méthode, introduite par [Ojala *et al.*, 1996] est fortement inspirée par la méthode précédente. L'un des intérêts de cette méthode est de pouvoir générer un nombre d'unités de textures plus restreint. Pour un voisinage de 8 pixels, il est possible d'obtenir  $2^8 = 256$  unités de textures différentes contre 6561 pour la méthode du spectre de textures. Cette méthode est utilisée avec succès dans divers domaines comme la détection ou la reconnaissance faciale [Ahoen *et al.*, 2006, Zhao et Pietikäinen, 2007]. En terme d'efficacité discriminante, cette méthode offre de bonnes performances et contient des informations structurelles et statistiques. L'opérateur LBP (local binary pattern) peut être invariant aux rotations et changements d'échelles et offre une grande tolérance face aux changements de luminosité. De plus, c'est une méthode de faible complexité qui rend possible l'analyse d'images 2D en temps réel.

L'opérateur LBP propose d'assigner un code binaire à un pixel en fonction de son voisinage. La valeur d'intensité d'un pixel central est utilisée pour seuiller ses pixels voisins afin de générer un motif binaire (Figure 1.6). Les pixels de ce motif binaire sont alors multipliés par des poids et sommés afin d'obtenir un code LBP.

Dans [Ojala *et al.*, 2001, Ojala et Pietikäinen, 2002] des extensions de cette méthode sont proposées. Celles-ci permettent :

- un choix sur le voisinage afin de calculer un code LBP

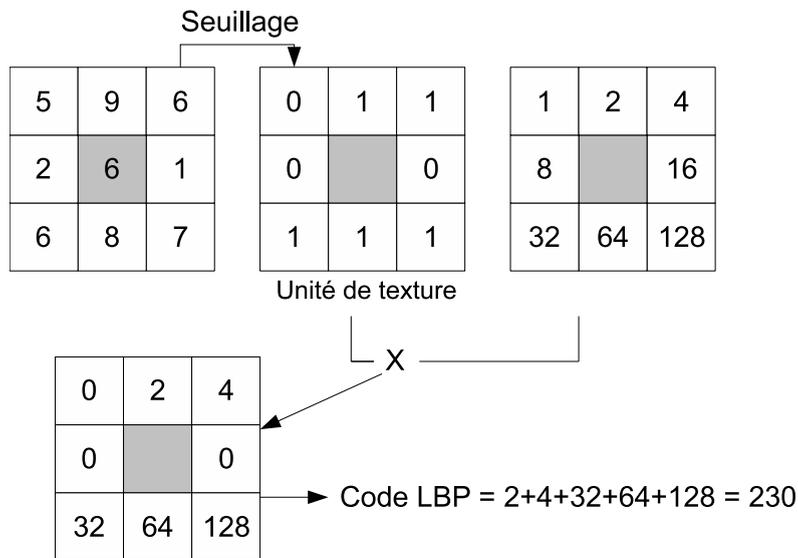


FIG. 1.6 – Construction d'un motif binaire et calcul du code LBP

– une invariance à la rotation

Un voisinage pour un pixel central est réparti sur un cercle et construit à partir de deux paramètres : le nombre de voisins sur le cercle et un rayon pour définir une distance entre un pixel central et ses voisins. Définissons une texture :  $T = t(g_c, g_0, \dots, g_{P-1})$ ,  $g_c$  correspond à la valeur de niveau de gris du pixel central et  $g_p$ , avec  $p = 0, \dots, P-1$ , correspond au niveau de gris de  $P$  pixels espacés régulièrement sur un cercle de rayon  $R$ . Si les coordonnées de  $g_c$  sont égales à  $(0, 0)$ , alors les coordonnées de  $g_p$  sont données par  $(-R \sin(2\pi p/P), -R \cos(2\pi p/P))$ . Comme nous pouvons le voir à la figure 1.7, les coordonnées d'un voisin ne sont pas forcément situées au centre d'un pixel. Dans ce cas, le niveau de gris est calculé par l'intermédiaire d'une interpolation. Voici une représentation graphique de  $T$  faisant varier les paramètres  $P$  et  $R$  :

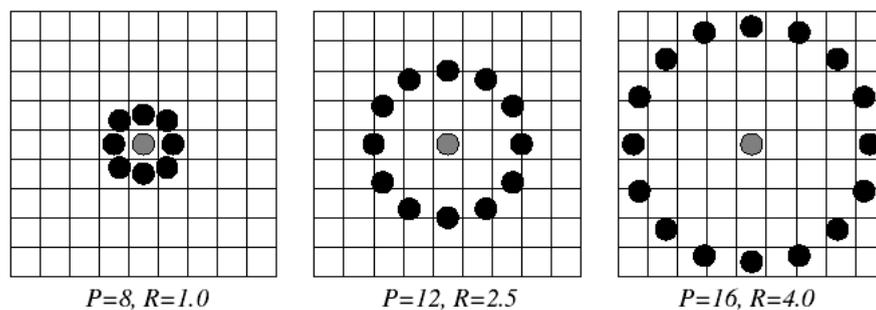


FIG. 1.7 – Représentation du pixel central  $g_c$  et de son voisinage  $g_p$

Comme dans la méthode LBP classique, le pixel central n'est pas utilisé pour la caractérisation des textures. En effet, indépendamment du voisinage  $g_p$ , ce pixel décrit uniquement une intensité lumineuse ce qui n'est pas forcément utile [Ojala et Pietikäinen, 2002].

$T$  peut alors s'écrire  $T = t(g_0, \dots, g_{P-1})$ . Par la suite,  $g_c$  est utilisé comme un seuil de la façon suivante :

$$T = t(s(g_0 - g_c), \dots, s(g_{P-1} - g_c))$$

avec  $s(x)$  égal à 1 si  $x$  positif et  $s(x)$  égal à 0 si  $x$  est strictement négatif.

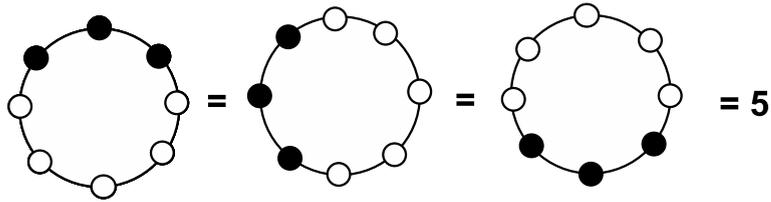
En multipliant chaque fonction  $s(g_p - g_c)$  par un poids de valeur  $2^p$ , on obtient alors un code binaire tel que :

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (1.12)$$

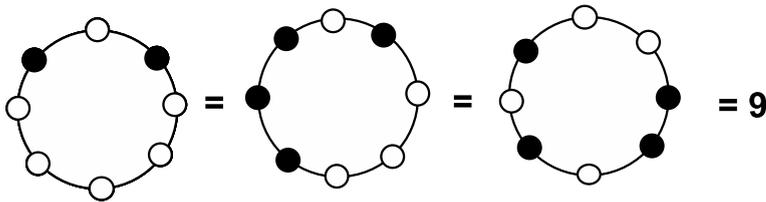
Pour supprimer l'effet de rotation, un code LBP ( $LBP_{P,R}^{ri}$ ) est défini en vue de retourner une valeur unique quelle que soit la rotation appliquée :

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i)\} \quad (1.13)$$

avec  $i = 0, 1, \dots, P - 1$ , et  $ROR(x, i)$  permettant de faire une rotation bit à bit de  $x$  et  $i$  fois. Avec cette méthode, 36 motifs différents sont considérés au lieu de  $2^P$ .



Motifs uniformes



Motifs non-uniformes

FIG. 1.8 – Exemples de motifs uniformes et non-uniformes

Pour une valeur  $R = 1$ ,  $LBP_{P,R}^{ri}$  est l'équivalent de  $LBPROT$  dans [Pietikäinen *et al.*, 2000]. Dans cet article, il est montré que  $LBPROT$  ne permet pas d'obtenir de bonnes performances. La principale raison est que l'occurrence des fréquences de certains des 36 motifs possibles varie fortement (instabilité). Pour résoudre ce problème, une mesure est effectuée sur les motifs afin de déterminer s'ils sont uniformes. Pour cela une fonction  $U$  retourne le nombre de transitions spatiales (passage de 0 à 1 ou de 1 à 0). Si le nombre de transitions est supérieur à deux, alors le motif est considéré comme

non-uniforme. L'opérateur appelé  $LBP_{P,R}^{riu2}$  est défini de la façon suivante :

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{si } U(LBP_{P,R}) \leq 2 \\ P + 1 & \text{sinon} \end{cases} \quad (1.14)$$

avec  $U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$  une mesure d'uniformité.

Concernant les motifs uniformes, l'opérateur  $LBP^{riu2}$  correspond à la somme des bits du voisinage. Au contraire si les motifs sont non-uniformes, l'opérateur est affecté à la valeur  $P + 1$  (Figure 1.8). Le nombre d'apparitions de chaque code est accumulé dans un histogramme de taille  $P + 2$ . Chaque tranche de cet histogramme donne une estimation de la probabilité d'apparition d'un motif dans une texture. Les textures sont caractérisées en utilisant directement l'histogramme des codes binaires.

Dans [Guo *et al.*, 2009], les auteurs proposent une variation de cette méthode en coordonnant une invariance à la rotation à la fois globale et locale. Dans un premier temps, ils utilisent la distribution LBP pour déterminer l'orientation principale d'une texture et l'utilisent afin d'aligner l'histogramme LBP. Dans la figure 1.8, le motif uniforme obtient le même code LBP quelle que soit son orientation. Le fait d'aligner l'histogramme LBP selon l'orientation principale (information globale) permet de différencier ces 3 motifs uniformes en leur affectant des codes binaires différents. Enfin, les auteurs proposent également un nouveau descripteur de textures LBP *variance* (LBPV) permettant de mesurer les informations locales de contraste.

### 1.2.1.3 La loi de Zipf

La loi de Zipf, également connue comme loi de puissance, a été déterminée de façon empirique par Zipf. Cette technique fonctionne de façon très similaire à la méthode LBP en proposant des mesures construites à partir de la fréquence d'apparition de motifs présents dans une texture. Dans [Caron *et al.*, 2007], les auteurs donnent une explication sur son utilisation à la fois en image et dans d'autres domaines d'application. A la base, cette méthode est utilisée pour l'étude de textes. En effet, selon la loi de Zipf, la fréquence d'occurrence de motifs dans un texte suit une loi de puissance. Si nous définissons un ensemble de motifs  $P_1, P_2, \dots, P_R$  et leurs fréquences d'apparition correspondantes  $N_1, N_2, \dots, N_R$  alors la fréquence d'apparition  $N_{\sigma(i)}$  du motif  $P_{\sigma(i)}$  de rang  $i$  est donnée par :

$$N_{\sigma(i)} = ki^{-\alpha} \quad (1.15)$$

avec  $k$  et  $\alpha$  deux constantes.

Durant ses travaux sur les textes naturels, [Zipf, 1935] a découvert un autre modèle de la loi puissance appelé loi de Zipf inverse. Selon cette loi, le nombre  $I$  de mots distincts avec une fréquence d'apparition  $f$  est donné par :

$$I(f) = af^{-b} \quad (1.16)$$

avec  $a$  et  $b$  deux constantes.

La loi de Zipf et la loi de Zipf inverse sont toutes les deux applicables à l'analyse d'image

[Caron *et al.*, 2007]. En analyse d'image 2D, les mots sont considérés comme des blocs adjacents de pixels de taille  $3^2$ . A l'intérieur des images, les niveaux de gris ne sont pas utilisés directement. Sans une étape de quantification, le nombre de motifs possibles serait bien trop important pour avoir une répartition statistique significative des fréquences (figure 1.9). Dans [Caron *et al.*, 2007], les auteurs proposent de requantifier l'image de la façon suivante :

$$c(x, y) = \text{int}\left[\frac{Ng(x, y)}{255}\right] \quad (1.17)$$

avec  $N$  le nombre de classes,  $g(x, y)$  l'intensité aux coordonnées  $(x, y)$  et  $c(x, y)$  la nouvelle valeur de niveaux de gris.

2	2	40
2	1	216
210	255	34

0	0	1
0	0	7
7	8	1

FIG. 1.9 – A gauche le motif original et à droite le motif requantifié avec  $N = 9$ .

De cette façon, les motifs d'une image sont répertoriés par un parcours effectué à l'aide d'un masque ( $3^2$  dans [Caron *et al.*, 2007]). Il est alors possible de compter la fréquence d'apparition des motifs (loi de Zipf) ou le nombre de motifs avec une fréquence  $f$  (loi de Zipf inverse). Ces deux lois sont généralement représentées à l'aide d'un graphique bi-logarithmique. L'aire ou la pente de la courbe obtenue sont deux caractéristiques qu'il est possible d'utiliser pour identifier une texture ou pour la détection d'objets dans une image [Caron *et al.*, 2007].

#### 1.2.1.4 Polygones de Voronoï

Dans [Tuceryan et Jain, 1990], les auteurs proposent d'utiliser les polygones de Voronoï pour segmenter les textures. Dans cette approche, ils cherchent à caractériser les textures à travers la distribution spatiale locale des germes qui constituent les différents polygones de Voronoï.

Si nous considérons un espace euclidien  $E$  avec  $S$  un ensemble fini de  $n$  points (germes) de  $E$ , alors nous appelons région de Voronoï ou cellule de Voronoï associée à un élément  $p$  de  $S$ , l'ensemble des points qui sont plus proches de  $p$  que de tout autre point de  $S$ .

$$Vor_s(p) = \{x \in E / \forall q \in S, d(x, p) \leq d(x, q)\} \quad (1.18)$$

Pour un ensemble de germes, le diagramme de Voronoï se construit en déterminant les médiatrices de chaque couple de germes. Un point d'une médiatrice appartient alors à une frontière de Voronoï s'il est équidistant d'au moins deux germes et s'il n'existe pas de distance plus faible entre ce point et un autre germe de l'ensemble.

Pour appliquer le maillage de Voronoï à une image en niveaux de gris, il faut pouvoir extraire des germes de cette image. [Tuceryan et Jain, 1990] proposent l'algorithme simple suivant :

1. Convoluer l'image avec le laplacien d'une gaussienne (LoG).
2. Sélectionner les maximums locaux dans l'image filtrée. Un pixel sera sélectionné parmi les maximums locaux si sa valeur est supérieure à six ou plus de ses huit voisins. L'image est ensuite binarisée pour ne conserver que ces mêmes maximums locaux.
3. Réaliser le diagramme de Voronoï en analysant les composantes connexes sur l'image binarisée. Chaque composante connexe est définie comme étant un germe pour le diagramme de Voronoï.

Par la suite, des caractéristiques sont calculées pour chaque cellule de Voronoï. Les germes possédant des caractéristiques identiques sont regroupés pour construire des régions de textures uniformes. Les moments de la surface des polygones de Voronoï sont des caractéristiques intéressantes permettant de refléter la distribution spatiale et les formes de germes dans une image texturée. Le  $(p + q)^{eme}$  ordre des moments de la surface d'une région  $R$  centrée sur un germe de coordonnées  $(x_0, y_0)$  est défini de la façon suivante :

$$m_{pq} = \int_R \int (x - x_0)^p (y - y_0)^q dx dy \quad (1.19)$$

A partir des moments, il est proposé dans [Tuceryan et Jain, 1990] de calculer les caractéristiques de textures spécifiées dans le tableau 1.1.

Caractéristiques	Méthodes de calcul
$f_1$	$m_{00}$
$f_2$	$\sqrt{\bar{x}^2 + \bar{y}^2}$
$f_3$	$\tan^{-1}(\bar{y}/\bar{x})$
$f_4$	$\left[ \frac{[(m_{20} - m_{02})^2 + 4m_{11}^2]^{0.5}}{[(m_{20} - m_{02})^2 + 4m_{11}^2]^{0.5} + m_{20} + m_{02}} \right]^{0.5}$
$f_5$	$\tan^{-1}(2m_{11}/(m_{20} - m_{02}))$

TAB. 1.1 – Tableau des caractéristiques calculées sur les polygones de Voronoï

Dans [Tuceryan et Jain, 1990], les polygones de Voronoï sont utilisés dans une problématique de segmentation de textures. Les régions ayant des propriétés de textures homogènes sont fusionnées pour produire une segmentation. Il serait intéressant d'évaluer cette méthode dans le cadre d'une classification en calculant des caractéristiques capables de décrire une texture dans sa globalité et non localement.

### 1.2.1.5 Extraction de textures par l'analyse de composantes connexes

La notion de composantes connexes est très utilisée dans l'analyse d'images binaires et dans l'analyse de documents. Elle est également répandue dans le domaine des textures et permet d'analyser les structures ainsi que les motifs binaires composant une texture. Dans [Mital *et al.*, 1994, Goyal *et al.*, 1994a, Goyal *et al.*, 1994b, Goyal *et al.*, 1995], les auteurs

proposent de caractériser de façon structurelle les textures. Pour cela, ils utilisent notamment la caractéristique de compacité d'un motif qui est calculée en 2D en faisant le quotient de son périmètre sur son aire. De nouvelles méthodes pour calculer le périmètre d'une composante connexe sont également proposées afin d'obtenir une caractéristique de compacité plus stable et de meilleure qualité. Cette caractéristique est invariante à tout type de transformation comme la translation, la rotation et les changements d'échelle [Zhang et Tan, 2002].

Récemment, une méthode de classification de textures utilisant la notion de composantes connexes a été proposée dans [Shoshany, 2008]. Une texture en niveaux de gris est décomposée en une succession de textures binaires 2D ce qui permet une étude des motifs (Figure 1.10). Dans cette approche, l'ensemble des images binaires est construit en utilisant des binarisations globales effectuées avec différentes valeurs de seuil.  $k$  images binaires sont ainsi calculées en faisant varier le seuil de binarisation, avec  $k \in \{1, 256\}$ . Pour classifier les textures, Shoshany propose ensuite d'utiliser les trois caractéristiques suivantes, dont la complémentarité a été prouvée dans [Luo et Savakis, 2001] :

le nombre de composantes connexes,

l'hétérogénéité des motifs,

$$SW = - \sum S_i \ln S_i \quad (1.20)$$

avec  $S_i$  l'aire relative d'une composante connexe  $i$ ,

la dominance,

$$ED = \sum (S_i)^2 \quad (1.21)$$

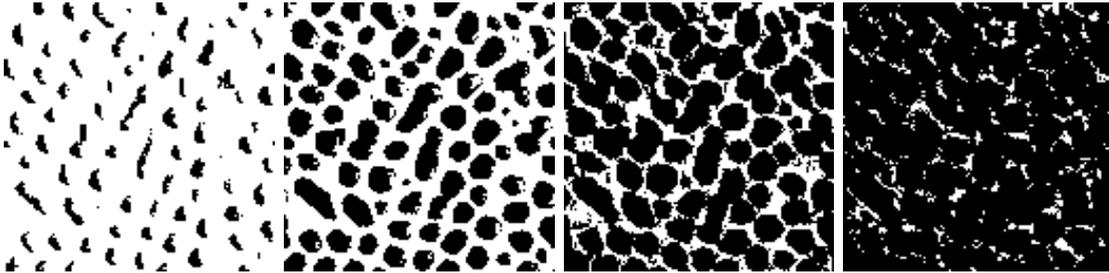


FIG. 1.10 – Représentation d'une suite de textures binaires 2D construite à partir de seuils de binarisation évoluant de façon croissante.

L'hétérogénéité des motifs représente leur niveau de fragmentation, et la dominance le nombre de motifs larges dans une distribution. Ces caractéristiques sont calculées sur l'ensemble des composantes des textures binaires, c'est à dire à la fois sur les motifs blancs (*foreground*) et les motifs noirs (*background*). A partir de ces attributs de textures, une technique de *clustering* hiérarchique [Borgatti, 1994] est utilisée pour analyser les dissimilarités entre les différents types de textures. D'un point de vue résultat, cette approche permet d'obtenir de bonnes performances en classification sur des images texturées et il pourrait être intéressant de l'adapter afin de pouvoir l'utiliser dans des problématiques de segmentation de textures.

### 1.2.1.6 Autres méthodes

De nombreuses autres méthodes permettent de caractériser une texture par l'analyse de ses structures et de ses primitives de base. Dans [Zucker, 1976], il est proposé un modèle permettant à partir d'une texture idéale, composée de primitives régulières, de synthétiser une texture observable. Cette texture idéale est considérée comme un maillage représenté par un graphe isomorphe. Pour générer la texture observée, des transformations déterministes ou stochastiques sont appliquées sur ce graphe. Il est alors possible de mesurer la compatibilité entre la texture idéale et la texture observée. [Fu, 1982] décrit la texture à l'aide d'une grammaire d'arbre. Dans cette approche, l'image est divisée en fenêtres/primitives qui peuvent à leur tour être divisées jusqu'à obtenir un élément de base. Ces primitives définissent une grammaire à laquelle sont associées une ou plusieurs règles de placement. De cette façon, une texture est vue comme une phrase dans le langage défini par la grammaire considérée. Cette méthode est adaptée à l'analyse mais également à la synthèse de textures. Dans [Voorhees et Poggio, 1987], les auteurs affirment que les taches jouent un rôle important dans la perception de la texture. Pour extraire les taches de l'image, ils utilisent le filtre *Laplacian of Gaussian* (LoG) à plusieurs échelles. [Varma et Zisserman, 2004] proposent de déterminer des textons à partir d'un banc de filtres. Ce banc de filtres est construit en utilisant des filtres gaussien et LoG selon plusieurs orientations et échelles. Les réponses de filtres par classe de textures sont regroupées et classifiées par l'algorithme des *K-Means*. Les centroïdes de chaque classe ainsi trouvés sont considérés comme des textons. Une texture contient  $K$  textons et peut être caractérisée par la proportion des pixels labellisés comme un texton particulier.

### 1.2.1.7 Bilan

L'objectif de ce type de méthodes est d'identifier les structures de base qui se répètent périodiquement pour constituer une texture. Le principe le plus courant est ensuite de comptabiliser le nombre d'occurrences de chacun de ces motifs ou de les regrouper en régions pour classifier ou segmenter les textures. Le problème réside alors dans la sélection des motifs pertinents permettant d'obtenir une caractérisation robuste. Ce type de méthode paraît bien adapté au traitement de textures déterministes et observables constituées de motifs répétitifs. Le problème est de pouvoir gérer la variabilité que peut subir chaque motif. Nous avons aussi mentionné la difficulté d'utiliser les bonnes régionalité et résolution d'observation qui permettent d'identifier correctement les motifs. Notons enfin que ce type de méthode peut fournir une description des textures facilement appréhendable par un humain, nous le retenons donc comme pertinent vis à vis de nos objectifs.

## 1.2.2 Les méthodes fréquentielles

Ces méthodes, souvent utilisées en traitement du signal, permettent d'analyser une texture en identifiant les différentes fréquences qui la composent. Parmi elles, on retrouve notamment la transformée de Fourier, les filtres de Gabor ainsi que les ondelettes. De façon générale, la transformation linéaire d'une image  $f$  est obtenue en calculant la

corrélation de  $f$  et d'une fonction  $\psi_\lambda$  :

$$\int_{-\infty}^{+\infty} f(x, y) \bar{\psi}_\lambda(x, y) dx dy \quad (1.22)$$

avec  $\lambda$  un ensemble de paramètres propre à  $\psi$  et  $\bar{(\ )}$  le complexe conjugué de la quantité en argument.

### 1.2.2.1 La transformée de Fourier

La transformée de Fourier [Bracewell, 1986] est une des méthodes utilisées pour l'analyse linéaire d'une image. La fonction  $\psi$  utilisée pour obtenir la transformée linéaire d'une image prend comme valeur :

$$\psi_\lambda(x, y) = e^{i2\pi(ux+vy)} \quad (1.23)$$

avec  $\lambda = (u, v)$  les fréquences à étudier. A partir de la transformée de Fourier, il est possible de calculer le spectre de puissance qui est défini comme le carré du module de la transformée de Fourier :

$$P(u, v) = F_R(u, v)^2 + F_I(u, v)^2 \quad (1.24)$$

avec  $F_R(u, v)$  et  $F_I(u, v)$  la partie réelle et la partie imaginaire de la transformée de Fourier. A partir du spectre de puissance, il est possible de calculer plusieurs caractéristiques pour décrire une image. Dans [Petrou et Sevilla, 2006], les auteurs donnent une explication sur la mesure de la périodicité d'une texture à partir du spectre de puissance. En considérant les opérateurs  $P_u(u) = \sum_v P(u, v)$  et  $P_v(v) = \sum_u P(u, v)$ , l'observation répétitive de pics démontre une certaine périodicité dans l'image. En utilisant le spectre de puissance, il est également possible de calculer la dimension fractale d'une image. Pour cela, il suffit de tracer une courbe des intensités selon une ligne du spectre de puissance. Formellement, une ligne  $L$  d'un spectre de puissance de taille  $(M, N)$  sera égale à  $P(u, v)$  avec  $u = a \cos \theta$ ,  $v = a \sin \theta$  et  $a = \{1, 2, 3, \dots, \min(M, N)\}$ . La dimension fractale est calculée à partir de la pente de la courbe.

La transformée de Fourier est un outil bien adapté à l'analyse des comportements harmoniques, cependant elle n'est pas appropriée à l'analyse de formes localisées spatialement. Si l'on veut conserver une analyse spectrale tout en ayant une localisation spatiale, il faut adopter une approche espace-fréquence en choisissant des fonctions  $\psi_\lambda$  qui soient à la fois oscillantes et de supports limités. On dit alors que les fonctions  $\psi_\lambda$  constituent des atomes espaces-fréquences.

### 1.2.2.2 Transformée de Fourier fenêtrée et fonction de Gabor

Motivé par la mécanique quantique, Gabor définit en 1946 [Gabor, 1946] les atomes temps-fréquence (espace-fréquence en dimension supérieure) comme des ondes ayant une propagation minimale dans l'espace temps-fréquence. Pour mesurer le contenu de l'information temps-fréquence, il propose de décomposer un signal à partir d'ondes atomiques

## 1.2. MÉTHODES D'EXTRACTION DE CARACTÉRISTIQUES DE TEXTURES

élémentaires. Pour cela la transformée de Fourier est légèrement modifiée par l'introduction d'une fenêtre  $g$ .

$$\psi_\lambda(x, y) = g(x - x_0, y - y_0) \exp(i(2\pi(u_0x + v_0y) + P)) \quad (1.25)$$

avec  $\lambda = (x_0, y_0, u_0, v_0)$ ,  $(u_0, v_0)$  définissant la fréquence spatiale de la sinusoïde, et  $P$  sa phase.

Si l'on exprime la fréquence spatiale à partir d'une coordonnée polaire selon une amplitude  $F_0$  et une direction  $\omega_0$ , on obtient :

$$u_0 = F_0 \cos \omega_0 \quad v_0 = F_0 \sin \omega_0 \quad (1.26)$$

L'énergie de  $\psi_\lambda$  est concentrée autour de  $(x_0, y_0)$  sur un espace de taille  $(\sigma_x, \sigma_y)$ .

La transformée de Fourier de la fonction  $\psi_\lambda$  est :

$$\Psi_\lambda(v_x, v_y) = \hat{g}(v_x - u_0, v_y - v_0) \exp(-i2\pi(x_0(v_x - u_x) + y_0(v_y - u_y) + P)) \quad (1.27)$$

Dans le domaine fréquentiel, l'énergie de  $\Psi_\lambda$  est localisée autour de  $(u_0, v_0)$  dans des intervalles de taille  $(\sigma_{v_x}, \sigma_{v_y})$ .

Lorsque  $g$  prend une forme gaussienne, on dit qu'on effectue une analyse de Gabor.  $g$  représente alors une enveloppe gaussienne qui peut s'écrire :

$$g(x, y) = K \exp(-\pi((\frac{(x - x_0)\cos\theta + (y - y_0)\sin\theta}{\sigma_x})^2 + (\frac{(y - y_0)\cos\theta - (x - x_0)\sin\theta}{\sigma_y})^2))$$

avec  $(x_0, y_0)$  le sommet de la fonction,  $\sigma_x$  et  $\sigma_y$  les écart-types de l'enveloppe gaussienne,  $K$  l'échelle de l'amplitude de l'enveloppe gaussienne,  $\theta$  l'angle de rotation de l'enveloppe gaussienne. La fonction complexe de Gabor s'écrit alors :

$$\psi_\lambda(x, y) = K \exp - \pi((\frac{(x - x_0)\cos\theta + (y - y_0)\sin\theta}{\sigma_x})^2 + (\frac{(y - y_0)\cos\theta - (x - x_0)\sin\theta}{\sigma_y})^2) \cdot \exp(j(2\pi(xF_0\cos\omega_0 + yF_0\sin\omega_0) + P)) \quad (1.28)$$

La figure 1.11 représente le filtre de Gabor dans le domaine fréquentiel mais aussi les différents paramètres agissant sur son orientation et sa forme, et qui mettent en évidence certaines fréquences. Dans la littérature, la sinusoïde complexe et l'enveloppe gaussienne sont souvent orientées dans la même direction avec  $\theta = \omega$ . Il arrive également que les valeurs des écart-types de l'enveloppe gaussienne soient choisies identiques ( $\sigma = \sigma_x = \sigma_y$ ). Enfin, certaines méthodes n'utilisent que la partie réelle qui est considérée comme la partie la plus intéressante et évitent d'utiliser le module qui n'est pas une opération linéaire. L'utilisation de cet outil demande une bonne connaissance du choix des paramètres. Dans un banc de filtres, la fonction de Gabor est utilisée avec plusieurs valeurs d'angles et de fréquences pour balayer au mieux les différentes fréquences présentes dans l'image à analyser. L'angle  $\theta$  prend ses valeurs entre 0 et  $2\pi$  séparées de façon équidistante (souvent tous les 30 ou 45 degrés). L'amplitude  $F_0$  peut s'exprimer en fonction de la longueur d'onde  $f_0$  avec  $F_0 = \frac{1}{f_0}$ . La réponse de la fréquence spatiale de la bande passante  $b$  du filtre de Gabor s'exprime en octave. Elle s'écrit :

$$b = \log_2 \frac{\frac{\sigma}{f_0} \pi + \sqrt{\frac{\ln 2}{2}}}{\frac{\sigma}{f_0} \pi - \sqrt{\frac{\ln 2}{2}}} \quad (1.29)$$

A partir de cette équation, on peut alors exprimer le rapport  $\frac{\sigma}{f_0}$  :

$$\frac{\sigma}{f_0} = \frac{1}{\pi} \sqrt{\frac{\ln 2}{2} \frac{2^b + 1}{2^b - 1}} \quad (1.30)$$

Une valeur de  $b = 1$  octave offre de bonnes performances. On obtient ainsi la relation  $\sigma = 0.56f_0$ . Dans le cas où  $\sigma_x$  et  $\sigma_y$  sont différents, il est conseillé dans [Xie *et al.*, 2005] la relation  $\sigma_y = 2\sigma_x = 1.12f_0$ .

De par leurs qualités à approximer les caractéristiques de cellules du cortex visuel des mammifères, ces filtres sont très utilisés en analyse d'image et notamment dans la caractérisation et la segmentation des textures [Fogel et Sagi, 1989, Bovik *et al.*, 1990, Farrokhnia et Jain, 1991, Jain et Farrokhnia, 1990, Higgins et Wakeley, 1994].

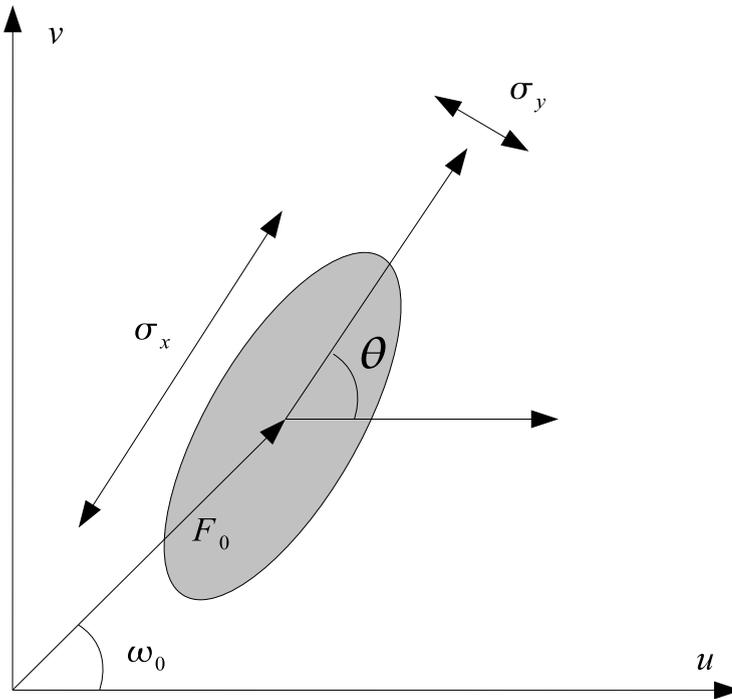


FIG. 1.11 – Représentation du filtre de Gabor et de ses paramètres dans le domaine fréquentiel

### 1.2.2.3 Les ondelettes

Tout comme les filtres de Gabor, les ondelettes permettent une bonne approximation du système visuel des mammifères. [Mallat, 1989] prétend qu'elles permettent une analyse plus proche du système psycho-visuel par rapport à celle obtenue avec la famille de Gabor. De plus, [Chang et Kuo, 1993] décrit les ondelettes comme plus naturelles et efficaces pour des textures composées majoritairement de fréquences moyennes. Les ondelettes sont donc un outil puissant pour la caractérisation de textures et sont très utilisées dans la littérature [Chang et Kuo, 1993, Laine et Fan, 1993, Unser, 1995,

Laine et Fan, 1996, Liapis *et al.*, 1998, Arivazhagan et Ganesan, 2003, Borah *et al.*, 2007, Kim et Kang, 2007].

La transformée en ondelette remplace les ondes sinusoïdales de la transformée de Fourier par une famille générée par translations et dilatations d'une fenêtre appelée ondelette. La fonction  $\psi_\lambda$  prend comme valeur :

$$\psi_\lambda(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-x_0}{a}\right) \quad (1.31)$$

avec  $\lambda = (x_0, a)$ ,  $x_0$  étant un coefficient de translation et  $a$  un coefficient de dilatation. L'analyse d'une fonction  $f$  se fait ensuite par dilatation et par translation d'une ondelette mère  $\psi$ . Les coefficients d'ondelette  $C_\lambda$  d'un ensemble de coefficient  $\lambda$  sont obtenus de la façon suivante :

$$C_\lambda(x) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x)\bar{\psi}\left(\frac{x-x_0}{a}\right)dx \quad (1.32)$$

avec  $a$  le coefficient d'échelle,  $x_0$  le coefficient de translation,  $\psi$  l'ondelette mère et  $(\bar{\phantom{x}})$  le conjugué complexe de la quantité en argument.

[Unser et Blu, 2003] affirme l'existence de fonctions d'ondelettes qui génèrent une base dans  $L_2$ . Une fonction  $f \in L_2$  peut être représentée de façon stable à partir de coefficients de détails  $d_a(x_0)$  et de fonctions de base  $\psi_{a,x_0} = 2^{-\frac{a}{2}}$  :

$$f(x) = \sum_{a \in \mathbb{Z}} \sum_{x_0 \in \mathbb{Z}} d_a(x_0)\psi_{a,x_0} = 2^{-\frac{a}{2}} \quad (1.33)$$

Les coefficients d'ondelette  $d_a(x_0)$  sont obtenus à partir du produit scalaire suivant :

$$d_a(x_0) = \langle f, \bar{\psi}_{a,x_0} \rangle_{L_2} \quad (1.34)$$

où  $\bar{\psi}_{a,x_0}$  est la base biorthogonale de  $\psi_{a,x_0}$  tel que  $\langle \bar{\psi}_{a,x_0}, \psi_{a',x'_0} \rangle = \delta_{a-a'} \cdot \delta_{x_0-x'_0}$ .

Pour construire une ondelette mère, il faut utiliser une fonction  $\phi$  appelée fonction d'échelle [Mallat, 1989]. Une ondelette  $\psi$  est exprimée selon la fonction d'échelle  $\phi$  de la façon suivante :

$$\psi(x) = \sqrt{2} \sum g(x_0)\phi(2x - x_0) \quad (1.35)$$

sachant que

$$g(x_0) = (-1)^{1-x_0}h(1 - x_0) \quad (1.36)$$

avec  $h$  le filtre conjugué miroir de la fonction d'échelle  $\phi$ .

Pour qu'une fonction d'échelle  $\phi$  soit admissible, elle doit vérifier plusieurs contraintes. [Unser et Blu, 2003] propose la définition suivante :

**Definition 1.2.1.**  $\phi$  est une fonction d'échelle admissible de  $L_2$  si et seulement si elle satisfait les trois conditions suivantes :

- Critère de la base de Riesz : il existe deux constantes  $A > 0$  et  $B < +\infty$  telles que

$$A \leq \sum_{n \in \mathbb{Z}} \left| \hat{\phi}(w + 2\pi n) \right|^2 \leq B, \quad \forall w \quad (1.37)$$

– *Raffinabilité*

$$\phi(x) = \sqrt{2} \sum h(x_0)\phi(2x - x_0) \quad (1.38)$$

– *Partition d'unité*

$$\sum_{x_0 \in \mathbb{Z}} \phi(x - x_0) = 1 \quad (1.39)$$

#### 1.2.2.4 Décomposition en ondelette discrète

L'algorithme de Mallat propose de décomposer un signal par une succession de filtres passe haut et filtre passe bas [Mallat, 1989]. La décomposition en ondelette n'est pas faite directement à partir des formes explicites de  $\phi(x)$  et  $\psi(x)$  mais à partir des coefficients  $g(x)$  (filtre passe haut) et  $h(x)$  (filtre passe bas). La figure 1.12 présente la décomposition d'un signal 1D à partir de ces deux filtres. A chaque niveau de décomposition, le signal est sous échantillonné par un facteur de deux. En sortie du filtre passe bas  $g(x)$ , nous obtenons des coefficients d'approximations pour une résolution donnée. Ce sont ces coefficients qui sont utilisés pour une décomposition à un niveau plus grossier. La sortie du filtre passe haut permet d'obtenir les coefficients de détail qui donnent une description des hautes fréquences présentes dans un signal.

Pour traiter des images, le mécanisme de décomposition reste le même. Une première convolution est appliquée dans le sens des lignes puis une deuxième dans le sens des colonnes, avec pour chaque convolution une étape de sous échantillonnage (Figure 1.13). Pour une image 2D, quatre images résultats sont obtenues pour un niveau de décomposition :  $A_1$  correspond aux basses fréquences horizontales et verticales,  $D_1$  aux hautes fréquences verticales et basses fréquences horizontales,  $D_2$  aux basses fréquences verticales et hautes fréquences horizontales,  $D_3$  aux hautes fréquences verticales et horizontales. De cette façon, il est possible d'appliquer une décomposition en ondelette discrète sur un signal à  $n$  dimensions [Mallat, 1989].

#### 1.2.2.5 Les frames d'ondelette

La décomposition en ondelette discrète classique n'est pas invariante à la translation. A chaque niveau de décomposition, il y a une étape de sous-échantillonnage et un déplacement même léger pourrait entraîner une représentation différente. Dans le cas d'une image contenant plusieurs textures, celles-ci peuvent apparaître n'importe où et les caractéristiques obtenues à partir d'une décomposition en ondelette standard peuvent ne pas correspondre avec les caractéristiques associées à ces textures.

Les frames d'ondelette permettent de contourner ce problème en éliminant l'étape de sous-échantillonnage de la décomposition en ondelette classique. Unser est le premier à utiliser cette méthode et les résultats qu'il présente sont meilleurs qu'avec une décomposition en ondelette standard [Unser, 1995]. Il fait également la comparaison avec des méthodes d'analyses de textures classiques comme les matrices de cooccurrences, les champs aléatoires de Markov Gaussien et des transformées linéaires locales, qui s'avèrent être moins performantes. Les frames d'ondelette permettent une description invariante aux translations dans une texture. Cependant, une partie de l'information est dupliquée et les caractéristiques

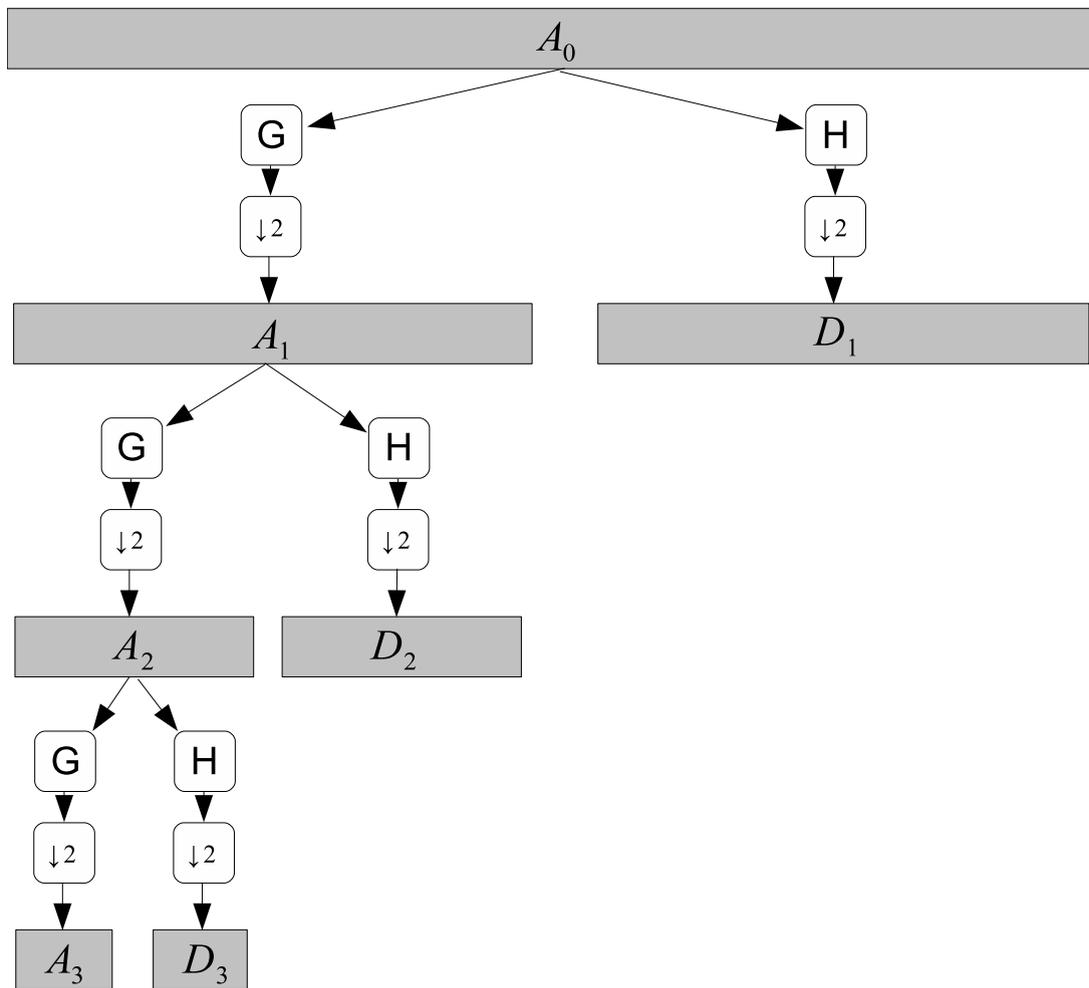


FIG. 1.12 – Décomposition en ondelette d'un signal 1D

obtenues à partir du signal ou de l'image d'origine sont redondantes. Enfin, cette méthode possède une complexité non-négligeable ce qui peut être problématique dans certains cas d'applications.

### 1.2.2.6 Les paquets d'ondelette

Les paquets d'ondelette sont une extension de la décomposition en ondelette standard et permettent une analyse plus précise des fréquences présentes dans une image. En effet, la transformée en ondelette discrète utilise uniquement les coefficients d'approximations pour une éventuelle décomposition à un niveau plus grossier. Dans le cas de la transformée en paquets d'ondelette, le passage à un niveau inférieur se fait en décomposant les coefficients d'approximations ainsi que les coefficients de détails. Cette méthode permet d'avoir à la fois une analyse sur les basses fréquences mais aussi sur les hautes fréquences. Enfin, de par le sous-échantillonnage, le nombre de coefficients d'ondelette n'augmente pas par rapport à la décomposition en ondelette discrète et il n'y a pas de redondance. La figure 1.14 montre

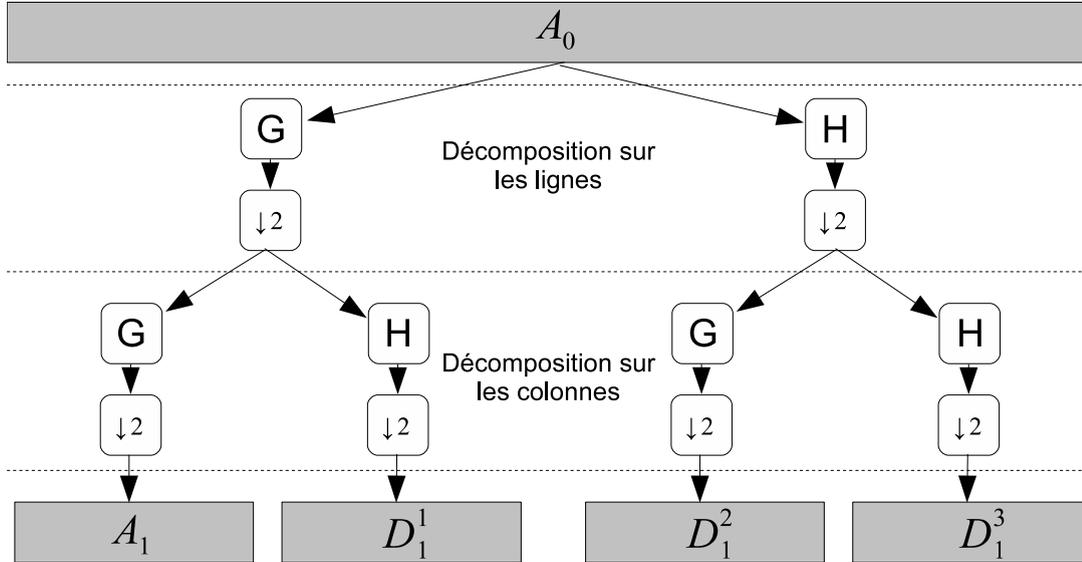


FIG. 1.13 – Décomposition en ondelette d'une image 2D

une décomposition en paquets d'ondelette sur un signal 1D.

Cette approche est utilisée dans [Laine et Fan, 1993]. Les auteurs proposent une décomposition en paquets d'ondelette et calculent l'énergie en chaque sous-bande pour former un vecteur de caractéristiques afin de classifier des textures à l'aide d'un réseau de neurones. Laine et Fan utilisent une extension de cette technique et la comparent à la méthode des frames d'ondelette afin de segmenter des textures de mosaïques [Laine et Fan, 1996]. A partir d'un algorithme de classification non-supervisé (*K-means*), les résultats obtenus s'avèrent de meilleure qualité avec leur méthode en décomposition par paquet.

La décomposition en ondelette par paquet donne plus de détails que nécessaire pour analyser une texture. Afin de corriger ce problème, il faudrait pouvoir récupérer uniquement les caractéristiques intéressantes et discriminantes. Pour cela, il est nécessaire de déterminer quelles sous-bandes doivent ou ne doivent pas être décomposées. Coifman et Wickerhauser présentent une première méthode de sélection pour la décomposition des sous-bandes [Coifman et Wickerhauser, 1992]. Ils proposent de calculer un terme d'entropie pour une sous-bande  $\alpha$ . Il est possible de la calculer de la façon suivante :

$$H_\alpha = - \sum_{i=0}^{N_\alpha} E_i \log E_i \quad (1.40)$$

avec  $E_{\alpha,i}$  l'énergie normalisée du  $i^{eme}$  coefficient d'ondelette de la sous-bande  $\alpha$  et  $N_\alpha$  le nombre de coefficients d'ondelette dans la sous-bande  $\alpha$ . Dans un premier temps, ils effectuent une décomposition complète en ondelette par paquets. Par la suite, ils cherchent à minimiser la somme des entropies des sous-bandes se trouvant au bout de la décomposition. En partant du niveau le plus bas, ils comparent l'entropie du parent de chaque paire de sous-bandes à la somme de l'entropie de ses fils. Si l'entropie du parent est inférieure,

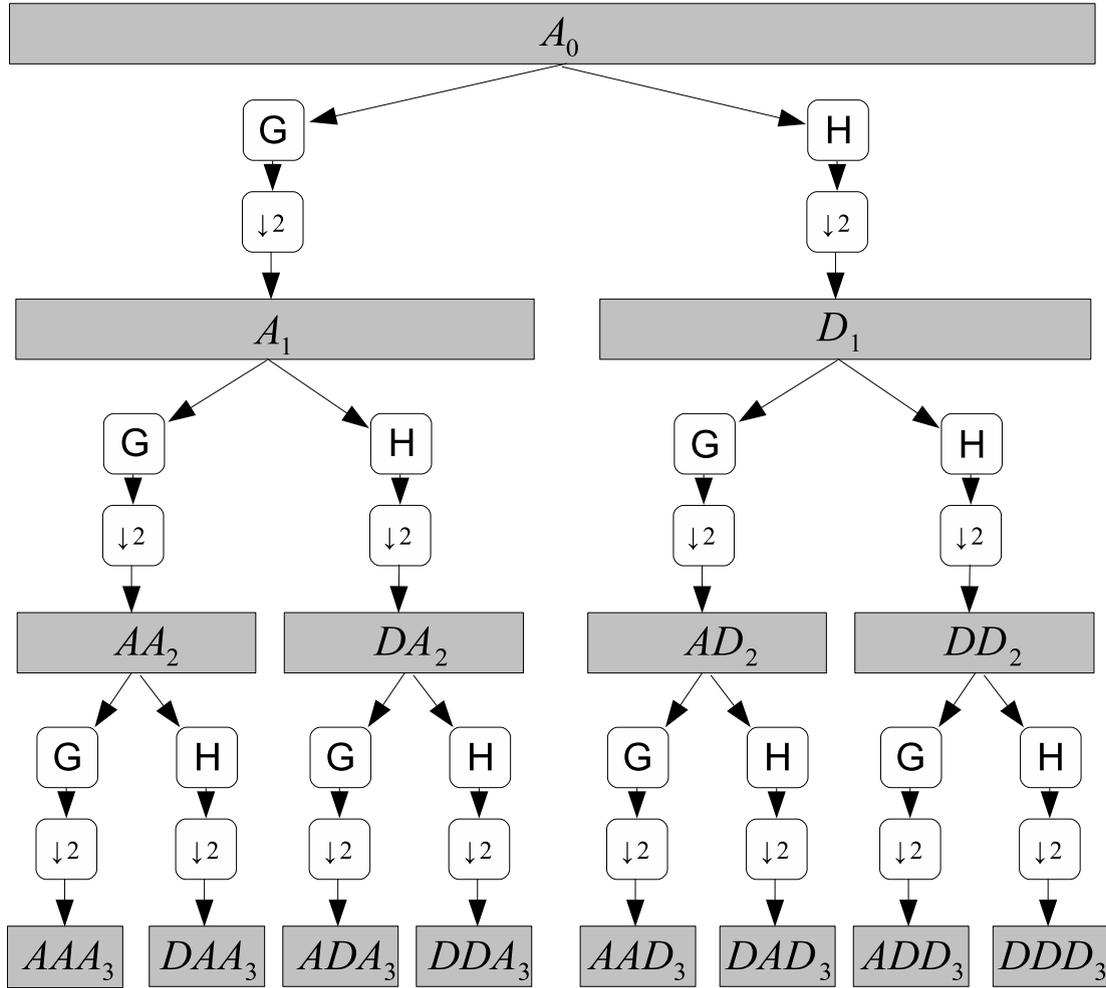


FIG. 1.14 – Décomposition en paquets d'ondelette sur un signal 1D

alors les fils sont supprimés de l'arbre de décomposition. Cette vérification est par la suite appliquée à l'ensemble de l'arbre.

Dans [Chang et Kuo, 1993], les auteurs appliquent l'idée de la décomposition par paquet d'ondelette à l'analyse des textures. La norme L1 moyennée est utilisée pour identifier si une sous-bande  $\alpha$  doit être décomposée ou pas. Voici son expression :

$$E_\alpha = \frac{1}{N_\alpha} \sum_{i=0}^{N_\alpha} |w_{\alpha,i}| \quad (1.41)$$

avec  $w_{\alpha,i}$  le  $i^{eme}$  coefficient d'ondelette de la sous-bande  $\alpha$  et  $N_\alpha$  le nombre de coefficients d'ondelette dans la sous-bande  $\alpha$ . La norme L1 moyennée est calculée pour chaque sous-bande en partant du premier niveau de décomposition. Si pour une sous-bande, nous obtenons une valeur faible, alors celle-ci n'est pas décomposée. Les sous-bandes ayant les normes les plus élevées sont conservées comme caractéristiques de textures.

### 1.2.2.7 Le banc de filtres de Schmid

Dans [Schmid, 2001], l'auteur propose d'utiliser un ensemble d'opérateurs dit générique afin d'analyser les structures et les textures d'une image. L'ensemble de Schmid se compose de 13 filtres invariants à la rotation :

$$F(x, y, \sigma, \tau) = F_0(\sigma, \tau) + \cos\left(\frac{\sqrt{x^2 + y^2}\pi\tau}{\sigma}\right)e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (1.42)$$

avec  $(\sigma, \tau)$  prenant les valeurs suivantes : (2, 1), (4, 1), (4, 2), (6, 1), (6, 2), (6, 3), (8, 1), (8, 2), (8, 3), (10, 1), (10, 2), (10, 3) et (10, 4).  $\tau$  représente le nombre de cycles de la fonction harmonique à l'intérieur de l'enveloppe Gaussienne du filtre. La figure 1.15 présente les différents filtres.

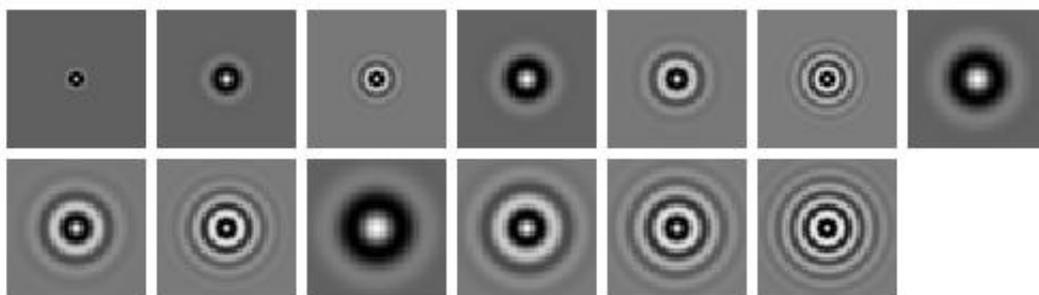


FIG. 1.15 – Les 13 filtres du banc de filtre de Schmid

### 1.2.2.8 Le banc de filtre de Leung-Malik (LM)

L'ensemble LM (figure 1.16) est un banc de filtres à la fois multi-échelle et multi-orientation. Il se compose de 48 filtres construits à partir des dérivées premières et secondes de Gaussiennes selon 6 orientations et 3 échelles, de 8 Laplaciens de Gaussiennes et de 4 Gaussiennes. Dans [Leung et Malik, 2001], les auteurs utilisent ce banc de filtres dans un problème de classification de textures 3D. Rappelons que les textures 3D sont des textures avec des variations de luminosité et de surface. Leur méthode propose de construire un vocabulaire de textons, utilisant l'ensemble LM, afin d'obtenir une description générique des surfaces des textures. Ce vocabulaire est utilisé avec un algorithme de chaînes de markov afin de reconnaître les textures 3D.

### 1.2.2.9 Les bancs de filtres de réponse maximum (MR)

Le banc de filtres MR8 [Caenen et Gool, 2004] est similaire au banc de filtres LM et se compose de 38 filtres : 2 filtres anisotropiques pour la détection de frontières et de bandes selon 6 orientations et 3 échelles  $(\sigma_x, \sigma_y) = \{(1, 3), (2, 6), (4, 12)\}$  pixels, 2 filtres isotropiques avec une Gaussienne et le Laplacien d'une Gaussienne ( $\sigma = 10$  pixels) (Figure 1.17). Cependant, seulement 8 réponses de filtres sont conservées en prenant à chaque échelle, la réponse maximum des filtres anisotropiques selon toutes les orientations. Le nombre de réponses est ainsi réduit de 38 (6 orientations à 3 échelles pour 2 filtres anisotropiques, plus

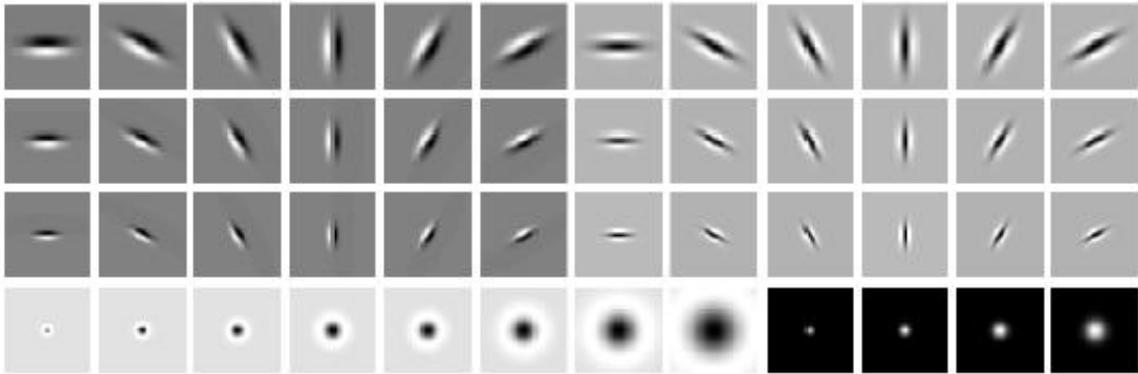


FIG. 1.16 – Le banc de filtres de Leung-Malik est composé de 48 filtres avec des filtres frontières, tâches et bandes selon plusieurs résolutions et orientations

2 filtres isotropiques) à 8 (3 échelles pour 2 filtres, plus 2 isotropiques). Des sous-ensembles du banc de filtres MR8 sont utilisés avec, par exemple, le banc de filtres MR4 qui est un sous-ensemble du MR8. Ici, une seule échelle est considérée avec  $\sigma_x = 4$  et  $\sigma_y = 12$ .

De façon générale, les filtres invariants à la rotation ne répondent pas de façon précise dans les images contenant des orientations. Les ensembles MR utilisent à la fois des filtres isotropiques et anisotropiques qui permettent d'obtenir des caractéristiques robustes à la fois pour les données orientées et non-orientées. Enfin le fait de conserver uniquement la réponse maximum limite le nombre de caractéristiques en regroupant les informations de textures.

### 1.2.2.10 Bilan

Les méthodes de description de textures fréquentielles fournissent des informations très intéressantes, cependant celles-ci sont souvent difficilement assimilables par les utilisateurs (il est difficile d'interpréter un spectre de Fourier ou un banc de filtre). Pour toutes ces méthodes basées sur des filtres, les caractéristiques de textures sont calculées directement sur les réponses générées pour chaque filtre. Certains chercheurs [Arivazhagan et Ganesan, 2003, Fauzi et Lewis, 2003] utilisent des statistiques simples comme la variance ou la moyenne. D'autres [Unser, 1995] utilisent des caractéristiques statistiques plus complexes comme le contraste, l'énergie, l'entropie, etc. Des statistiques d'ordre deux peuvent également être utilisés : dans [Arivazhagan et Ganesan, 2003] les auteurs calculent des caractéristiques d'Haralick en construisant une matrice de cooccurrence sur des coefficients d'ondelette. Il s'agit encore une fois de caractéristiques difficiles à manipuler pour un humain. Cependant, le point fort de ce type de méthodes est leur faculté à apporter une vision/information multi-niveaux sur le contenu de l'image. Ce point constitue une solution potentielle (bien que partielle) vis à vis des inconvénients mentionnés précédemment pour les méthodes d'analyse de motifs (régionalité et échelle d'observation). Nous retenons donc cet intérêt pour la suite de nos travaux.

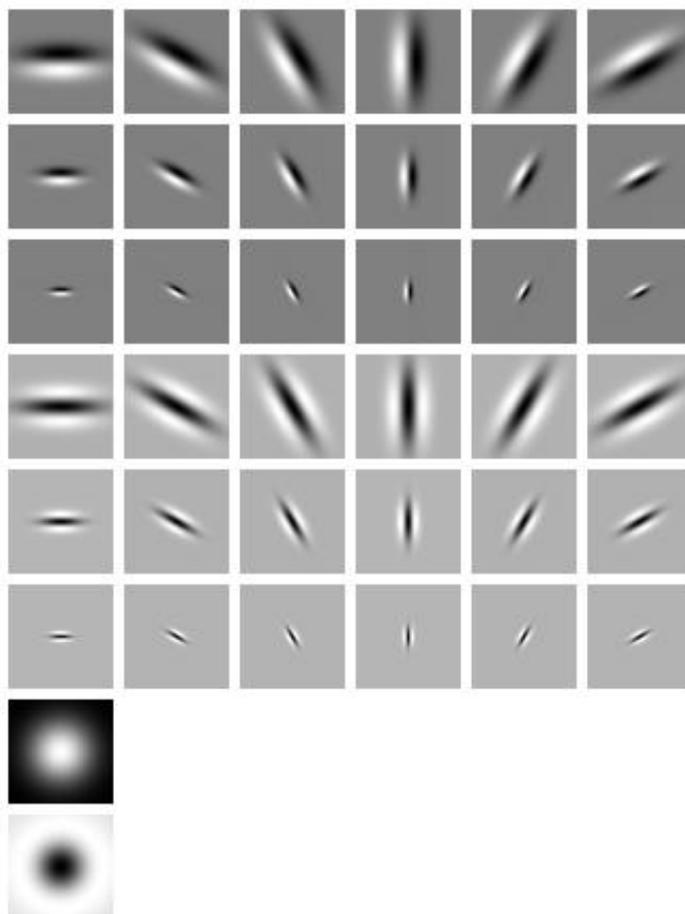


FIG. 1.17 – Banc de filtres MR8 composé de 38 filtres à la fois isotropiques et anisotropiques

### 1.2.3 Les méthodes basées sur un modèle

Parmi les modèles les plus utilisés pour l'analyse de textures, nous retrouvons les champs de Markov (MRF), les fractales, les modèles de différenciation fractionnaire et les modèles auto-régressifs. L'estimation des paramètres de ces modèles permet de caractériser les textures d'une image.

#### 1.2.3.1 Les champs de Markov

Les champs de Markov sont beaucoup utilisés dans le domaine de l'analyse et de la synthèse de textures. [Li, 1995, Li, 2001] nous donne un panorama de leur utilisation en vision par ordinateur.

##### *Notion de sites*

Les Champs de Markov décrivent une image  $I$  par l'intermédiaire d'un ensemble fini  $S$  de sites  $s_i$  qui correspondent aux pixels. A chaque site est associé un système de voisinage

$V$ .

Mathématiquement,  $I = \{s_1, s_2, s_3, \dots, s_N\}$  avec  $N$  la taille de l'image, et  $V = \{V_s, s \in I\}$ .

Ce voisinage respecte certaines propriétés qui sont les suivantes :

- $s \notin V_s, \forall s \in I$
- $\forall s, t \in I, s \in V_t \Leftrightarrow t \in V_s$

2	1	2
1	s	1
2	1	2

FIG. 1.18 – Représentation d'un système de voisinage d'ordre 2

A partir de ce voisinage, un système de cliques est construit. Une clique peut être soit un singleton de  $S$ , soit un ensemble de sites voisins (Figure 1.19). Selon le nombre de connexité utilisé, le système de cliques sera différent et fera intervenir plus ou moins de sites. Nous définissons ainsi le voisinage influant sur le niveau de gris du pixel considéré. Les interactions locales entre niveaux de gris de sites voisins peuvent alors s'exprimer comme un potentiel de clique. Si l'on définit  $c$  une clique associée à un potentiel  $U_c$  alors l'énergie globale de l'image s'écrit :

$$U = \sum_{c \in C} U_c \tag{1.43}$$

avec  $C$  l'ensemble des cliques relatives à  $V$ .

L'énergie locale en un site peut alors s'écrire :

$$U_s = \sum_{c \in C/s \in c} U_c \tag{1.44}$$

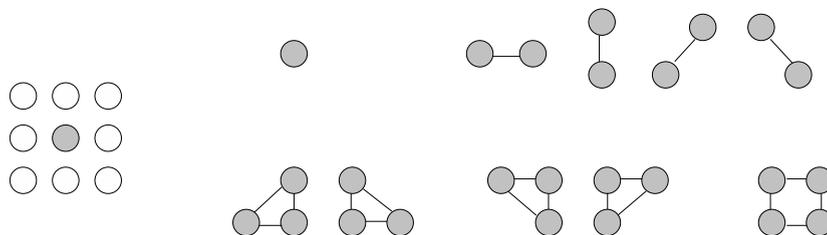


FIG. 1.19 – Ensemble de cliques pour un système de voisinage d'ordre 2 (8-connexe)

*Modélisation probabiliste*

L'utilisation des champs de Markov en imagerie nécessite une modélisation probabiliste de l'image. Ainsi, une image est considérée comme une réalisation d'un champ aléatoire. Pour chaque site  $s$  de l'image, il existe différentes réalisations  $x_s$  (niveau de gris) possibles pouvant être perçues comme une famille de variables aléatoires  $X = \{X_1, \dots, X_N\}$  indexées sur  $I$ . Chaque variable aléatoire  $X_s$  est à valeur dans  $\Omega$ , l'espace des états.

[Geman et Geman, 1984] donne la définition du champ de Markov 2D suivant :

**Definition 1.2.2.** *X est un champ de Markov 2D sur  $(I, V)$  si :*

- $P(X = x) > 0, \forall x \in \Omega^N$
- $P(X_s = x_s / X_r = x_r, r \in S - \{s\}) = P(X_s = x_s / X_r = x_r, r \in V_s)$
- $P(X_s = x_s / V_s)$  est dépendante de la configuration des voisins et est invariante à la translation

Le premier point représente la propriété de positivité qui garantit l'unicité de la distribution jointe  $P(X = x)$  du champ de Markov. La deuxième propriété traduit le fait que l'influence des autres pixels de  $V$  sur un pixel  $i$  est réduite à l'influence du voisinage de  $i$ .

Cette formulation ne permet pas d'obtenir de façon simple une expression de  $p(X = x)$ . Un champ de Markov est donc généralement caractérisé par une distribution jointe, ce qui est rendu possible par le théorème de Hammersley-Clifford [Besag, 1974, Geman et Geman, 1984]. Celui-ci lie la propriété de localité des champs de Markov avec la propriété de globalité des champs de Gibbs.

**Théorème 1.2.1.** *X est un champ de Markov  $\iff p(X = x)$  est une distribution de Gibbs.*

La définition d'un champ de Gibbs est la suivante :

**Definition 1.2.3.** *Si  $(X, \Omega^N, P)$  est un champ de Markov par rapport au système de voisinage  $V$  alors sa distribution  $P(X = x)$  est une distribution de Gibbs de portée bornée :*

$$P(X = x) = \frac{1}{Z} \exp(-U(x))$$

avec  $Z = \sum_{X \in \Omega^N} \exp(-U(X))$  une fonction de normalisation appelée fonction de partition de Gibbs.

*Estimation des paramètres des champs de Markov*

Concernant l'estimation des paramètres d'un champ de Markov pour une image, on peut recenser deux grandes familles de problèmes :

- les données dites complètes, avec lesquelles on cherche à retrouver les paramètres d'une réalisation connue d'un champ de Markov. On se retrouve dans cette problématique par exemple lorsque l'on cherche à caractériser une texture via les paramètres d'un champ de Markov.
- les données dites incomplètes, pour lesquelles le champ de Markov recherché est caché, et ses paramètres sont à estimer. On peut citer comme exemple le cas de la segmentation ou bien encore de la restauration d'images.

*Exemple de caractérisation de textures avec le modèle markovien binomial*

Ce modèle est un cas particulier du modèle markovien général et permet une bonne description de variables discrètes telles que le niveau de gris d'un site d'une image. Dans [Cross et Jain, 1983], les auteurs l'utilisent pour synthétiser et caractériser des textures.

## 1.2. MÉTHODES D'EXTRACTION DE CARACTÉRISTIQUES DE TEXTURES

D'un point de vue formel, la probabilité conditionnelle d'avoir un niveau de gris  $x_s$  dans le site  $S$  est :

$$p(X_s = x_s | x_t, \forall t \in V_s) = C_{F-1}^{x_s} p^{x_s} (1-p)^{F-1-x_s}$$

avec  $V_s$  le voisinage du site  $S$ . On est ici dans le cas d'une loi binomiale avec une probabilité  $p$  et un nombre de tirage de  $F-1$ ,  $F$  étant le nombre total de niveaux de gris. La probabilité  $p$  qui est un paramètre de la loi binomiale s'exprime de la façon suivante :

$$\frac{p}{1-p} = e^{U_s} \Rightarrow p = \frac{e^{U_s}}{1 + e^{U_s}}$$

avec  $U_s$  dépendant de l'ordre considéré par le modèle, et défini comme une fonction linéaire calculée à partir des niveaux de gris du voisinage au site  $S$ .

La figure 1.18 présente les pixels pris en compte dans un système de voisinage d'ordre 2 (pixels 1 et 2). Dans un modèle d'ordre 1, seuls les pixels labellisés "1" auraient été utilisés. En prenant comme exemple le modèle d'ordre 2 de la figure 1.19,  $U_s$  peut être calculé de la façon suivante :

$$U_s = a + b_1(x_{i,j-1} + x_{i,j+1}) + b_2(x_{i-1,j} + x_{i+1,j}) + b_3(x_{i-1,j-1} + x_{i+1,j+1}) + b_4(x_{i+1,j-1} + x_{i-1,j+1})$$

Les coefficients de pondération  $a$ ,  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$  correspondent alors aux descripteurs de la texture.  $a$  donne une représentation du niveau de gris moyen,  $b_1$  décrit l'homogénéité horizontale,  $b_2$  verticale, et  $b_3$  et  $b_4$  l'homogénéité diagonale.

Une notation vectorielle pour  $U_s$  donnerait :

$$U_s = \theta^T \cdot \Phi \quad (1.45)$$

avec  $\theta^T = [a, b_1, b_2, b_3, b_4]$  et  $\Phi^T = [1, x_{i,j-1} + x_{i,j+1}, x_{i-1,j} + x_{i+1,j}, x_{i-1,j-1} + x_{i+1,j+1}, x_{i+1,j-1} + x_{i-1,j+1}]$ .

Une des façons pour estimer  $\hat{\phi}$  est d'utiliser la méthode du maximum de vraisemblance. Cependant, celle-ci est en pratique extrêmement lente. D'autres méthodes emploient des algorithmes itératifs avec notamment l'algorithme du gradient stochastique [Younes, 1988].

### *Cas de la segmentation : Données incomplètes*

Dans le cas de la segmentation, le champ de Markov cherché est caché. Nous disposons d'une image  $y$  qui est une réalisation d'un champ aléatoire  $Y$ . Nous cherchons alors une réalisation  $x$  de l'image segmentée qui peut être modélisée comme un champ de Markov  $X$  appelé champ des étiquettes. Ces deux champs ont des espaces de configurations différents. En effet, dans le champ  $X$ , les étiquettes considérées correspondent aux différentes classes cherchées. Si l'on considère que chaque classe  $i$  a une distribution gaussienne de moyenne  $\mu_i$  et d'écart-type  $\sigma_i$ , alors la probabilité d'avoir un niveau de gris pour une classe donnée s'écrit :

$$P(y_s/x_s = i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_s - \mu_i)^2}{2\sigma_i^2}\right) \quad (1.46)$$

La probabilité *a priori* est utilisée pour introduire des contraintes sur la solution. Cette

probabilité dans le cas d'un modèle gaussien s'écrit :

$$P(X = x) = \frac{1}{Z} \exp(-U(x)),$$

$$\text{avec } U(x) = \beta \sum_{(s,t) \in C_2} (x_s - x_t)^2 + \alpha \sum_{s \in S} (x_s - \mu_s)^2 \quad (1.47)$$

Dans  $U$ , le premier terme correspond à un terme de régularisation alors que le deuxième se définit comme un terme d'attache aux données.  $\alpha$  et  $\beta$  permettent de pondérer ces deux expressions.

L'énergie correspondant à la distribution de Gibbs du champ *a posteriori* peut alors s'écrire :

$$U(x/y) = \sum_s \frac{(y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2} + \log(\sqrt{2\pi}\sigma_{x_s}) + \beta \sum_{(s,t) \in C_2} (x_s - x_t)^2 + \alpha \sum_{s \in S} (x_s - \mu_s)^2 \quad (1.48)$$

Dans cet exemple, nous utilisons le modèle gaussien, mais d'autres modèles de champs de Markov peuvent être utilisés comme le modèle d'Ising [Ising, 1925] ou le modèle de Potts [Wu, 1982].

Pour trouver les états d'énergie minimum, il existe plusieurs estimateurs [Besag, 1986] :

- l'estimateur MAP (maximum *a posteriori*) : avec cet estimateur, on cherche à pénaliser toute différence entre deux configurations. Les algorithmes associés à cet estimateur sont le recuit simulé ainsi que l'algorithme des modes conditionnels itérés (ICM).
- l'estimateur MPM (maximum *a posteriori* marginal) : la fonction de coût associé à cet estimateur pénalise une configuration proportionnellement au nombre de différences entre deux configurations.
- l'estimateur TMP (*Thresholded Posteriori Mean*) : la fonction de coût pénalise la somme des différences entre deux configurations. Si l'on compare avec l'estimateur MPM, elle tient compte non seulement du nombre de différences mais également de leurs valeurs.

Parmi tous les modèles Markoviens, le plus utilisé est le champ de Markov Gaussien (GMRF) [Chellappa et Jain, 1993, Cohen et Cooper, 1987, Chellappa et Chatterjee, 1985]. Dans [Descombes *et al.*, 1999], les auteurs proposent deux modèles markoviens gaussiens, l'un étant isotrope et l'autre anisotrope. Pour chacun de ces modèles, une méthode d'estimation est utilisée, la première étant fondée sur la construction de matrices de probabilités conditionnelles, la deuxième sur la théorie de la renormalisation. [Krishnamachari et Chellappa, 1997] utilisent la méthode GMRF pour modéliser l'image à plusieurs résolutions en faisant l'hypothèse que les variables aléatoires sont indépendantes quelle que soit la résolution. Les propriétés de Markov étant perdues pour des résolutions trop grossières, elles ne peuvent pas être approximées par des champs aléatoires de Markov. Ils proposent alors une approximation avec les Champs de Markov. Dans le même principe, [Comer et Delp, 1999] utilisent un modèle gaussien auto-régressif. Cependant, leur méthode prend en compte la corrélation entre deux niveaux adjacents de résolution.

### 1.2.3.2 Les fractales

Les fractales ont été utilisées avec beaucoup de succès pour synthétiser les textures naturelles. Sa grande capacité à caractériser la rugosité fait de la dimension fractale un candidat naturel pour la mesure de caractéristiques de textures. Mandelbrot [Mandelbrot, 1983, Mandelbrot, 1977] est le premier à relater l'existence de la géométrie fractale dans le monde.

Dans la géométrie classique, un objet à une dimension est symbolisé par une ligne, un objet à deux dimensions un plan et un objet à trois dimensions par un volume. De cette façon la dimension des objets est un nombre entier allant de 0 jusqu'à  $n$ . La géométrie fractale peut être définie comme une notion complémentaire à la géométrie euclidienne. Celle-ci définit la manière dont un objet occupe l'espace euclidien à travers un nombre, généralement non entier, qui caractérise un espace où la dimension de l'objet prend une valeur finie non nulle. La courbe de Koch constitue un bon exemple pour mieux comprendre la signification de ce nombre. Dans [Mandelbrot, 1977], cette courbe est construite de la façon suivante : en partant d'un segment de longueur unité, le tiers central est retiré et remplacé par un triangle équilatéral sans base (Figure 1.20). Par rapport au segment initial, la longueur de la courbe augmente en passant de  $3/3$  à  $4/3$ . En continuant cette opération sur chacun des segments de longueur  $1/3$ , la longueur de la courbe continue d'augmenter avec  $(4/3)^2$ . Si cette opération est répétée un nombre infini de fois, alors nous pouvons construire la courbe de Koch de longueur infinie.

La mesure de la longueur de cette courbe dépend de l'étalon choisi. Pour un étalon de 1 la longueur de la courbe sera de 1, pour un étalon de  $1/3$  la longueur sera de  $4/3$  etc. Soit  $L(\delta)$  la longueur mesurée avec un segment de mesure  $\delta$ . Avec un étalon de longueur  $\delta/3$ , la longueur  $L(\delta/3)$  sera égale à la longueur mesurée avec un étalon de longueur  $\delta$  multipliée par  $4/3$ . La longueur de la courbe de Koch à la  $n^{ième}$  étape peut donc être exprimée de la façon suivante :

$$L(\delta) = \left(\frac{4}{3}\right)^n \quad (1.49)$$

avec  $\delta = \left(\frac{1}{3}\right)^n$  la longueur d'un étalon d'où  $n = -\frac{\text{Log}\delta}{\text{Log}3}$  La longueur de la courbe peut alors être exprimée comme suit :

$$L(\delta) = \left(\frac{4}{3}\right)^{-\frac{\text{Log}\delta}{\text{Log}3}} \quad (1.50)$$

D'après la loi de Richardson, nous avons :

$$L(\delta) = \left(\frac{4}{3}\right)^{-\frac{\text{Log}\delta}{\text{Log}3}} = \delta^{1-D} \quad (1.51)$$

En prenant le logarithme de la dernière égalité, nous trouvons :

$$D = \frac{\text{Log}4}{\text{Log}3} = 1.2628 \quad (1.52)$$

Ce nombre D, appelé dimension fractale, peut être considéré comme la dimension de la courbe de Koch. Pour cette courbe, le quart de l'objet est identique à l'objet initial s'il est dilaté 3 fois. Or la valeur de D est le logarithme de 4 sur le logarithme de 3.

Un concept important concernant les fractales est l'auto-similarité. Un objet invariant par la transformation d'un vecteur  $r$  est dit auto-affine de vecteur  $r$ . De plus, si les composantes du vecteur  $r$  sont égales, alors nous pouvons parler d'auto-similarité. Par exemple la courbe de Koch est auto-similaire de vecteur  $r = (3, 3)$ . Si nous prenons le quart gauche et que nous le dilatons, nous retrouvons cette même courbe. L'auto-similarité d'une structure, c'est donc le fait qu'elle puisse être réduite à des éléments plus petits qui ne sont que des répliques de la structure initiale.

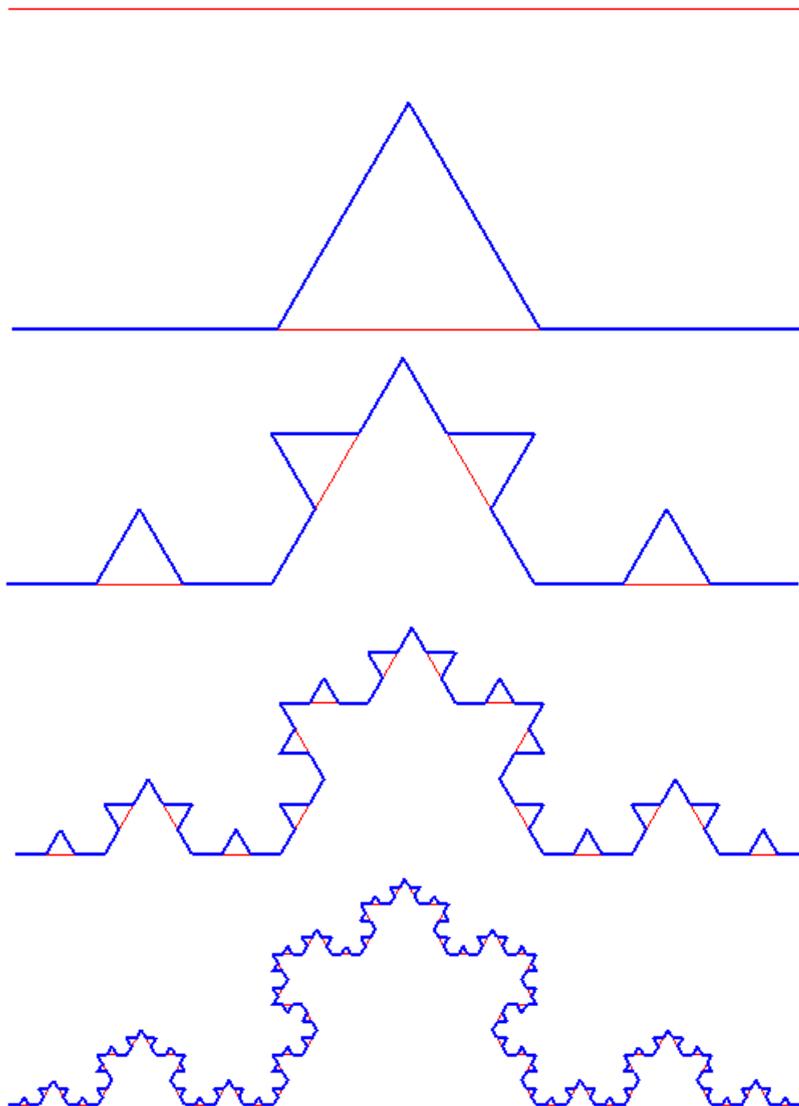


FIG. 1.20 – Construction de la courbe de Koch

Dans [Soille et Rivest, 1996], les auteurs proposent un état de l'art sur les méthodes d'estimation de la dimension fractale. Celles-ci peuvent être divisées en deux familles comprenant les méthodes d'estimations géométriques et les méthodes stochas-

tiques. Parmi les méthodes d'estimations géométriques, nous pouvons citer les méthodes d'analyse morphologique [Peleg *et al.*, 1984], la méthode des triangles plans [Clarke, 1986], la méthode de l'élément structurant plat et enfin la méthode dite de *Box counting* [Sarkar et Chaudhuri, 1994]. Les méthodes stochastiques permettent d'estimer la dimension fractale en analysant le spectre de puissance obtenu à partir de la transformée de Fourier [Soille et Rivest, 1996] mais également en utilisant les variogrammes [Dubuc *et al.*, 1989].

La méthode du *Box counting* est l'une des plus utilisées grâce à sa simplicité de mise en œuvre ainsi que pour son faible coût en temps d'exécution. L'image est binarisée puis découpée en un quadrillage de carrés élémentaires de taille  $\epsilon$ . Par la suite, le nombre de carrés  $N_\epsilon$  nécessaires à l'approximation de la texture est compté. Cela revient à faire la somme des boîtes contenant au moins un pixel noir. La *box dimension* d'une zone considérée  $X$  est alors donnée par :

$$D(X) = \lim_{\epsilon \rightarrow 0} \frac{\log[N_\epsilon(X)]}{\log \epsilon} \quad (1.53)$$

Afin de déterminer cette limite, plusieurs valeurs de  $\epsilon$  sont considérées. La pente de régression linéaire de la courbe ainsi obtenue correspond à la dimension fractale.

En complément de cette dimension, certains auteurs proposent d'ajouter une mesure caractérisant l'homogénéité d'une distribution fractale. Cette mesure est nommée lacunarité. Avec la méthode des *box counting*, il est possible de l'estimer de la façon suivante :

$$\Lambda = \frac{E[M^2(\epsilon)] - E^2[M(\epsilon)]}{E^2[M(\epsilon)]} \quad (1.54)$$

avec  $E$  l'espérance mathématique,  $M$  la masse fractale définie par  $M(\epsilon) = \sum_{m=1}^{\epsilon^2} m PR_{G(f)}(m, \epsilon)$ .  $PR_{G(f)}(m, \epsilon)$  la probabilité de trouver un point de l'objet fractal  $G(f)$  dans un carré de taille  $\epsilon$  (probabilité de trouver un pixel noir pour le cas d'une texture).

Une des adaptations les plus connues de la méthode *box counting* est la méthode *Differential Box Counting* proposée dans [Sarkar et Chaudhuri, 1992, Sarkar et Chaudhuri, 1994]. Cette méthode propose des solutions à un certain nombre de limitations de *box counting* et permet notamment de travailler directement sur les images en niveaux de gris.

[Pentland, 1984] est l'un des premiers à utiliser la dimension fractale pour la segmentation. Il estime une dimension fractale omnidirectionnelle sur des blocs de pixels de taille  $8^2$  en utilisant la méthode des spectres de puissance. Il arrive à une segmentation grossière à l'intérieur des régions texturées et non-texturées. De la même façon, Medioni et Yasumoto utilisent une dimension fractale omnidirectionnelle pour segmenter une image contenant plusieurs textures [Medioni et Yasumoto, 1984]. Cependant les résultats obtenus ne sont pas satisfaisants. Pour eux la dimension fractale n'est pas suffisante pour caractériser tous les aspects d'une texture et permet de séparer uniquement les textures ayant une rugosité suffisamment éloignée.

Pour améliorer les résultats, de nouveaux paramètres comme la lacunarité doivent être ajoutés. Les auteurs dans [Keller *et al.*, 1989] utilisent la lacunarité ainsi qu'une estimation omnidirectionnelle de la dimension fractale. Leurs résultats montrent que ce paramètre

permet une amélioration des résultats de segmentation des textures.

Tout comme dans la plupart des méthodes d'analyses de textures, il est possible d'utiliser la notion de direction afin d'estimer la dimension fractale. [Mosquera *et al.*, 1992] estiment la dimension fractale selon quatre directions avec les deux diagonales, la direction verticale et horizontale. Dans [Jeannane, 1995], l'auteur utilise un modèle non-stationnaire fractal dans lequel il analyse des lignes de niveau de gris d'une image dans une direction donnée. Dans [Lopes et Betrouni, 2009], les auteurs présentent un état de l'art sur les méthodes d'analyse fractale et multifractale ainsi qu'une description de leur application dans le domaine médical. Les multifractales sont décrites comme une extension des fractales. Un objet multifractal est plus complexe dans le sens où il est toujours invariant par translation. Dans [Lopes et Betrouni, 2009], les méthodes permettant l'approximation du spectre multifractal sont réparties en deux classes : les méthodes dites *box counting* et les méthodes basées ondelettes. Les méthodes de la première classe sont basées sur le même principe que les méthodes *box counting* utilisées pour l'estimation de la dimension fractale. Une image est découpée avec des boîtes de tailles variables et une mesure normalisée est calculée à l'intérieur de chacune d'elles. Les méthodes basées ondelettes utilisent la transformée en ondelette d'une image comme une boîte "oscillante" afin de représenter ses différents composants. Certaines méthodes utilisent la transformée en ondelette discrète, d'autres utilisent la transformée en ondelette continue.

### 1.2.3.3 Bilan

Le point commun à toutes ces méthodes basées sur un modèle est qu'il est possible de les utiliser pour générer des textures ainsi que pour les analyser. Leur utilisation en caractérisation et segmentation peut être considérée comme plus difficile de par la nécessité de définir de nombreux paramètres complexes à manipuler. De plus, les temps de calcul de ce genre de méthodes sont parfois très importants. Ces méthodes semblent donc plus difficilement utilisables dans le cadre de la mise en place d'un système d'analyse de textures volumétriques se voulant générique et facilement manipulable par les utilisateurs.

### 1.2.4 Les méthodes basées sur une analyse spatiale et statistique des intensités de niveau de gris

Ces méthodes sont basées sur des évaluations quantitatives de la distribution des niveaux de gris définissant une texture. Ces approches ne cherchent pas à isoler un motif de base mais permettent d'évaluer les propriétés statistiques d'une région ou d'un certain voisinage autour d'un pixel.

#### 1.2.4.1 Statistiques du premier ordre

Les statistiques du premier ordre mesurent la probabilité d'observer un niveau de gris à un emplacement aléatoire d'une image. Elles sont calculées à partir de l'histogramme des intensités. Parmi les attributs du premier ordre nous retrouvons :

- les moments d'ordre  $k$  :  $\mu_k = \sum_n n^k p(n)$
- les moments centrés d'ordre  $k$  :  $\tilde{\mu}_k = \sum_n (n - \mu_1)^k p(n)$  avec :

- la moyenne des niveaux de gris :  $\mu_1$
- la variance :  $\sigma^2$
- le biais :  $\frac{\tilde{\mu}_3}{\sigma^3}$
- l'aplatissement :  $\frac{\tilde{\mu}_4}{\sigma^4} - 3$
- l'énergie :  $\sum_n |p(n)|^2$
- l'entropie :  $-\sum_n p(n) \log p(n)$
- le contraste :  $\frac{\max(n) - \min(n)}{\max(n) + \min(n)}$
- la dynamique :  $\max(n) - \min(n)$
- le coefficient de variation :  $\frac{\mu}{\sigma}$
- l'exposant de Holder [Pentland, 1984]

$p(n)$  correspond à la probabilité empirique du niveau de gris  $n$  dans l'histogramme des intensités. Pour améliorer la qualité des statistiques, il est conseillé de requantifier l'histogramme des intensités. En effet, le fait de considérer 256 niveaux de gris réduit le nombre d'occurrences moyen par pixel. Si nous effectuons des mesures dans un voisinage restreint, les statistiques mesurées risquent de devenir peu significatives.

#### 1.2.4.2 Caractéristiques d'auto-corrélation

La fonction d'auto-corrélation permet d'évaluer si une texture est plutôt fine ou grossière. La figure 1.22 présente plusieurs exemples de fonctions d'auto-corrélation calculées sur une texture large, une texture fine et une texture périodique. Pour des primitives larges, la fonction d'auto-corrélation décroît lentement alors que pour des primitives fines, elle décroît rapidement. Enfin, pour des textures périodiques, la fonction d'auto-corrélation croît et décroît périodiquement. La fonction d'auto-corrélation permet également d'étudier la présence d'orientation dans une texture [Bres, 1994].

Formellement, la fonction d'auto-corrélation 2D se définit de la façon suivante :

$$x(p, q) = \frac{M \cdot N}{(M - p)(N - q)} \cdot \frac{\sum_{i=1}^{M-p} \sum_{j=1}^{N-q} f(i, j) \cdot f(i + p, j + q)}{\sum_{i=1}^M \sum_{j=1}^N f^2(i, j)} \quad (1.55)$$

avec  $M$ ,  $N$  les dimensions de l'image,  $(p, q)$  un vecteur permettant de définir un déplacement par rapport à la position  $(i, j)$ , et  $f(i, j)$  l'intensité à la position  $(i, j)$ .

Cette fonction dispose de nombreuses propriétés :

- elle a un maximum à l'origine égal à 1
- elle est bornée à l'intervalle  $[-1, +1]$
- elle est paire
- si elle possède un autre maximum égal à 1 en un autre point  $P$ , alors elle possède une infinité de maxima aux positions  $nP$ , et elle est périodique de période  $P$ , ainsi que la fonction  $f$  elle-même
- il existe des relations simples entre les dérivées de la fonction d'auto-corrélation et les fonctions d'auto-corrélation des dérivées de  $f$

La fonction d'auto-corrélation est également liée au spectre de puissance de la transformée de Fourier. En effet le théorème de Wiener-Kinchine (<http://mathworld.wolfram.com/PlancherelsTheorem.html>) établit que la fonction d'auto-corrélation est la transformée de Fourier du spectre de densité de puissance (Figure 1.21). Sur cette base, la rose des

directions utilise la fonction d'auto-corrélation pour déterminer l'orientation et l'anisotropie d'objets ou de textures [Bres, 1994].

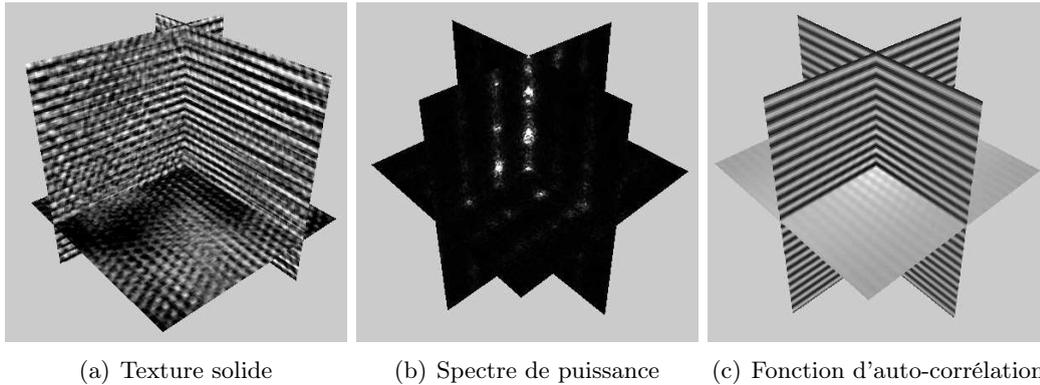


FIG. 1.21 – Fonction d'auto-corrélation calculée à partir du spectre de puissance d'une texture volumétrique

Une extension des fonctions d'auto-corrélations sont les fonctions d'auto-corrélations d'ordre  $N$ . Elles sont définies de la façon suivante :

$$x(a_1, a_2, \dots, a_N) = \int f(r)f(r + a_1)\dots f(r + a_N)dr \quad (1.56)$$

avec  $f(r)$  l'intensité du pixel  $r$  et  $a_1, a_2, \dots, a_N$   $N$  déplacements. [Otsu et Kurita, 1988] proposent une adaptation locale de cette approche appelée *High order local autocorrelation features* (HLAC). Cependant, cette méthode utilise au second ordre et reste très locale puisqu'elle utilise à un masque de dimension  $3^2$ . [Kreutz *et al.*, 1996] proposent de calculer des caractéristiques sur une région plus large ( $5 \times 5$ ). Le problème est que cette méthode génère un nombre de caractéristiques importantes à cause du grand nombre de masques possibles. En effet, les caractéristiques sont calculées en scannant l'image avec les différents masques et en calculant les sommes des produits des pixels correspondants. D'autres méthodes utilisent une structure d'image en pyramide [Kurita *et al.*, 1992]. Plus récemment, [Toyoda et Hasegawa, 2007] proposent une extension de HLAC pour la classification de textures et la reconnaissance de visages. Cette méthode permet de prendre en compte des ordres plus importants (jusqu'à l'ordre 8) et utilise également le principe de multi-résolution.

### 1.2.4.3 Les Matrices de Cooccurrence

[Haralick *et al.*, 1973, Haralick, 1979] proposent d'extraire des statistiques à partir d'une matrice de cooccurrence (GLCM) afin de caractériser une texture. Les matrices de cooccurrence d'Haralick, encore appelées matrices de dépendance spatiale des niveaux de gris, utilisent les statistiques d'ordre 2 et permettent de déterminer la fréquence d'apparition d'un motif formé de deux pixels séparés par une certaine distance  $d$  dans une direction particulière  $\theta$ .

Cette matrice est une matrice carrée de taille  $n^2$  où  $n$  correspond au nombre des niveaux de

## 1.2. MÉTHODES D'EXTRACTION DE CARACTÉRISTIQUES DE TEXTURES

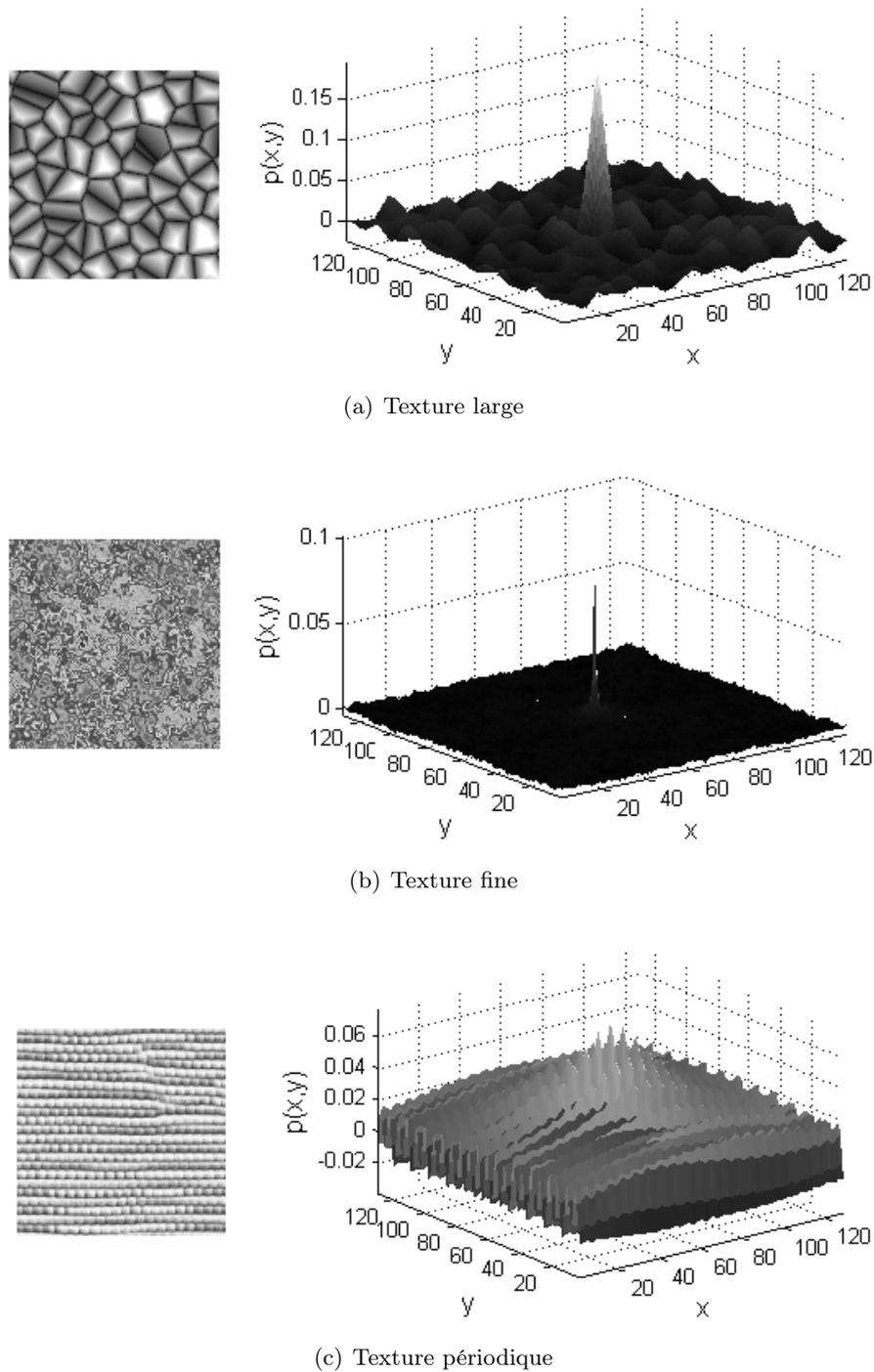


FIG. 1.22 – Exemple de fonctions d'auto-corrélation

gris de l'image. Généralement les niveaux de gris de l'image sont requantifiés (16, 32 ou 64 niveaux de gris) ce qui permet d'avoir un temps de traitement plus restreint. Également,

la requantification permet d'avoir un nombre d'occurrences moyen plus important pour chaque paire de niveaux de gris de la matrice. La normalisation de cette matrice produit une distribution de probabilité. Pour une image  $I$ , une matrice de cooccurrence des niveaux de gris normalisés  $P_v$  pour un vecteur de séparation donné  $v = (v_x, v_y)$  est définie par ses composantes  $P_v(i, j)$  données par l'expression suivante :

$$P_v(i, j) = \frac{\#\{(a, b) : I(a, b) = i, I(a + v_x, b + v_y) = j\}}{\#\{I\}} \quad (1.57)$$

avec  $\#$  la fonction cardinale,  $I(a, b)$  le niveau de gris de l'image  $I$  aux coordonnées  $(a, b)$  et  $i, j \in \{1, 2, \dots, n\}$

[Haralick, 1979] suggère de calculer des matrices de cooccurrence avec différentes directions et d'en faire la moyenne. De cette façon une matrice de cooccurrence peut être un descripteur de textures invariant à la rotation.

La figure 1.23 présente différentes matrices de cooccurrence de niveaux de gris obtenues à partir de plusieurs textures de Brodatz. Celles-ci sont définies sur 16 niveaux de gris et calculées selon le vecteur de séparation  $v = (1, 0)$ . Plus une zone est claire, et plus le couple correspondant de niveaux de gris est présent.

A partir de cette matrice, il est possible de calculer plusieurs caractéristiques [Haralick *et al.*, 1973, Haralick, 1979]. Parmi les plus fréquemment utilisées, nous retrouvons : la moyenne, la variance, le contraste, l'énergie (second moment angulaire), l'entropie, la corrélation, l'homogénéité et l'uniformité. Dans l'annexe A, nous présentons les caractéristiques d'Haralick que nous utilisons par la suite dans cette thèse.

Un certain nombre de variations des matrices de cooccurrence a été proposé dans la littérature. Une matrice de cooccurrence multidimensionnelle généralisée est présentée dans [Kovalev et Petrou, 1996]. Ils proposent d'utiliser des matrices à  $N$  dimensions, ce qui permet de mesurer le nombre d'occurrences d'attributs comme les niveaux de gris ou la magnitude d'un gradient local. Il est également possible d'associer des relations aux éléments de données. Cette variation des matrices de cooccurrence est utilisée dans plusieurs domaines et notamment dans le domaine médical : l'analyse d'images IRM du cerveau [Kovalev *et al.*, 2001], la détection des effets de l'âge et du sexe sur l'asymétrie structurelle du cerveau [Kovalev *et al.*, 2003a], la détection des patients schizophrènes [Kovalev *et al.*, 2003b], la mesure de l'anisotropie d'une texture [Kovalev *et al.*, 1999] etc.

#### 1.2.4.4 Les filtres de convolution

Les filtres de convolution sont utilisés comme des fenêtres glissantes permettant d'extraire les mesures désirées dans une image. Selon les formes présentes (frontières horizontales, verticales, diagonales, bulles etc.) ces opérateurs donneront en résultat des réponses différentes. Ces filtres sont représentés en discret sous forme de grilles de taille  $2^2, 3^2, 5^2, \dots$  en 2D,  $2^3, 3^3, 5^3, \dots$  en 3D.

[Laws, 1980b] propose d'utiliser un ensemble de filtres construits à partir de 3 vecteurs comprenant :  $[-1 \ 0 \ 1]$  utilisé pour les frontières,  $[1 \ 2 \ 1]$  utilisé pour moyenner,  $[-1 \ 2 \ -1]$  utilisé pour détecter les taches. En utilisant ces 3 vecteurs, ces filtres peuvent facilement être étendus à la 3D. Cependant, le nombre de combinaisons possibles en 3D est important,

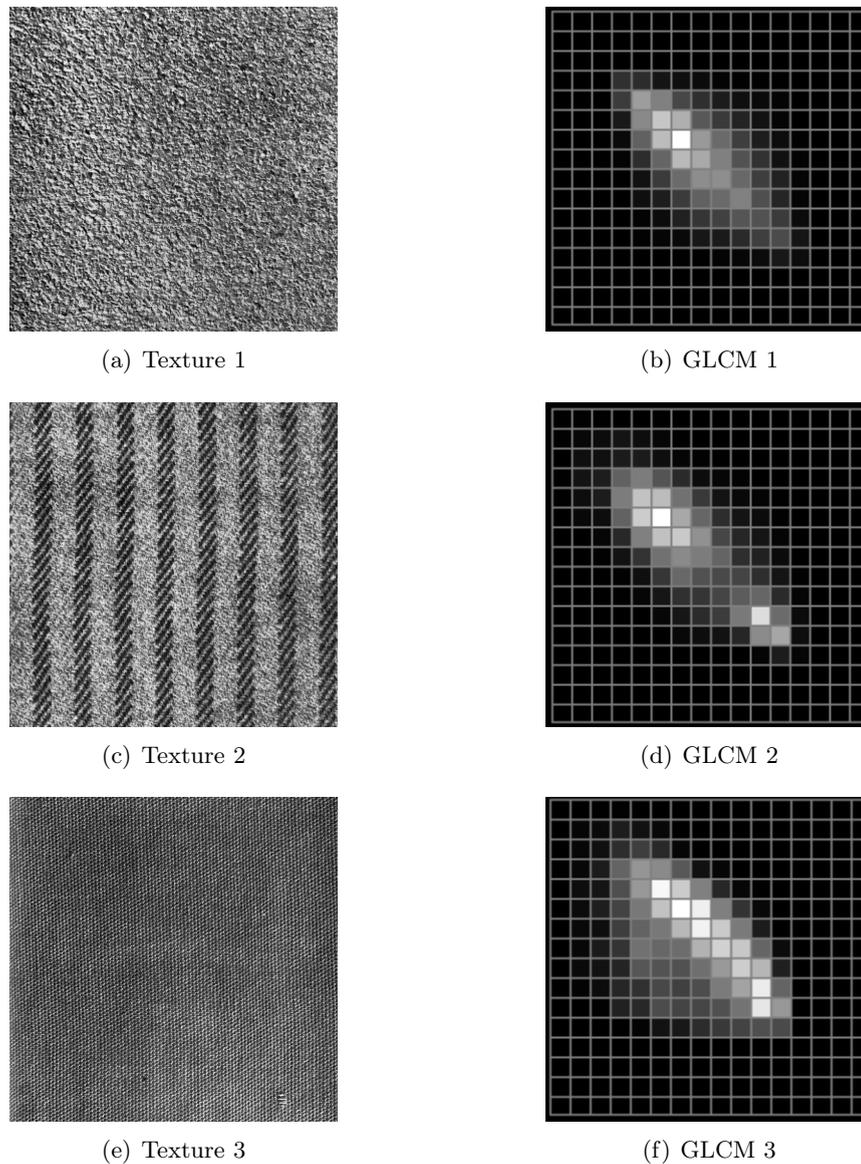


FIG. 1.23 – Exemple de matrices de cooccurrences de niveaux de gris sur des textures de Brodatz

ce qui pose problème pour le temps de génération des caractéristiques mais aussi pour le choix des caractéristiques discriminantes. Dans [Lang *et al.*, 1991], l'auteur les utilise pour analyser les structures fibreuses dans les muscles à partir d'images microscopiques confocales.

Les filtres différentiels sont également utilisés en analyse de textures. La dérivée partielle en chaque orientation  $(x, y, z)$  est obtenue par une simple soustraction avec

$\begin{bmatrix} +1 & -1 \end{bmatrix}$  ou  $\begin{bmatrix} +1 \\ -1 \end{bmatrix}$ . L'opérateur de Roberts permet d'obtenir les différences sur les

## 1.2. MÉTHODES D'EXTRACTION DE CARACTÉRISTIQUES DE TEXTURES

diagonales  $\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix}$ ,  $\begin{bmatrix} -1 & 0 \\ 0 & +1 \end{bmatrix}$ . D'autres opérateur comme Sobel  $\begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}$ ,

$\begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$  ou Prewit  $\begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix}$ ,  $\begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$  permettent de calculer

le gradient. Le filtre de Zucker-Hummel [Zucker et Hummel, 1981]

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ,  $\begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -1 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix}$  est également un opérateur de calcul du gradient et celui-

ci est utilisé dans de nombreuses applications d'analyse de textures [Kovalev *et al.*, 2001, Kovalev *et al.*, 2003a, Kovalev *et al.*, 1999, Kovalev *et al.*, 2003b].

Dans [Bernasconi *et al.*, 2001], l'auteur utilise le module du gradient comme une mesure de textures afin d'analyser les transitions entre la matière grise et la matière blanche dans des images IRM du cerveau. [Randen *et al.*, 2000] utilisent le gradient afin de déterminer les attributs de textures d'orientation et d'enfoncement dans des applications sismiques. Pour cela, ils calculent une matrice de cooccurrence locale construite à partir du gradient plutôt que sur l'image initiale.

### 1.2.4.5 Bilan

La description de textures à partir de telles analyses statistiques constitue le moyen le plus fréquemment utilisé, quel que soit le type d'images à traiter (notamment pour ce qui concerne les matrices de cooccurrence d'Haralick). Nous pensons que ceci est dû à la simplicité de mise en œuvre et de compréhension des concepts manipulés. De même, les caractéristiques produites (entropie, homogénéité, moyenne, contraste, etc.) correspondent à des notions fréquemment manipulées et assimilables par les utilisateurs. Par contre, ces méthodes paraissent assez sensibles au paramétrage utilisé qui doit être adapté pour chaque type d'images et nécessite d'avoir d'importantes connaissances *a priori* sur les textures à traiter. Ce point négatif en contradiction avec nos objectifs de généralité ne nous incitera pas à mettre en œuvre ce type de méthodes.

## 1.2.5 Vers une caractérisation plus visuelle des textures

D'autres auteurs proposent de combiner plusieurs attributs de textures afin d'approximer le système visuel humain [Tamura *et al.*, 1978, Amadasun et King, 1989]. En effet, le système visuel humain est efficace sur tous types de textures même en l'absence d'un contexte favorable. Dans [Tamura *et al.*, 1978], les auteurs proposent de décrire une texture à partir de 6 caractéristiques quantifiables par un être humain : la rugosité, le contraste, la directionnalité, la linéarité, la finesse et la régularité. Afin de tester les méthodes de calcul proposées pour chaque caractéristique, ils proposent de construire un ensemble d'expérimentations psychologiques et de les comparer. Pour cela, ils demandent à 46 sujets de clas-

sifier (d'ordonner) différentes textures selon les différentes caractéristiques proposées. Des tests de corrélation montrent une très bonne correspondance de leurs caractéristiques avec la perception humaine. Ils concluent également en disant que la rugosité, la directionnalité et le contraste sont des caractéristiques importantes pour la séparabilité des textures. De la même façon, [Amadasun et King, 1989] définissent plusieurs caractéristiques de textures avec la finesse, le contraste, *busyness* (correspondant à la rugosité) et la complexité (mesure du nombre de primitives de base). En utilisant les caractéristiques proposées par Tamura et al ainsi que Amadasun et King, Ravishankar Rao et Lohse dans [Rao et Lohse, 1993a] proposent d'identifier les plus pertinentes. Pour cela, ils demandent à 20 sujets d'effectuer une classification sur 30 textures et de les regrouper dans autant de classes que nécessaire. A partir de cette classification, ils construisent une matrice de similarité et en concluent qu'une texture peut être classifiée à partir de trois composantes : la répétition, l'orientation et la complexité. Dans [Rao et Lohse, 1993b], les mêmes auteurs, présentent une série de tests identiques à ceux présentés dans l'article précédent. La différence est qu'ils utilisent à la fois des données avec et sans métriques et le nombre de textures 2D devient plus important. Les résultats présentés affinent ceux décrits dans [Rao et Lohse, 1993a] et les auteurs proposent une représentation des textures selon trois dimensions : "répétitive contre non-répétitive", "fort-contraste et non-directionnel contre faible-contraste et directionnel" et "Granuleux, grossier et faible-complexité contre non-granuleux, fin et forte-complexité". Cette modélisation présente l'avantage de décrire une texture de façon très simple avec une métrique sur trois dimensions, chacune d'entre elles étant représentative d'une combinaison de caractéristiques. Cependant, même s'il est très probable qu'une texture fine ait une forte complexité, il est également hautement probable qu'une texture soit séparable selon ces deux caractéristiques. Nous pouvons alors nous demander si une caractérisation sur trois dimensions est suffisamment précise pour séparer deux textures différentes ayant des propriétés similaires.

Les différents travaux semblent montrer qu'il est plus efficace, pour produire ces descripteurs, de combiner des méthodes de caractérisation de textures appartenant aux différentes familles que nous avons identifiées préalablement. [Arivazhagan et Ganesan, 2003] proposent d'utiliser à la fois les matrices de cooccurrence des niveaux de gris (méthode statistique) et la décomposition en ondelette discrète (méthode fréquentielle). Parmi l'ensemble des méthodes d'analyse de textures comparées dans [Stachowiak *et al.*, 2005], cette dernière se révèle être la plus performante.

Parmi les "caractéristiques visuelles", la rugosité et la directionnalité sont très souvent utilisées. Dans la suite, nous proposons une revue des méthodes principalement utilisées pour les estimer.

### 1.2.5.1 Méthodes de calcul de la rugosité

Une texture rugueuse est une texture dont la surface présente différentes aspérités. Dans la littérature, plusieurs méthodes sont utilisées pour estimer la rugosité présente dans une image. Parmi elles, nous pouvons citer les méthodes d'estimation de la dimension fractale dont nous avons déjà discuté à la section 1.2.3.2. En effet, la dimension fractale correspond à la perception humaine de la rugosité d'une image [Novianto *et al.*, 2003]. Dans

[Soille et Rivest, 1996], les auteurs décrivent plusieurs méthodes pour estimer la dimension fractale. L'une des méthodes les plus connues est la méthode *Differential Box Counting* proposée dans [Sarkar et Chaudhuri, 1992, Sarkar et Chaudhuri, 1994]. En terme de temps d'exécution cette approche est la plus efficace. [Tsai et Tseng, 1999] proposent d'estimer la rugosité d'une surface en utilisant le spectre de puissance de la transformée de Fourier. Celui-ci est alors divisé en plusieurs régions anneaux. Une caractéristique est calculée pour chacune d'entre elles afin d'obtenir un vecteur représentatif de la rugosité de l'image de départ. De nombreuses mesures statistiques ont également été proposées pour déterminer la rugosité d'une image. Dans [Al-Kindi *et al.*, 1992], les auteurs mesurent la rugosité d'une surface en se basant sur l'espacement mais aussi la cardinalité des pics de niveaux de gris sur une ligne de l'image. Dans [Luk et Huynh, 1987], il est proposé d'utiliser la moyenne et la variance de l'histogramme des niveaux de gris afin de caractériser la rugosité. De la même façon, [Zhongxiang *et al.*, 2008] utilisent l'histogramme des niveaux de gris pour en déduire la rugosité à partir de 4 caractéristiques : la moyenne, la variance, la racine carrée moyenne de la déviation de la surface représentant l'amplitude de la surface, la pente des pics.

### 1.2.5.2 Méthodes de caractérisation de la directionnalité

De nombreuses approches ont été développées pour estimer les orientations principales d'une image. Parmi celles-ci, nous retrouvons notamment les approches dérivatives. [Lyvers et Mitchell, 1988] proposent une étude de ce type de méthodes et font une distinction sur les filtres gradients selon que leur réponse impulsionnelle est finie (FIR) ou infinie (IIR). Parmi les filtres FIR, les plus populaires sont les filtres de Sobel ou de Prewitt. Néanmoins, ce type de filtres montre une certaine sensibilité au bruit de part la faible taille de leur support. Concernant les filtres IIR, leur construction se fait souvent à partir de dérivées de Gaussienne [Kass et Witkin, 1987, Rao, 1990, Rao et Schunck, 1991, Rao et Jain, 1992]. Ces approches sont peu sensibles au bruit, cependant elles ne garantissent pas une estimation précise des orientations.

D'autres méthodes comme les tenseurs d'orientations permettent d'obtenir des informations sur l'orientation à partir de la matrice de covariance du champ local de gradient [Kass et Witkin, 1987, Rao, 1990, Rao et Schunck, 1991, Rao et Jain, 1992]. D'après [Germain, 2007], ces approches s'éloignent de la notion d'estimation de l'orientation locale dans le sens où elles calculent des statistiques spatiales sur des gradients locaux.

Les méthodes basées sur les statistiques d'ordre 2 permettent également d'estimer l'orientation. En effet, avec les matrices de cooccurrence [Haralick *et al.*, 1973, Haralick, 1979], il est possible de retrouver la structure de placement des primitives d'une texture. Cependant, le nombre de matrices à générer est très important, ce qui augmente grandement les temps d'exécution. Les histogrammes des différences des niveaux de gris étendus [Chetverikov, 1995] sont une autre alternative des statistiques d'ordre 2 pour le calcul de l'orientation. Néanmoins, tout comme les matrices de cooccurrence, cette méthode engendre un coût calculatoire élevé.

Proposé par [Bres, 1994], la rose des directions d'une image peut être définie comme une courbe polaire permettant de mesurer la force des orientations à partir de la fonction d'auto-corrélation de cette même image. Dans [Eglin, 1998], l'auteur définit la fonction

d'auto-corrélation comme étant le regroupement de l'ensemble des valeurs que l'on peut obtenir en faisant la somme de tous les produits des niveaux de gris. Ainsi un point de la fonction d'auto-corrélation correspond à la somme des produits de l'ensemble des pixels d'une image pour une translation donnée. De cette façon, une image est explorée selon plusieurs directions, ce qui permet par la fonction d'auto-corrélation de mettre en évidence les directions principales. Pour étudier les orientations d'une image, cette méthode se révèle comme étant très précise. Cependant, dans un faible voisinage, cette méthode est sensible au bruit. De plus, dans le cadre d'une segmentation, l'exploration des fonctions d'auto-corrélation sur chacun des pixels/voxels de l'image se révèle coûteuse en temps de calcul. Les bancs de filtres sont également utilisés pour l'étude des orientations dans les images. Contrairement aux méthodes dérivatives, ils permettent de détecter les orientations locales multiples. Il existe plusieurs types de filtres orientés avec par exemple les filtres en quadrature [Granlund et Knutsson, 1995], les filtres de Gabor [Chen *et al.*, 2000], ou les filtres Steerable [Sommer *et al.*, 1998]. Cependant, la précision de ce type de méthodes dépend beaucoup de leurs paramètres, à savoir le nombre de filtres, la taille des masques, etc. Dans certains cas d'application, elles peuvent également se révéler coûteuses en temps de calcul.

#### 1.2.5.3 Bilan

Ces approches combinent des méthodes de caractérisation et tentent de créer des "caractéristiques visuelles" facilement manipulables par les utilisateurs et correspondent parfaitement à nos objectifs initiaux. Il nous semble donc pertinent d'aller plus loin dans ce sens afin d'essayer de combler certaines lacunes des propositions passées. Dans [Tamura *et al.*, 1978, Amadasun et King, 1989], les auteurs démontrent la correspondance de leurs caractéristiques avec le système visuel humain mais ne comparent pas la performance de leurs descripteurs à d'autres attributs de textures. De plus, la plupart de ces propositions ne prennent pas en considération la multirésolution. Enfin, le dernier aspect primordial réside dans la proposition de modes d'exploitation concrets de ces combinaisons de caractéristiques visuelles dans le cadre d'applications réelles. En effet, il s'agit de démontrer que l'utilisabilité annoncée de ces caractéristiques est effective pour les utilisateurs.

## 1.3 Caractérisation des textures solides

Avant toute chose, il est préférable de rappeler que les textures solides, aussi appelées textures volumétriques, sont des textures représentées dans l'espace 3D qui peuvent être considérées comme des séries d'images texturées 2D ou comme un ensemble de voxels situé dans un volume. Celles-ci sont différentes des textures 3D ou des textures volumiques. Les textures 3D [Cuba et Dana, 2004] désignent la texture 2D d'un objet 3D vu d'un angle particulier avec différentes conditions d'éclairage. Les textures volumiques [Neyret, 1995] correspondent au rendu de reflets et géométries répétitives à l'intérieur une image 2D.

Comme nous l'avons vu tout au long de cet état de l'art, beaucoup d'efforts ont été menés pour analyser les textures en 2 dimensions. Cependant, peu de méthodes ont été étendues en vue d'analyser les textures solides. L'analyse des textures est présente dans de nombreux domaines comme l'analyse d'images médicales ou le traitement de

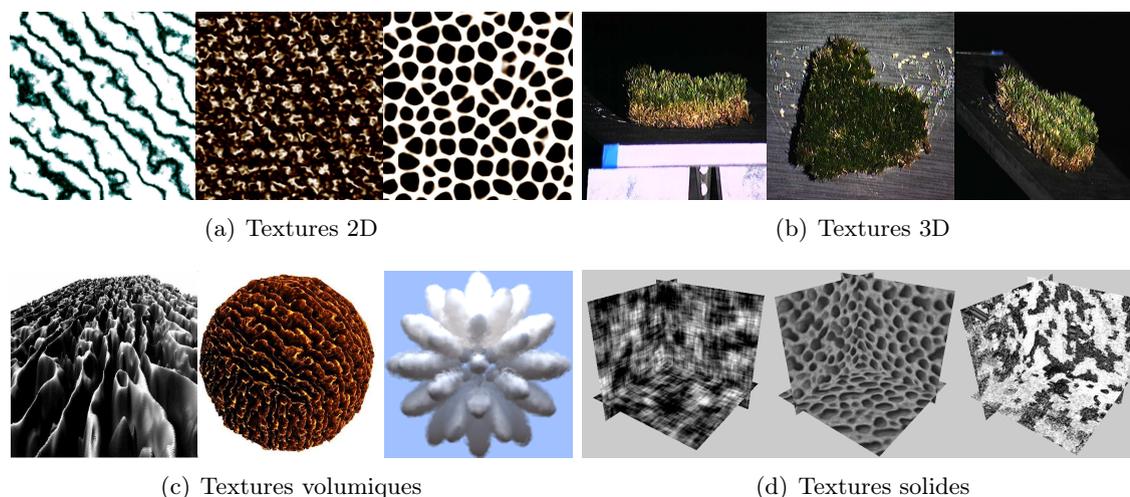


FIG. 1.24 – Illustration sur différents types de représentations des textures

documents. En imagerie médicale, l'analyse des textures est fréquemment utilisée car elle permet de segmenter de façon robuste des images complexes comme par exemple les images ultrasonores. Elle est également très utile pour caractériser un tissu afin d'aider les médecins dans leurs diagnostics. Le progrès des technologies a comme conséquence l'accroissement de l'utilisation d'images 3D, notamment en imagerie médicale avec les images à résonance magnétique, la radiographie digitale, les images échographiques etc. Ce besoin de traiter des images médicales 3D, a fait que certaines méthodes d'analyses de textures ont été étendues au domaine 3D assez récemment.

La plus simple façon d'étudier une image 3D est d'utiliser l'une des méthodes 2D d'analyse de textures sur chacune des coupes 2D texturée qui la composent. Nous ne parlerons pas ici de ce type de méthodes qui n'exploite aucunement les informations 3D disponibles.

Dans [Suzuki *et al.*, 2004], les auteurs proposent une extension de HLAC (*High order local autocorrelation*) à la troisième dimension. Les données 3D sont analysées à partir de masques HLAC 3D qui sont des cubes divisés en grilles de dimension  $3 \times 3 \times 3$ . Pour tester leur méthode, ils génèrent deux types de textures volumétriques : des textures basées sur des fonctions *Perlin's noise* [Perlin, 1985] et des textures construites à partir de fonctions *Perlin's noise* avec un bruit fractal.

Dans la littérature, beaucoup d'articles proposent d'utiliser une extension 3D des matrices de cooccurrence des niveaux de gris. [IP et Lam, 1994] sont les premiers à utiliser une matrice de cooccurrence avec des données 3D. Les textures solides sont séparées via une structure d'*octree* en plusieurs partitions. Chaque noeud de l'*octree* représente une partition, et les arêtes, les relations de voisinage entre 2 partitions. Pour chacun des noeuds, des caractéristiques d'Haralick 3D sont calculées. Les noeuds ayant des caractéristiques similaires sont fusionnés afin de produire une segmentation. [Kovalev *et al.*, 1999] proposent deux approches pour la caractérisation de textures solides. La première méthode utilise un histogramme des orientations 3D calculé à partir de vecteurs de gradient selon plusieurs orientations. La deuxième méthode est une extension 3D des matrices de cooccurrence de Chetverikov. A partir de ces deux méthodes, ils caractérisent des textures

### 1.3. CARACTÉRISATION DES TEXTURES SOLIDES

---

synthétiques avec différents niveaux de bruit, ainsi que des images médicales en vue de surveiller l'évolution de pathologies. Ces deux approches permettent de caractériser précisément l'anisotropie d'une texture, cependant, pour un problème de classification il faudrait associer ces caractéristiques avec d'autres descripteurs de textures pour avoir des informations supplémentaires.

Dans [Ghoneim *et al.*, 2003], les auteurs proposent une classification de tumeurs dans le cerveau par l'analyse de textures. Pour cela, ils calculent des caractéristiques de textures à partir des matrices de cooccurrence d'Haralick 2D mais aussi 3D. Dans tous les cas, les caractéristiques d'Haralick calculées à partir de la 3D donnent de meilleurs résultats que les matrices de cooccurrence 2D. En utilisant la troisième dimension, l'information concernant la distribution des niveaux de gris est plus riche. De la même façon, [Showalter *et al.*, 2006] utilisent les caractéristiques de textures d'Haralick 3D pour prédire des propriétés de micro-architectures des os.

[Kovalev *et al.*, 2003b] utilisent le module du gradient calculé à partir du filtre de Zucker-Hummel. Les amplitudes calculées leur permettent de construire une matrice de cooccurrence 3D. En interprétant les résultats, les auteurs concluent que toutes les coupes du cerveau ne sont pas nécessaires pour déterminer si un sujet est schizophrène. Dans [Kovalev *et al.*, 2001], les auteurs proposent de construire une matrice de cooccurrence en utilisant les niveaux de gris et le module du gradient. Cette matrice est utilisée pour la discrimination d'ensemble de données sur le cerveau de patients ayant des perturbations cognitives légères. Ils proposent également d'utiliser cette technique pour la segmentation de lésions dans le cerveau.

Pour la classification de motifs de sous-cellule, les auteurs utilisent dans [Chen et Murphy, 2004] une combinaison de deux méthodes de caractérisation de textures volumétriques avec les matrices de cooccurrence d'Haralick 3D et des caractéristiques morphologiques et frontières. Les auteurs concluent en disant que la combinaison de caractéristiques permet d'améliorer les résultats de classification.

Parmi les méthodes utilisées pour analyser les textures solides, nous retrouvons les méthodes fréquentielles avec notamment les filtres de Gabor et les ondelettes discrètes. [Jafari-Khouzani *et al.*, 2004] cherchent à caractériser des hippocampes dans un cerveau à partir d'images 3D à résonance magnétique. Pour cela, ils calculent l'énergie moyenne sur les coefficients d'ondelette obtenus à partir d'une décomposition en ondelette discrète 2D sur chacune des coupes de l'hippocampe. De la même façon, ces caractéristiques sont calculées, mais cette fois-ci en utilisant une décomposition en ondelette 3D sur le volume de l'hippocampe. Leurs résultats montrent que sur ce type d'images, la décomposition en ondelette 2D permet une meilleur séparabilité qu'avec les ondelettes 3D.

[Zhan et Shen, 2006] présentent un modèle déformable et l'appliquent sur des images échographiques 3D. Pour approximer l'ensemble des caractéristiques qu'il est possible d'obtenir avec des filtres de Gabor 3D, ils calculent des caractéristiques en utilisant deux bancs de filtres de Gabor 2D positionnés dans les deux plans orthogonaux. Cette méthode permet d'alléger les temps de calcul par rapport à un filtre 3D, mais elle entraîne une perte d'information non-négligeable. [Bigun, 1994] fait une présentation des filtres de Gabor 3D. [François *et al.*, 2003] les appliquent pour segmenter des images ultrasonores 3D de carotide. En utilisant l'information de textures obtenue à partir des filtres de Gabor 3D, ils définissent une mesure de similarité qui permet de labéliser les différentes zones

de l'image. Dans [Fernandez *et al.*, 2000], les auteurs segmentent des zones sismiques 3D en utilisant des caractéristiques obtenues à partir de filtres de Gabor 3D ainsi qu'à partir d'une décomposition en ondelette par paquet. [Shen et Bai, 2008] proposent de régler les paramètres de l'algorithme SPM (*Statistical Parametric Mapping*) en utilisant l'entropie calculée à partir de la réponse maximale locale d'une ondelette de Gabor 3D. [Aldasoro et Bhalerao, 2007] proposent une méthode de segmentation multirésolution pour la segmentation de textures solides. Pour cela, ils extraient des caractéristiques du domaine de Fourier à partir de sous-bandes [Wilson et Spann, 1988] en utilisant une pyramide d'orientation. Ils proposent également une sélection de caractéristiques à partir de l'espace *Bhattacharya* basée sur une distance *Bhattacharya*. Après avoir construit un *oct tree*, ils proposent d'identifier les différentes classes à une résolution spatiale faible puis de les répercuter à la résolution de l'image d'origine. Ils appliquent cette méthode sur des images à résonance magnétique de genoux humains et obtiennent de bons résultats. Il pourrait être intéressant de voir les performances de cette méthode sur des images médicales plus difficiles à traiter comme par exemple des images échographiques 3D. La conclusion de ce panorama de méthodes d'analyse volumétrique est qu'il s'agit pour l'instant de travaux émergeant. Il est donc nécessaire d'explorer plus profondément le sujet pour produire des méthodes exploitant encore mieux les informations 3D disponibles. Il faut noter qu'aucune expérimentation ou étude n'a été menée jusqu'à présent pour essayer de proposer ou construire une "caractérisation visuelle ou qualificative" dans le cadre de la description de textures solides.

### 1.4 Méthodes d'exploitation des caractéristiques de textures

Dans la section précédente, nous avons décrit une grande partie des méthodes utilisées en analyse de textures. Il nous paraît également important de faire un rapide bilan sur les différents modes d'exploitation possibles des descripteurs obtenus. Deux grandes catégories de problématiques se dégagent : les cadres applicatifs faisant intervenir des problèmes de classification et les problèmes de segmentation d'images.

#### 1.4.1 Méthodes de classification

La classification de textures consiste à associer une classe ou un label, à une texture donnée, en fonction des caractéristiques présentes. Une première étape importante est donc l'extraction des caractéristiques de textures. Ces attributs de textures sont calculés à partir des méthodes que nous avons présentées à la section 1.2. Ces caractéristiques sont regroupées sous forme de vecteurs numériques et sont, par la suite, directement utilisées par un classificateur. Une deuxième étape importante du processus de classification est donc le choix du classificateur. Il est possible d'utiliser soit une approche supervisée, soit une approche non-supervisée.

Dans le cadre de la classification non-supervisée [Qin *et al.*, 2008], il est généralement nécessaire d'indiquer le nombre de classes recherchées. Les textures similaires sont alors regroupées à l'intérieur d'une même classe sans l'intervention de connaissances *a priori* et simplement à partir d'estimations de similarité entre caractéristiques.

Dans le cadre de l'indexation d'images, les approches supervisées [Li *et al.*, 2003] sont

généralement préférées. Parmi elles, nous pouvons citer les  $k$  plus proches voisins [Chitre et Dhawan, 1999], l'analyse discriminante linéaire [Clausi et Jernigan, 2000], la quantification vectorielle supervisée [Randen et Husóy, 1999], les classificateurs Bayésiens [Manian et Vasquez, 1998], les réseaux de neurones [Laine et Fan, 1993] mais aussi les machines à vecteur support [Vapnik, 1995]. L'utilisation des techniques supervisées nécessite une première phase d'apprentissage durant laquelle le système apprend les caractéristiques associées à chaque classe d'une base d'apprentissage. Une fois cette étape réalisée, ces méthodes supervisées sont capables de classer les individus d'une base de test en fonction des caractéristiques de textures qui leur sont associées. Les applications en classification de textures sont très nombreuses et constituent encore aujourd'hui un pôle de recherche important [Laine et Fan, 1993, Valkealahti et Oja, 1998, Randen et Husóy, 1999, Li et Shawe-Taylor, 2005, Aldasoro et Bhalerao, 2007, Kim et Kang, 2007, Ferri *et al.*, 2008, Guo *et al.*, 2009].

### 1.4.2 Méthodes de segmentation de textures

La segmentation d'une image texturée consiste à diviser cette même image en plusieurs régions homogènes comportant des caractéristiques de textures identiques. Plusieurs approches utilisent des descripteurs de textures afin de segmenter une image. Parmi elles, nous pouvons citer les méthodes basées sur les contours actifs, la ligne de partage des eaux, et les méthodes de type *clustering* de pixels/voxels.

De manière générale, lorsqu'il s'agit de textures, les méthodes de segmentation basées contours sont proscrites puisqu'une texture ne comporte en général pas de frontières franches. Les approches régions sont plus adaptées, mais la faible homogénéité de l'intensité des pixels/voxels, habituellement exploitée par ce type d'approche, doit être remplacée par un mécanisme plus complexe.

La ligne de partage des eaux est utilisée afin de segmenter une image à partir d'une fonction de hauteur classiquement représentée par un gradient des intensités de niveaux de gris. Dans sa version de base, cette méthode n'est pas toujours efficace pour segmenter des régions texturées qui sont perceptuellement homogènes. Afin d'utiliser cette méthode pour la segmentation de textures, il est nécessaire d'établir une fonction de hauteur permettant de prendre en considération les différences entre l'ensemble des textures présentes dans l'image. Dans [Hill *et al.*, 2002, Malpica *et al.*, 2003], les auteurs proposent d'utiliser la notion de gradient de textures qui correspond à une fonction de hauteur construite à partir des vecteurs de caractéristiques sur chaque pixel/voxel.

Concernant les méthodes basées sur les contours actifs, les dernières avancées en recherche se focalisent sur l'intégration de caractéristiques de textures servant à guider le contour. Dans [Savelonas *et al.*, 2008], une méthode de contours actifs, guidée à partir des caractéristiques de textures LBP est proposée. Dans cet article, les auteurs présentent un état de l'art sur l'ensemble des techniques de contours actifs utilisant une information de textures. Ils les regroupent en deux catégories : les méthodes basées sur les ondelettes et les filtres de Gabor, et les méthodes basées sur des mesures statistiques.

Parmi les contours actifs utilisant des méthodes fréquentielles, nous retrouvons la méthode présentée dans [Sandberg *et al.*, 2002], proposant d'utiliser un banc de filtres de Gabor avec le modèle proposé dans [Chan *et al.*, 2002]. Dans [Paragios et Deriche, 1999], les au-

teurs proposent une méthode de segmentation supervisée utilisant un banc de filtres de Gabor permettant de guider le modèle de contour actif géodésique [Caselles *et al.*, 1997]. Il en existe encore un certain nombre utilisant des principes de fonctionnement similaires [He *et al.*, 2004, Liapis *et al.*, 2004, Sagiv *et al.*, 2006, Lecellier, 2009].

Parmi les méthodes basées sur les statistiques, nous pouvons citer [Lehmann *et al.*, 2001, Pujol et Radeva, 2004] qui proposent des méthodes supervisées de segmentation de textures utilisant les caractéristiques d'Haralick. D'autres méthodes, comme celle présentée dans [Rousson *et al.*, 2003], proposent une approche supervisée utilisant à la fois des caractéristiques statistiques et d'orientation.

Les méthodes de *clustering* sont non-supervisées et permettent de classifier et de regrouper les pixels/voxels d'une image ayant des caractéristiques de textures identiques. Contrairement aux méthodes de contours actifs, ces méthodes ne se limitent pas à l'exploration d'une zone de l'image et permettent d'obtenir une segmentation incluant l'ensemble des pixels/voxels d'une image. Parmi ces méthodes de clustering, souvent exploitées en segmentation de textures, nous pouvons citer les K-means [Coleman et Andrews, 1979], CLARA [Kaufman et Rousseeuw, 1990], CLARANS [Ng et Han, 1994] mais également l'algorithme de BIRCH [Zhang *et al.*, 1996].

Les méthodes présentées dans cette sous-section sont celles qui sont le plus classiquement utilisées et permettent la génération d'une segmentation à partir de l'utilisation de vecteurs de caractéristiques de textures ou bien d'un gradient de textures. D'autres méthodes comme les champs de Markov proposent une approche différente et permettent d'obtenir des segmentations à l'aide de métaheuristiques appliquées sur le modèle choisi (voir sous-section 1.2.3.1).

## 1.5 Comparaison de méthodes d'analyse de textures

Dans ce paragraphe, nous passons en revue les comparaisons et les évaluations de méthodes d'analyse de textures fournies par la littérature. Nous avons vu dans le paragraphe précédent que beaucoup d'efforts avaient été menés pour analyser les textures en 2D mais beaucoup moins dans le cadre de la 3D. De façon naturelle, les articles d'évaluation proposent uniquement des comparaisons de méthodes 2D. L'absence de travaux pour l'analyse de performance d'algorithme 3D s'explique sûrement par l'absence de bases de données de textures 3D suffisamment riche pour mettre en place un protocole d'évaluation. La mise en place d'une telle base de donnée constitue l'une de nos contributions que nous détaillons section 3.1.1.

Dans [Stachowiak *et al.*, 2005], les auteurs proposent de comparer six méthodes d'analyse de textures invariantes à la rotation et aux changements d'échelle : les matrices de cooccurrence [Haralick, 1979], les transformées en ondelette discrète, une combinaison des ondelettes et des caractéristiques de cooccurrence [Arivazhagan et Ganesan, 2003], les filtres de Gabor, les champs de Markov aléatoires Gaussiens [Chellappa et Jain, 1993, Cohen et Cooper, 1987, Chellappa et Chatterjee, 1985] et la méthode des motifs binaires (LBP)[Ojala *et al.*, 1996, Ojala *et al.*, 2001, Ojala et Pietikäinen, 2002]. Dans leur étude, ils proposent d'évaluer les différentes méthodes d'analyse de textures à l'aide d'un pro-

blème de classification utilisant quatre ensembles d'images comprenant des textures de Brodatz, des images fractales isotropiques, des images microscopiques (images naturelles) de surfaces sablées et de surfaces érodées. Pour chaque méthode, des résultats de classification sont générés à partir de deux classificateurs : les  $k$  plus proches voisins, et le classificateur en analyse discriminante linéaire. Leur étude montre que les filtres de Gabor et les champs de Markov aléatoires gaussiens donnent de bons résultats sur les textures de Brodatz. Pour les autres images, les résultats sont pauvres et peu satisfaisants. La méthode LBP/VAR affiche de très bonnes performances pour les textures de Brodatz et les images de type sablées. Cependant, le nombre d'erreurs de classification augmente lorsque la taille de l'image analysée diminue. L'erreur de classification est également élevée pour les images de type érodées et les images fractales. Les matrices de cooccurrence sont efficaces sur les images de type sablées mais leur performance est faible sur des images anisotropiques comme les surfaces érodées. La transformée en ondelette discrète donne de mauvais résultats sur les images fractales, par contre, sa performance sur les images de type sablées et de type érodées est très satisfaisante. Enfin, la combinaison entre la transformée en ondelette discrète et les caractéristiques de cooccurrence donne de bons résultats dans l'ensemble des cas ce qui nous conforte dans notre idée de mieux exploiter la combinaison de méthodes de caractérisation.

Dans [Sharma et Singh, 2001], les auteurs étudient 5 méthodes d'analyse de textures à travers un problème de classification en utilisant la base de donnée *Meastex*. Parmi elles, nous retrouvons : les matrices de cooccurrence des niveaux de gris, la fonction d'auto-corrélation, les filtres de Laws [Laws, 1980b], des détecteurs de frontières pour la description de textures, ainsi que des primitives de taille [Sharma et Singh, 2001]. Dans leur analyse, les auteurs utilisent une analyse discriminante linéaire mais aussi deux algorithmes modifiés des  $k$  plus proches voisins. Parmi toutes ces méthodes, les matrices de cooccurrence et la méthode de Laws donnent les meilleurs taux de reconnaissance. Les auteurs concluent en affirmant que les différentes méthodes d'analyse de textures permettent d'obtenir différents aspects des textures et qu'il pourrait être intéressant de les combiner.

[Unser, 1995] compare, à l'aide d'un problème de classification, les frames d'ondelette (DWF) avec la décomposition en ondelette discrète classique (DWT). Pour cela, 12 textures de la base de données de Brodatz et de dimension  $256^2$  pixels sont utilisées. Pour chaque texture, un total de 64 vecteurs de caractéristiques indépendantes est évalué (série de sous-régions non-superposées et de taille  $32^2$  pixels). Par la suite, les textures sont classifiées à l'aide d'un classificateur bayésien. Dans tous les cas, la DWF donne de meilleurs résultats que la DWT. En effet, l'approche DWF réduit la variabilité concernant les caractéristiques de textures, ce qui améliore les performances de classification. Il souligne également l'importance de la multirésolution pour obtenir une discrimination efficace des textures. Pour lui, les approches traditionnelles (matrices de cooccurrence [Haralick, 1979], les champs de Markov gaussien, etc.) sont restreintes à une analyse très locale, ce qui en fait des méthodes performantes pour l'analyse de micro-textures. Au contraire, les méthodes fréquentielles comme les filtres de Gabor ou les ondelettes permettent d'avoir des informations sur la texture qui sont à la fois locales et globales, ce qui permet une caractérisation de meilleure qualité.

[Ojala *et al.*, 1996] étudient, grâce à un test de classification, quatre techniques d'analyse de textures avec les matrices des différences de niveaux de gris GLDM [Unser, 1986] (*Grey*

*Level Difference matrices*(GLDM)), les mesures de covariance symétriques centrées (*center-symmetric covariance measures*(CSCM)), les filtres de Laws (*Laws texture method*(LTM)) et le descripteur LBP(*Local Binary Pattern*). Dans leur expérimentation, les textures à reconnaître sont obtenues en sous-échantillonnant aléatoirement les images texturées d'origine provenant de l'album de Brodatz. Les résultats de classification sont générés à partir d'une méthode basée sur la discrimination de Kullback. GLDM donne dans l'ensemble de très bons résultats. De plus, cette méthode n'est pas très coûteuse en temps de calcul. De la même manière, LBP offre de très bonnes performances pour une faible complexité. Cependant, cette méthode n'est pas robuste au changement d'intensité. Pour résoudre ce problème, Ojala fait une combinaison des caractéristiques LBP avec une simple mesure de contraste qui permet d'améliorer de façon considérable les performances de caractérisation. Enfin LTM et CSCM donnent de moins bons résultats et nécessitent des échantillons d'images plus importants que les méthodes LBP et GLDM.

De nombreuses études comparent les caractéristiques obtenues à partir de filtres et celles obtenues à partir de matrices de cooccurrence des niveaux de gris pour la classification de textures. Dans [Strand et Taxt, 1994], les auteurs concluent que les matrices de cooccurrence des niveaux de gris donnent de meilleurs résultats. Au contraire, les auteurs dans [Laws, 1980a, Pietikäinen *et al.*, 1983, Clausi et Jernigan, 1996] concluent que les méthodes avec filtres sont plus performantes.

[Chen et Chen, 1999] comparent, dans un problème de classification de textures, plusieurs méthodes de type fréquentiel avec la transformée de Fourier, le filtre spatial, les filtres de Gabor et la décomposition en ondelette discrète sur des images naturelles de la base de Brodatz et sur des images synthétiques obtenues à l'aide de champs de Markov. Lorsque le nombre de caractéristiques est restreint (4 caractéristiques), les filtres Gabor sont les plus performants. Cependant avec plus de caractéristiques (12 caractéristiques), les ondelettes calculées sur trois niveaux de décompositions donnent les meilleurs résultats avec très peu d'erreur de classification.

[Pichler *et al.*, 1996] comparent les ondelettes et les filtres de Gabor et ils concluent que les filtres de Gabor sont plus performants mais aux dépens d'un effort de réglage conséquent. Dans [Chang et Kuo, 1993], l'auteur affirme que les ondelettes sont plus naturelles et efficaces pour des textures avec une dominance de fréquences moyennes alors que les filtres de Gabor sont plus appropriés pour des images composées de basses fréquences.

## 1.6 Modélisation d'un système de reconnaissance de textures

Dans cette section, nous essayons d'identifier à travers une modélisation, les différents facteurs étant susceptibles d'intervenir dans un système de reconnaissance de textures. Pour caractériser une texture, le cheminement utilisé est très souvent le suivant : un ensemble de caractéristiques, généralement numériques, assimilables ou non par l'humain et en nombre plus ou moins important, sont calculées à partir d'une méthode d'analyse de textures. Une deuxième étape, facultative, consiste à sélectionner les caractéristiques les plus pertinentes dans l'ensemble initial. En lien avec cette étape, les différentes caractéristiques peuvent être combinées (linéairement ou non) afin de créer de nouveaux attributs de textures.

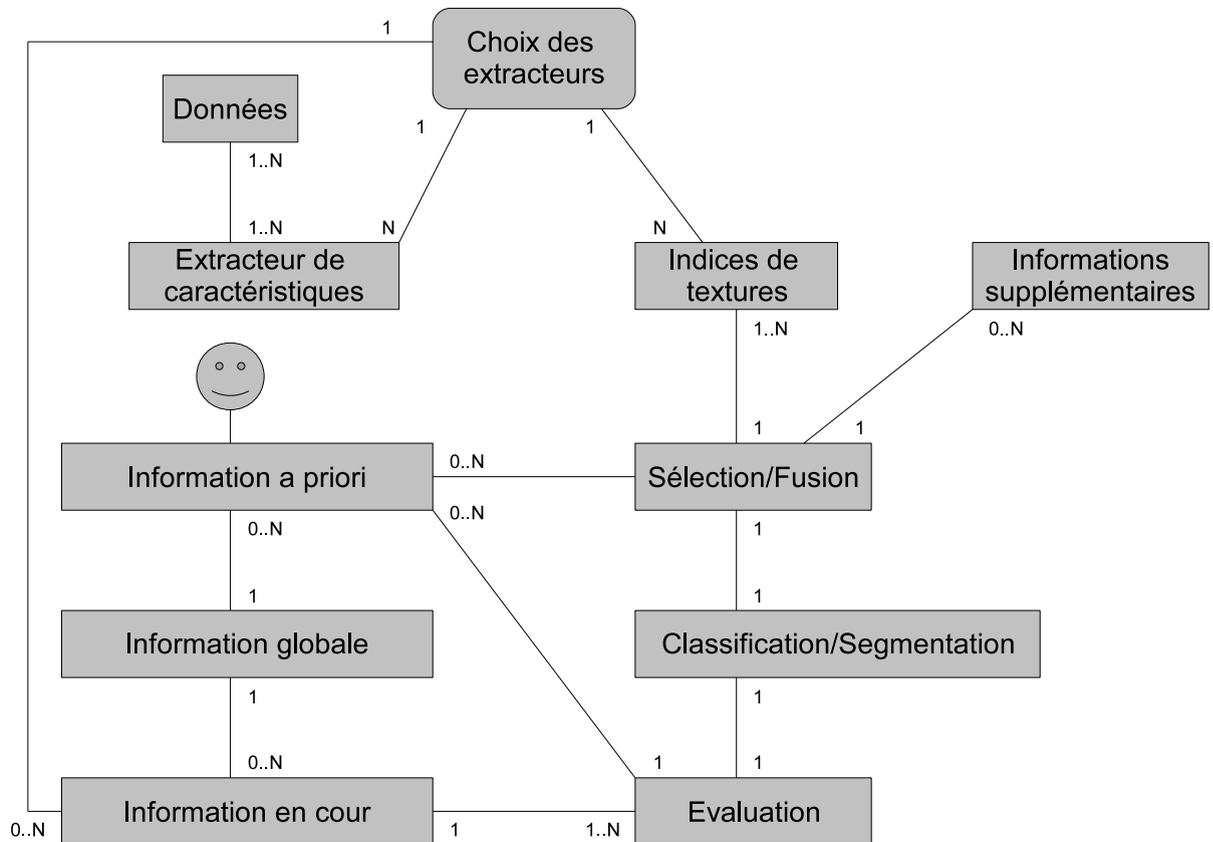


FIG. 1.25 – Système de reconnaissance de textures

Dans cette modélisation, nous faisons apparaître des notions supplémentaires avec en premier lieu ce que nous appelons les « caractéristiques supplémentaires ». En utilisant des informations dites non-texture, un système peut être capable de reconnaître une texture. Par exemple, l'être humain est capable d'utiliser des informations de forme (d'un objet) pour en déduire plus facilement des informations sur sa texture : l'objet identifié est un arbre donc la texture à reconnaître est de type « Bois ». Les « informations *a priori* » sont des connaissances apportées avant l'exécution du système et peuvent être fournies de différentes manières. L'utilisateur du système peut par exemple choisir de ne retenir qu'un certain nombre d'indices de textures car il juge qu'ils seront suffisants et nécessaires pour une bonne reconnaissance. Les « informations en cours » sont des informations amenées à mûrir au fur et à mesure de l'évolution du système. Dans cette modélisation, elles peuvent intervenir dans le choix des extracteurs de caractéristiques en se basant sur le résultat des différentes évaluations. En effet, il est possible d'utiliser plusieurs extracteurs pour obtenir une même caractéristique et le but est alors de ne retenir que les plus performants. Cette information peut également être utilisée pour sélectionner les indices de textures les plus pertinents. Enfin les « informations globales » font office de mémoire du système. Leur rôle est alors d'influencer le système dans le cas d'une situation déjà rencontrée ou très similaire.

Formellement, nous considérons  $D$  les données à traiter (images texturées) qui sont définies selon une dimension mais aussi selon leur espace de représentation (spatial ou fréquentiel). Nous définissons également  $A$  comme l'ensemble possible des extracteurs de textures sachant que chacun d'eux permet d'obtenir un ou plusieurs attributs de textures  $T$ . Nous pouvons alors écrire :

$$a_{i,j} \in A \text{ avec } i \in \{1..N\} \text{ et } j \in \{1..M_i\}$$

avec  $a_{i,j}$  l'extracteur  $j$  permettant d'obtenir la caractéristique  $i$ ,  $N$  le nombre de caractéristiques possibles et  $M_i$  le nombre d'extracteurs permettant d'obtenir la caractéristique  $i$ . Un indice de textures  $T_i$  s'exprime alors de la façon suivante :

$$a_{i,j} : D \rightarrow T_i$$

Un indice  $T_i$  est dépendant de la résolution mais aussi de la régionalité (voisinage) considérée.

Un ensemble de caractéristiques  $F = \{T_1, T_2, \dots\}$  est alors calculé à partir des données  $D$ , de l'ensemble des extracteurs  $A$  et des informations en cours  $K_{ec}$ .

$$Cei : D \times A \times K_{ec} \rightarrow F$$

avec  $Cei$  la fonction de choix des indices et des extracteurs.

Une sélection de caractéristiques peut alors être effectuée à la fois sur  $F$  l'ensemble des caractéristiques utilisées mais aussi sur  $O$  qui représente ce que nous appelons les caractéristiques supplémentaires :

$$Select : F \times O \times K_{ap} \rightarrow F'$$

$Select$  est la fonction de sélection,  $K_{ap}$  représente les informations *a priori* et  $F' \subseteq F$ .

Par la suite une fonction de classification ou segmentation  $ClassSeg$  est appliquée sur ces données :

$$ClassSeg : K_{ap} \times F' \rightarrow R$$

$R$  représentant un résultat sous forme de labels ou de frontières selon la problématique.

Enfin une fonction d'évaluation  $Eval$  permet de vérifier la pertinence des résultats. Selon le type d'évaluation (évaluation supervisée ou non-supervisée) une connaissance *a priori* intervient dans le système.

$$Eval : K_{ap} \times R \rightarrow L$$

avec  $L$  un indice ou un ensemble d'indices de qualités.

Cette modélisation met en avant toutes les difficultés qu'il est possible de rencontrer dans un système de reconnaissance de textures. En effet, il est nécessaire de traiter plusieurs problématiques telles que le choix ou le calcul des caractéristiques de textures, le choix des méthodes de classification ou de segmentation, la mise en place de méthodes d'intelligence artificielle permettant d'exploiter les informations du connexes etc. Dans cette thèse, nous nous focaliserons sur le choix des caractéristiques à exploiter avec une attention particulière sur l'interaction Homme-Machine. Beaucoup d'extracteurs de textures proposent des caractéristiques difficilement compréhensibles par un être humain ce qui rend impossible leur exploitation dans un cadre applicatif concret. Les propriétés que nous recherchons impliquent de définir une texture à partir d'attributs visuels correspondant à des adjectifs qualificatifs non ambigus du langage naturel.

## 1.7 Conclusion

L'analyse des textures est extrêmement répandue pour le traitement d'images 2D et s'étend actuellement à la 3D. Nous avons présenté dans cet état de l'art une grande panoplie des méthodes permettant de caractériser les textures afin de traiter des problèmes de classification et de segmentation d'images texturées. Cette étude nous a permis de mettre en évidence plusieurs points cruciaux concernant l'analyse de textures 2D et volumétriques, soulignant notamment les avantages et inconvénients de chaque approche. Les points fondamentaux que nous avons mentionnés sont la régionalité, la multirésolution, la facilité de manipulation, la sélection et combinaison de caractéristiques et leur adaptation aux textures solides.

Dans une problématique de segmentation, le choix du voisinage (régionalité) est primordial et dépendant du type de textures traitées. La régionalité doit être suffisamment large pour contenir l'ensemble ou une grande partie des motifs de base de la texture considérée. Dans [Jan et Hsueh, 1998], les auteurs proposent de prédire la taille de la fenêtre de voisinage sur des textures structurées (déterministes, observables). Pour cela, ils utilisent les propriétés de périodicité de ce type de textures. Cependant, cette problématique reste encore d'actualité et il serait intéressant de mettre en place des méthodes efficaces quel que soit le type de textures traité.

Comme mentionné par les auteurs dans [Randen et Husøy, 1999], il est difficile d'affirmer qu'une méthode d'analyse de textures est clairement la plus performante. Sur l'ensemble des articles faisant des comparatifs de méthodes d'analyse de textures, certains aboutissent à des conclusions contradictoires. Chaque évaluation est faite à partir de bases de données et de critères différents, ce qui peut expliquer la divergence de certains résultats. Si nous faisons une synthèse sur l'ensemble des comparaisons proposées dans la littérature, il en ressort tout de même une préférence pour les méthodes fréquentielles [Pietikäinen *et al.*, 1983, Unser, 1995, Clausi et Jernigan, 1996, Randen et Husøy, 1999, Reyes-Aldasoro, 2004, Stachowiak *et al.*, 2005]. De façon générale, ces méthodes décrivent une texture localement et globalement, ce qui permet une caractérisation plus complète. En effet, ces approches intègrent la notion de multi-résolution ce qui permet de caractériser une texture selon plusieurs échelles d'observation. Tout comme la régionalité, la résolution est un paramètre important et il n'est pas toujours évident de savoir quel sont les résolutions qu'il est nécessaire d'explorer. Il est également important de souligner les bonnes performances que procure la combinaison de plusieurs descripteurs de textures [Ojala *et al.*, 1996, Sharma et Singh, 2001, Stachowiak *et al.*, 2005]. Chaque méthode d'analyse permet de décrire différents aspects dans une texture et des combinaisons pertinentes peuvent aboutir à des descriptions très précises. De ce point de vue, la majeure partie des méthodes d'analyse de textures n'ont pas d'applicabilité générale et dans certains cas ne sont pas capables d'identifier certaines classes de textures. En comparaison, le système visuel humain est efficace sur tous types de textures même en l'absence d'un contexte favorable. De façon générale, les êtres humains utilisent des propriétés texturales comme la directionnalité, la rugosité, le contraste, la granularité, etc. Afin d'avoir des mesures générales sur les textures, certains auteurs [Tamura *et al.*, 1978, Amadasun et King, 1989] ont présenté des recherches dans cette direction en proposant des descripteurs de textures 2D compréhensibles par des êtres humains.

Parmi les méthodes statistiques les plus utilisées, nous retrouvons les matrices de cooccurrence des niveaux de gris d'Haralick (GLCM). Celles-ci sont facilement extensibles à la 3D cependant leur complexité de calcul n'est pas négligeable. Certaines des caractéristiques d'Haralick extraites à partir de ces matrices sont compréhensibles et quantifiables par un être humain (contraste). Pour d'autres, il est plus difficile de produire une interprétation visuelle (somme des entropies).

Les méthodes d'analyse de motifs s'attachent à déterminer et étudier les structures et les motifs composant une texture. Les méthodes comme les graphes de Voronoï permettent d'analyser géométriquement une texture en construisant un maillage à partir des informations de niveaux de gris. Celles-ci sont basées sur la construction de graphes, ce qui rend possible leur extension en 3D. D'autres méthodes, également utilisables sur des images 3D, proposent d'identifier les textures d'une texture à partir de bancs de filtres. La méthode LBP est une technique efficace avec une complexité très faible. Cependant, l'extension 3D de celle-ci n'est pas forcément aisée car dans le domaine 3D le nombre de motifs qu'il est possible de générer devient très important. Pour caractériser une texture, cette méthode recense dans un histogramme 1D l'ensemble des motifs binaires identifiés à l'intérieur de celle-ci. Il n'y a donc aucune correspondance entre cette façon de décrire une texture et celle utilisée par les êtres humains. Pour la plupart, ces méthodes diffèrent des mécanismes de qualification utilisés par les êtres humains. Il faut noter cependant quelques exceptions [Mital *et al.*, 1994, Goyal *et al.*, 1994a, Goyal *et al.*, 1994b, Goyal *et al.*, 1995, Shoshany, 2008] où les auteurs proposent de calculer des informations de compacité et de forme sur des motifs binaires. Ces descripteurs de textures se révèlent tout à fait exploitables et compréhensibles pour un utilisateur non expert en analyse d'image.

Parmi les méthodes basées sur un modèle, nous avons présenté les champs de Markov ainsi que les fractales. Les champs de Markov permettent de décrire une texture à partir du voisinage décrit par un ensemble de cliques selon un type de connexité choisi. En considérant un voisinage 3D, ces méthodes sont donc naturellement extensibles au domaine 3D. Il faut cependant noter que ces méthodes sont gourmandes en temps de calcul. Enfin, il est difficile de mettre en correspondance les paramètres d'un modèle Markovien, utilisés pour caractériser les textures, avec la description humaine des textures. La dimension fractale permet d'avoir une bonne estimation de la rugosité d'une texture. Celle-ci est une caractéristique facilement quantifiable par un être humain. De plus, certaines méthodes d'estimation de la dimension fractale, comme la méthode du *Box counting* sont facilement applicables aux textures solides. Avec les fractales, il est possible de calculer d'autres types de caractéristiques pour améliorer leur performance. Cependant, utilisés seuls, ces descripteurs restent peu efficaces.

Parmi les méthodes empruntées au traitement du signal, les plus populaires pour le traitement des textures sont les ondelettes discrètes et les filtres de Gabor. L'une comme l'autre, ces méthodes sont extensibles à la 3D. Si la décomposition en ondelette 3D d'une texture solide est rapide, il n'en est rien concernant son analyse avec un filtre de Gabor 3D. Ces derniers permettent une analyse très précise mais le nombre important de paramètres peut entraîner des calculs lourds. Les caractéristiques de ce type de méthodes sont calculées directement sur les réponses de filtres générés. Même si ces caractéristiques sont de type statistique, il est bien souvent impossible d'en déduire une correspondance avec le processus de description humain.

## 1.7. CONCLUSION

---

Dans la suite, nous nous attacherons à proposer de nouveaux descripteurs de textures permettant de caractériser de façon robuste et générique les textures volumétriques. Cet ensemble d'attributs sera le résultat de la combinaison de méthodes d'extraction de différents types (fréquentielle, géométrique, statistique). De plus, les caractéristiques choisies devront correspondre à des concepts facilement exploitables et manipulables par des utilisateurs non experts en vision par ordinateur. Le prochain chapitre présente les différents attributs de textures choisis ainsi qu'une description détaillée des méthodes développées pour les obtenir.