Chapitre 4

Les textures

Chapitre rédigé par Henri MAÎTRE

4.1 Qu'est ce qu'une texture ?

Dans le domaine du traitement de l'image et de la vision, il n'existe pas de définition satisfaisante de la texture. Les définitions mathématiques construites à partir de propriétés statistiques sont soit trop générales et imprécises soit trop restrictives pour s'adapter à la diversité des cas rencontrés. La définition que nous proposons n'est pas opérationnelle et prête sur ce point à critique. Elle s'appuie sur une constatation expérimentale : une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène, c'est-à-dire formant un tout pour un observateur. C'est cette propriété de cohérence de la texture placée dans son contexte d'être perçue comme un tout homogène par l'œil humain qui sera recherchée le plus souvent par le traiteur des images, dans le but d'isoler les textures, soit pour segmenter l'image, soit pour reconnaître des régions. La figure 4.1 illustre la diversité des textures et l'ouvrage de référence de Brodatz [Brodatz, 1966] offre une collection de textures naturelles qui constituent d'excellents exemples et sont souvent utilisées pour tester les algorithmes et les méthodes.

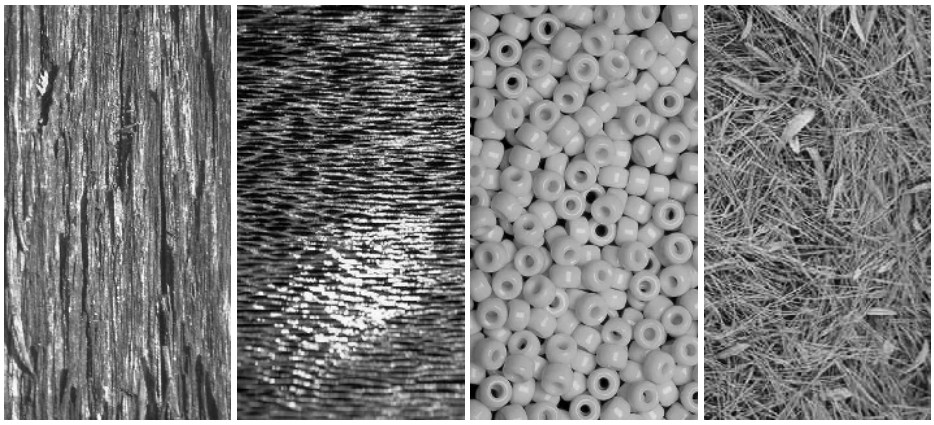

FIG. 4.1 – Quelques exemples de textures naturelles : écorce d'arbre, poil court, perles, tapis d'aiguilles de pin et de feuilles.

Les capacités du système visuel humain à discerner les textures différentes sont remarquables et très mal ex-

pliquées. Nos capacités à mémoriser et discriminer des textures sont très variables. Elles deviennent excellentes dès lors que l'on a reconnu l'objet de la vie courante dont la texture est issue. Notre mémorisation est alors remarquable et invariante à de très nombreuses transformations (changement d'échelle, d'orientation, d'éclairage, de couleur, etc.). Au contraire, des textures non reconnues (champs aléatoires créés mathématiquement par exemple, ou textures placées hors de leur contexte) sont mémorisées de façon très fugitive et sont peu robustes aux transformations. Nos capacités à distinguer des textures différentes peuvent parfois s'expliquer par les capacités de filtrage du système visuel (bande passante des voies optiques, discrimination angulaire, sensibilité aux variations locales de luminosité [Hubel et Wiesel, 1969, Julesz, 1971]). Parfois elles relèvent de mécanismes de psycho-vision et peuvent trouver des explications par exemple dans la Gestalt Theorie qui propose des mécanismes d'association (groupement perceptuel) ou de discrimination, à partir des symétries, proximités, similarités des stimulus élémentaires composant la texture [Wertheimer, 1944, Grossberg et Mingolla, 1985].

L'une des remarquables capacités du système visuel est de reconnaître des textures déjà mémorisées même sous de très fortes distorsions géométriques (vues perspectives, recouvrement de surfaces gauches, etc.). Dans ces conditions, et même en l'absence d'un contexte favorable (par exemple présence de silhouettes d'objet ou d'alignements perspectifs), 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, celle-ci restant globalement invariante dans un repère imprécisé (cf. figure 4.2). Ainsi un tronc d'arbre est vu comme une texture homogène enroulée sur un cylindre et non comme une texture plane dont les fréquences spatiales s'accroissent en approchant des bords.

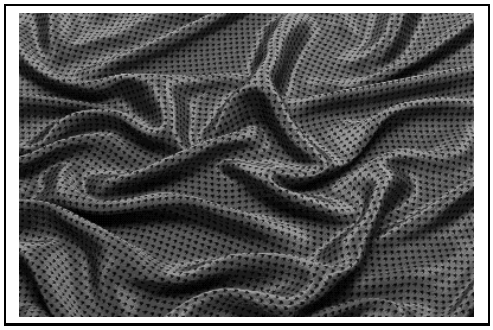

FIG. 4.2 – Lorsque la texture n'est pas homogène, on interprète naturellement les variations locales de ses propriétés statistiques comme des déformations de la surface sur laquelle elle est projetée.

Pour toutes ces raisons, (mais aussi pour simplifier considérablement la tâche du traiteur d'image), nous considérerons par la suite les textures comme spatialement invariantes, c'est-à-dire que nous négligerons généralement les effets de perspectives et les variations d'homogénéité du champ de distribution des intensités. Nous lèverons cette hypothèse au paragraphe 4.6.

4.1.1 Distribution aléatoire ou régulière ?

Une première constatation que l'on fait en examinant des textures naturelles est le rôle particulier que joue l'aléatoire dans la texture. On distingue assez naturellement deux modèles extrêmes de textures, entre lesquels se positionnent un peu toutes les textures :

1. les textures régulières, dans lesquelles la périodicité du motif est évidente : grilles, murs, tissus, etc.
2. les textures aléatoires pour lesquelles la distribution des intensités n'est l'objet d'aucune régularité apparente : sable, nuages, herbe, foule.

La première famille sera bien décrite par des approches fréquentielles ou des approches structurelles dans lesquelles on associera un motif et des règles de placement sur un pavage régulier.

La seconde approche sera plutôt décrite par des lois statistiques, des moments, une description spectrale en termes de densité de puissance, des propriétés de corrélation ou d'isotropie.

Mais une texture n'est jamais strictement périodique ni totalement aléatoire et les deux modèles se complètent naturellement. L'une des difficultés est de les mêler dans des modèles capables de s'adapter à la variabilité des textures étudiées.

4.1.2 Les échelles des textures

Une autre particularité très importante des textures est qu'elles présentent généralement plusieurs niveaux d'échelle auxquels on peut les étudier.

À petite échelle (donc pour des détails fins), on observe souvent un objet élémentaire qui constitue la base de la texture (poil, perle, aiguille de pin, sur la figure 4.1). Cette primitive peut être très régulière géométriquement (perle), ou photométriquement (aiguille de pin). Elle peut être au contraire relativement variable (personne dans une foule, caillou sur une plage, nuage). Elle possède une taille et des propriétés statistiques caractéristiques (qui joueront sur sa fonction de corrélation, son isotropie, ...). Cette composante élémentaire de la texture a tendance naturellement à disparaître lorsqu'on observe le champ global de la texture. Elle est intégrée à la texture.

À plus grande échelle (donc pour une analyse plus grossière), la texture apparaît comme une juxtaposition plus ou moins régulière des motifs ci-dessus. Cette juxtaposition possède ses propres lois d'isotropie, de périodicité, de régularité et induit donc d'autres lois statistiques (corrélation, densité de puissance) qui se combinent à celles du motif de base.

Une bonne analyse de texture donnera des informations sur ces deux composantes également.

4.1.3 Analyse ou synthèse

Un effort très important a été fait dans le domaine de la synthèse des images pour créer artificiellement ou copier des textures naturelles : forêts ou cultures pour des simulateurs de vols, bois, tissus, métaux pour des représentations d'intérieur, murs, toitures, eau, pour des jeux vidéo, etc. Ces travaux se distinguent notablement de ceux conduits en traitement d'image qui visent plutôt à extraire des paramètres discriminants et robustes permettant de séparer des textures différentes. Mais ces travaux se rejoignent en ce qu'ils contribuent simultanément à une meilleure connaissance et compréhension des textures à travers la boucle « analyse/synthèse » qui a fait ses preuves en reconnaissance des formes. Dans la suite, nous empruntons les méthodes aussi bien à l'analyse qu'à la synthèse des images en précisant lorsque c'est nécessaire les limites des méthodes dans l'une ou l'autre application.

4.2 Modèles de texture

4.2.1 Un modèle biologiquement plausible

Compte tenu du rôle important de la perception humaine dans la définition même de texture, des modèles de textures on été proposés, s'inspirant de ce que l'on connaît aujourd'hui de cette perception. Ces travaux s'appuient sur les mécanismes d'adaptation du système visuel aux fréquences spatiales et aux orientations [Hubel et Wiesel, 1969, de Valois et al., 1982] ainsi que sur des expérimentations sur les effets de masquage psychovisuel [Phillips et Wilson, 1984].

On trouve par exemple dans [Bergen et Landy, 1991] un tel modèle. On décompose tout d'abord le flot optique en une pyramide (cf. paragraphe 3.2.4) par une succession de filtrages gaussiens et de sous-échantillonnages adaptés (semblables aux pyramides de Burt et Adelson [Burt, 1981]). Chaque niveau de la pyramide est alors filtré par 4 filtres qui produisent une dérivée seconde directionnelle selon l'horizontale, la verticale et les deux diagonales. On procède ensuite à une mesure d'énergie dans chaque image par une intégration locale des sorties

précédentes et l'on soustrait ces énergies, dans un même niveau de résolution, entre orientations voisines (pour exprimer le contraste relatif entre directions). On procède enfin à une étape de normalisation pour tenir compte d'une sorte de gain variable de contrôle de contraste vérifié expérimentalement.

Un tel schéma de fonctionnement du système visuel humain est très schématisé et controversé. Il existe d'autres schémas tout aussi intéressants [Mayhew et Frisby, 1978], mais il a le mérite de comporter la plupart des étapes que nous reverrons plus loin pour analyser les textures.

4.2.2 Modèles stochastiques

Ces modèles, au contraire, mettent l'accent sur la distribution statistique des pixels et sur leur dépendance spatiale. La texture est alors considérée comme la réalisation d'un processus aléatoire gouverné par ses lois. Nous avons vu dans les chapitres antérieurs deux modèles qui ont été abondamment utilisés pour modéliser les textures :

- le schéma booléen (cf. section **??**), et plus généralement les ensembles fermés aléatoires, qui considèrent la texture comme une distribution poissonnienne de motifs et qui est particulièrement bien analysé par les outils développés en morphologie mathématique qui permettent de définir sa capacité de Choquet [Schmitt et Mattioli, 1994] ;
- les modèles autorégressifs et leurs dérivés, chaînes de Markov et champs de Markov, ces derniers étant aujourd'hui les plus universellement adoptés. Nous nous attarderons particulièrement sur ce point à la section 4.4 [Cross et Jain, 1983, Chellappa et Kashyap, 1985].

4.3 Analyse et reconnaissance de textures

Il existe de nombreuses revues des diverses approches de l'analyse de textures, et nous renvoyons le lecteur à trois textes qui couvrent assez bien l'évolution du domaine : [VanGool et al., 1985, Reed et du Buf, 1993, Randen et Husøy, 1999]

4.3.1 Schéma général

Le principe le plus général de l'analyse statistique des textures est le suivant.

1. On définit un voisinage $\mathcal{V}_{ij}$ de tout pixel (i, j), de taille et de forme appropriées.
2. Sur le voisinage $\mathcal{V}_{ij}$ on mesure des propriétés particulières de l'image : soit π_k la mesure attachée à la configuration k parmi les N mesurées. Ces mesures seront attachées au pixel (i, j).
3. On classifie les pixels à partir du vecteur formé par les π_k par l'une des nombreuses méthodes de la reconnaissance des formes (rappelées en section 3.1.3).
4. Éventuellement, avant cette classification, et si le vecteur de mesures est trop grand, on réduit la dimension de l'espace par une sélection judicieuse des composantes les plus significatives ou par analyse en composantes principales.

La fenêtre d'analyse

La dimension du voisinage $\mathcal{V}_{ij}$ est importante. C'est l'un des choix délicats de l'analyse des textures (cf. figure 4.3). Elle doit comprendre au moins un motif de base de la texture mesurée pour fournir des statistiques homogènes sur la texture, mais, si $\mathcal{V}_{ij}$ est trop grand, la précision de localisation des frontières des textures détectées sera médiocre. La forme est généralement carrée. Taille et forme peuvent éventuellement s'adapter au signal à mesurer si l'on a quelque méthode pour le faire (pré-segmentation, masque de contours, etc.), mais il faut s'assurer alors que la technique de classification n'est pas perturbée par des mesures faites sur des échantillons variables.

La classification des données

Ce sont celles que nous avons vues au chapitre 3.1.3 : nuées dynamiques, plans séparateurs, réseaux neuro-mimétiques, C moyennes floues, etc. Dans l'idéal on aimerait pouvoir obtenir à partir de l'image $f(i,j)$, une nouvelle image $g(i,j)$ dont les plages seront très distinctes et pourront, idéalement, être segmentées par simple seuillage (cas de la mesure d'un seul paramètre discriminateur) ou par des seuils multiples sur les diverses composantes. En pratique il est rare que l'on puisse séparer les nuages des mesures et il faut trouver des compromis lors des décisions. Les erreurs de classification peuvent bénéficier de techniques contextuelles de reclassification comme la relaxation ou les champs de Markov comme nous le verrons plus bas (section 4.4).

Lorsque l'on cherche non seulement à discriminer des plages de textures différentes, mais aussi à reconnaître des textures particulières, on utilise des techniques de classification superviséees : k-plus proches voisins, classifieurs bayésiens, réseaux neuro-mimétiques, algorithmes génétiques, préalablements entraînés sur des ensembles d'apprentissage [Vickers et Modestino, 1982].

La réduction des dimensions

Les diverses variables mesurées sont rarement indépendantes et le vecteur des π_k occupe rarement toutes les dimensions de N. On calcule alors la matrice Π de taille $N \times N$ et telle que $\Pi(k,l) =< \pi_k \pi_l >$. On diagonalise Π et on ne retient que ses p plus grands vecteurs propres sur lesquels on projette toutes les mesures. La classification se fait alors dans l'espace de dimension p, ce qui réduit les calculs et permet de prendre de meilleures décisions car les composantes sont alors non corrélées.

4.3.2 Approches par mesures statistiques

Dans ces approches on mesure les propriétés statistiques autour du point (i,j) dans le voisinage $\mathcal{V}_{ij}$.

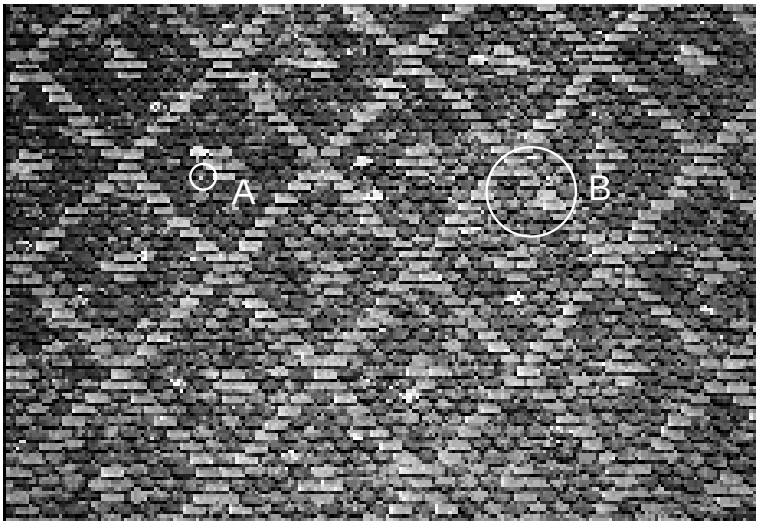

FIG. 4.3 – Le choix de la taille de la fenêtre d'analyse est important pour sélectionner les détails que l'on souhaite préserver. La fenêtre A permettra de déterminer des propriétés caractéristiques d'une tuile de ce toit et distinguera tuiles claires de tuiles sombres, la fenêtre B mesurera des propriétés moyennes sur tout le toit et ignorera les différences entre tuiles claires et sombres.

Statistiques du premier ordre

Ce sont celles qui ne prennent en compte qu'un pixel à la fois[1]. Elles se déduisent de la probabilité empirique $p(n)$ du niveau de gris n dans la fenêtre ou de l'histogramme $h(n) \simeq \nu p(n)$ où ν est le nombre total de pixels dans l'image. Ce sont :

- les moments d'ordre k (non centrés) : $\mu_k = \Sigma_n n^k p(n)$,
- les moments centrés d'ordre k : $\tilde{\mu_k} = \Sigma_n (n - \mu_1)^k p(n)$, et en particulier :
  1. la moyenne : μ_1,
  2. la variance : $\sigma^2 = \tilde{\mu_2}$,
  3. le biais : $\gamma_1 = \frac{\tilde{\mu_3}}{\sigma^3}$,
  4. l'aplatissement (ou kurtosis) : $\gamma_2 = \frac{\tilde{\mu_4}}{\sigma^4} - 3$.
- l'énergie : $W = \Sigma_n |p(n)|^2$,
- l'entropie : $E = -\Sigma_n p(n) \log p(n)$,
- le contraste : $C = \frac{\max(n) - \min(n)}{\max(n) + \min(n)}$,
- la dynamique : $D = \max(n) - \min(n)$,
- le coefficient de variation (surtout pour les images cohérentes dont le bruit est multiplicatif : radar, images ultra-sonores) : $v = \frac{\mu}{\sigma}$,
- l'exposant de Holder, caractéristique de la dimension fractale : [Pentland, 1984].

Si les fenêtres de mesure sont petites, les statistiques mesurées sont souvent peu significatives (pour une fenêtre de 30×30 pixels et un histogramme de 256 niveaux de gris, l'occurrence moyenne des pixels est inférieure à 4 par niveau de gris). On choisit alors souvent de réduire la dynamique des images. On a montré que les propriétés de discrimination des textures se conservaient remarquablement même pour des quantifications très fortes des textures (jusqu'à 8 ou 4 niveaux de gris seulement), à la condition d'adapter la quantification. Une quantification brutale se fait en ne conservant que les bits de poids fort. Une quantification adaptée choisit de partager les niveaux de gris en fonction des statistiques de la texture :

- par moyenne et écart type par exemple (cela donne 4 classes non équiréparties),
- par médiane et distance interquartile[2], ce qui donne des classes équiréparties.

Les statistiques de points particuliers

On recherche dans ce cas, dans la fenêtre $\mathcal{V}_{ij}$, la densité moyenne de certains points d'intérêt. Les plus souvent utilisés sont :

- les maximums locaux de l'intensité,
- les points de contour (après application d'un détecteur et seuillage),

Les statistiques d'ordre élevé

Ce sont surtout les statistiques d'ordre 2 qui sont exploitées, c'est-à-dire celles qui mettent en jeu deux pixels simultanément. En effet, B. Julesz a émis une conjecture, fondée sur une très vaste étude expérimentale, que le système perceptif humain ne distingue pas les textures qui ont des statistiques similaires aux ordres 1 et 2, même si elles diffèrent aux ordres supérieurs [Julesz, 1971]. Cette conjecture a été démontrée fausse par A. Gagalowicz qui a produit des textures artificielles ne différant qu'à l'ordre 3 [Gagalowics et Tournier-Lasserve, 1986], mais perceptivement différentes (cf. figure 4.4). La conjecture de Julesz demeure cependant un guide important pour les études sur les textures car elle est assez bien vérifiée pour les textures naturelles.

[1] Attention, on désigne également parfois par premier ordre les statistiques comme la moyenne, qui ne mettent en jeu que des moments du premier ordre.

[2] La distance interquartile est la distance qui, dans l'histogramme, sépare les niveaux de gris des 25% de pixels les plus sombres de ceux des 25% de pixels les plus clairs. C'est une mesure robuste de l'étalement d'une loi.

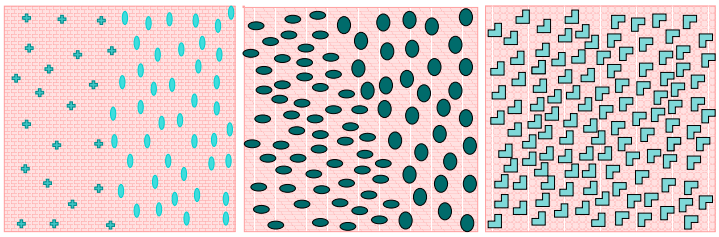

FIG. 4.4 – A gauche, les 2 textures diffèrent dès le premier ordre. Au centre, les 2 textures sont identiques au premier ordre mais diffèrent au second. A droite les 2 textures sont identiques à l'ordre 1 et à l'ordre 2, mais différentes à l'ordre 3. Ces deux textures sont distinguables visuellement, en contradiction avec la conjecture de Julesz.

Les statistiques d'ordre 2 sont bien appréhendées par la **fonction d'autocorrélation**[3] de la texture, et plus particulièrement par la fonction d'autocorrélation normée centrée :

$$C_f(k,l;i,j) = \frac{\Sigma_{\mathcal{V}_{ij}}(f(i,j)-\bar{f}).(f(i+k,j+l)-\bar{f})}{\Sigma_{\mathcal{V}_{ij}}(f(i,j)-\bar{f})^2}$$

On déduit de nombreuses propriétés de cette fonction :

1. dans une direction donnée, et au voisinage de l'origine, elle peut être fréquemment approchée par une fonction régulière, par exemple :
$$C_f(k,l;i,j) \sim e^{-\alpha(i,j)|k|}e^{-\beta(i,j)|l|}$$
et les paramètres α et β sont très significatifs de la dépendance spatiale (en i et j) des niveaux de gris à l'intérieur de la texture ;
2. les deux directions d'inertie maximale $\lambda_0(i,j)$ et minimale $\lambda_1(i,j)$ extraites de $C_f(k,l;i,j)$ expriment bien les directions privilégiées de la texture et, si le rapport $\frac{\lambda_0}{\lambda_1}$ est fort, la texture est fortement anisotrope.
3. si la fonction $C_f(k,l;i,j)$ possède des maximums locaux différents de $k=0, l=0$, alors la texture est fortement périodique et les périodes de la fonction de corrélation permettent d'estimer de façon robuste la période de la texture.
4. enfin par transformation de Fourier, la fonction d'autocorrélation donne accès au spectre de densité de puissance (théorème de Wiener-Kinchine, cf. section **??**), et à tous les traitements spectraux que nous verrons ci-dessous.

Plus encore que la fonction d'autocorrélation, les **matrices de cooccurrence** sont des outils adaptés à l'analyse des textures. Ces matrices sont définies pour un vecteur de translation V_i donné comme la probabilité jointe de l'amplitude d'un point M et d'un point $N = M + V_i$. Elles sont obtenues pratiquement par une estimation sur la texture, c'est-à-dire un décompte des occurrences des niveaux m en M et n en N. C'est donc un tableau généralement de taille 256×256 si la texture à 256 niveaux de gris :

$$\Gamma(m,n;V_i) \;=\; \text{proba}\,(f(M)=m,\; f(N=M+V_i)=n))$$

En raison de sa taille même, la matrice de cooccurrence est sous cette forme un outil ni très pratique, ni très fiable (les valeurs $\Gamma(m,n;V_i)$ sont des estimations très médiocres des probabilités réelles). On a donc intérêt à comprimer ces matrices. Cela se fait de plusieurs façons :

– en diminuant le nombre de niveaux de gris par une quantification régulière ou adaptative, par exemple 8 niveaux de gris déterminés de façon adaptée (par moyenne et écart-type ou par médiane et distance interquartile) donnent souvent une bonne qualité de représentation ;

[3] certains auteurs préfèrent étudier le **variogramme** qui s'exprime comme $1 - C_f(k,l)$.

FIG. 4.5 – Construction d'une matrice de cooccurrence : on définit un vecteur V_i, puis on mesure l'occurrence des couples de niveaux de gris m en M et n en N. La matrice de coocurrence ainsi calculée est $\Gamma(m, n, V_i)$.

- en tirant profit des symétries existant dans la texture pour regrouper des vecteurs V_i contribuant de façon identique à l'apparence de la texture (par exemple des vecteurs symétriques le long des axes, ou des vecteurs de même module si la texture est isotrope) ;
- en ne conservant que les seuls vecteurs significatifs ;
- enfin, on a pu montrer que quelques descripteurs des matrices de cooccurrences pouvaient très bien les représenter : position du centre de gravité, rapport des inerties, énergie, entropie, facteurs de symétrie, etc. [Chen et Pavlidis, 1979].

Les configurations particulières de pixels

On s'intéresse dans cette méthode à mesurer la probabilité d'apparition de certaines configurations particulières de pixels (cf. figure 4.6). Ce sont donc des techniques qui s'appliquent surtout à des textures binaires ou à très peu de niveaux de gris (ou rendues telles). Les mesures ainsi faites permettent de construire un vecteur d'attributs qui permettra de discriminer des textures différentes. Des configurations bien choisies pour un problème particulier peuvent être très discriminantes. Mais il est souvent difficile d'éviter que des mesures ne soient statistiquement dépendantes car les probabilités de certaines fenêtres sont évidemment dépendantes. Il convient donc souvent de réduire la dimension de l'espace des paramètres par analyse en composantes principales.

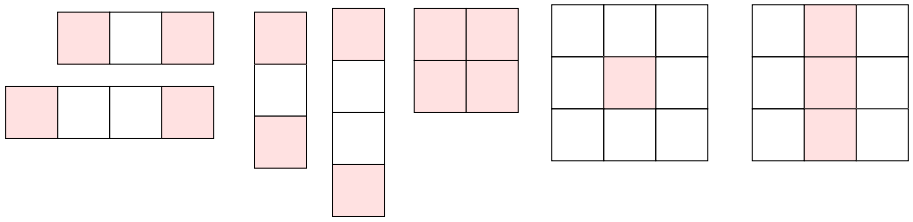

FIG. 4.6 – Quelques configurations particulières de pixels qui peuvent être mesurées pour fournir un descripteur statistique : par exemple on mesure le nombre d'occurrences de trois pixels alignés horizontalement, les 2 des bords étant noirs et celui du centre blanc (en haut à gauche).

4.3.3 La détection d'énergie dans des canaux fréquentiels

L'idée sous-jacente à toutes ces techniques est d'extraire l'énergie portée par le signal dans des bandes de fréquences diverses. Il n'est alors pas nécessaire d'utiliser de fenêtre d'estimation, car les analyses fréquentielles utilisées disposent généralement de leurs propres fonctions de fenêtrage, souvent gaussiennes. Il existe donc deux familles de techniques :

- celles qui analysent systématiquement tout l'espace fréquentiel (par exemple par une décomposition en ondelettes),
- celles qui sélectionnent un petit nombre de domaines fréquentiels significatifs (c'est par exemple l'analyse par filtres de Gabor).

Les filtres de Law

Ce sont des filtres très simples qui sont appliqués dans l'espace de l'image par des masques, dans l'esprit des filtres de détection de contour par masquage adapté. Les filtres sont au nombre de 25, obtenus par produit en x et y de 5 filtres de base (cf. figure 4.7). Leurs performances sont modestes, mais ils sont très rapides.

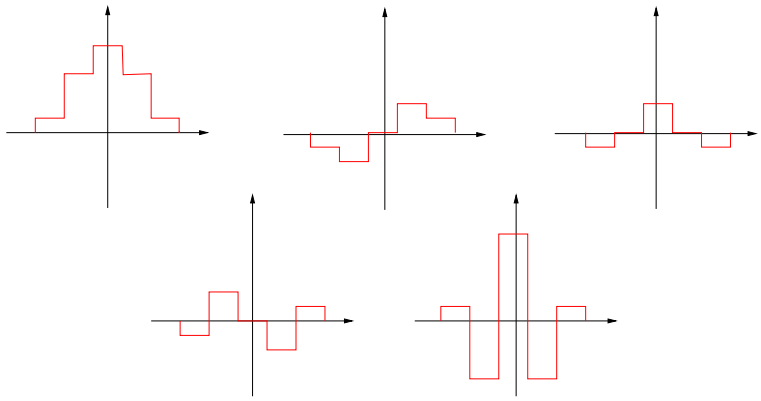

FIG. 4.7 – Les 5 filtres de base de Law. Un filtre de détection est le produit séparable de deux de ces filtres. Il est appliqué dans l'espace de l'image.

Les filtres en anneau et en coin

Ce sont des filtres définis dans l'espace de Fourier à 2 dimensions [Coggins et Jains, 1985]. Ils sont constitués du produit de deux familles élémentaires de filtres :

- 7 filtres en anneau, échelonnés de façon dyadique (c'est-à-dire centrés autour de fréquences suivant une progression géométrique de fréquences, le premier de fréquence $\nu_1 = \nu_0$, le second de fréquence $\nu_2 = 2\nu_0$, le troisième de fréquence $\nu_3 = 4\nu_0$, etc.), et de profil gaussien (cf. figure 4.8) :

$$\Phi_i(\nu = \sqrt{u^2 + v^2}) = \exp\left[-\frac{(\nu - \nu_i)^2}{2\alpha\nu_i^2}\right]$$

  le recouvrement de ces filtres est contrôlé par le paramètre α.
- 4 filtres directionnels selon les axes et les diagonales, également de profils gaussiens.

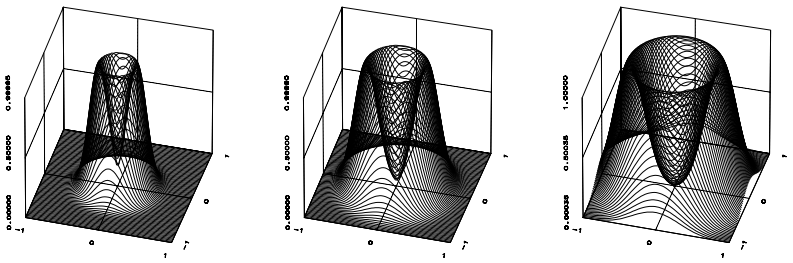

FIG. 4.8 – Les filtres en anneau dans le domaine spectral centrés sur les fréquences : 0,25, 0,40 et 0,60. Ce ne sont donc pas des filtres dyadiques.

Les filtres dyadiques de Gabor

Ce sont des filtres qui essaient de copier les fonctions de la vision des mammifères. Ils sont assez proches des précédents. Ce sont aussi des filtres à couverture dyadique qui se déduisent par rotation et homothétie du filtre de base qui a pour réponse impulsionnelle [Jain et Farrokhnia, 1991] :

$$\phi(x,y) = exp\left[-1/2\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right]\cos(2\pi\nu_0 x)$$

Dans [Jain et Farrokhnia, 1991], on conseille l'emploi de 5 fréquences radiales différentes et de 4 orientations. Dans [Randen et Husøy, 1999], on utilise aussi 4 orientations et les fréquences radiales suivantes :

$$\frac{\sqrt{2}}{2^6} \quad \frac{\sqrt{2}}{2^5} \quad \frac{\sqrt{2}}{2^4} \quad \frac{\sqrt{2}}{2^3} \quad \frac{\sqrt{2}}{2^2}$$

Les ondelettes et les QMF

La transformation en ondelettes discrètes permet de choisir une famille complète de filtres de décomposition en sous-bandes. Ces méthodes ont été présentées au chapitre **??**. Les ondelettes de Mallat [Mallat, 1989] utilisent elles aussi une décomposition dyadique (mais deux directions d'analyse seulement). Des travaux suggèrent qu'une décomposition plus serrée que l'octave (c'est-à-dire des décompositions non-dyadiques) serait plus efficace pour l'analyse des textures [Chang et Kuo, 1993]. On propose ainsi l'usage de paquets d'ondelettes [Laine et Fan, 1993, Saito et Coifman, 1995]. Les ondelettes de Daubechies demeurent l'une des bases d'ondelettes les plus utilisées pour l'analyse de textures en raison de leur bonne efficacité de calcul, de leurs bonnes performances pour séparer les fréquences ainsi que de la grande variété des fonctions qu'elles autorisent [Daubechies, 1992, Unser, 1995].

Des modèles ont également été créés alliant les propriétés de sélection des ondelettes et les modèles statistiques [Portilla et Simoncelli, 2000].

4.3.4 Les filtres optimisés

Lorsque l'on cherche à distinguer 2 textures, ou un petit nombre de textures connues par avance, la recherche systématique de l'énergie dans de nombreuses décompositions en sous-bandes peut être très lourde et hasardeuse. Il est possible de mettre en place des techniques adaptées pour les seules textures que l'on recherche qui permettent de se concentrer sur quelques mesures seulement de l'espace fréquentiel.

Les filtres de Gabor, qui disposent d'un petit nombre de paramètres seulement se prêtent assez bien à ce type

d'optimisation. On cherche alors à optimiser un critère comme :

$$J_1 = \frac{\mu_{t_1}}{\mu_{t_2}} \qquad J_2 = \frac{(\mu_{t_1} - \mu_{t_2})^2}{\mu_{t_1}\mu_{t_2}} \qquad \text{ou} \quad J_3 = \frac{(\mu_{t_1} - \mu_{t_2})^2}{\sigma_{t_1}^2 + \sigma_{t_2}^2}$$

où μ_{t_i} est la valeur moyenne de la mesure donnée par la texture t_i et σ_{t_i} sa variance [Unser, 1986]. Il est parfois possible de déterminer des valeurs optimales analytiques des filtres par ces critères, dans les autres cas on optimise par essai-erreur.

4.3.5 Les modélisations autorégressives

De la même façon que les signaux monodimensionnels sont analysés par des modélisations autorégressives uni-dimensionnelles (AR ou ARMA) afin de mettre en évidence les périodicités qu'ils contiennent, les signaux bidimensionnels peuvent être analysés par des modèles autorégressifs bi-dimensionnels. Un étude très complète de ces méthodes de description est présentée dans [Garello, 2001].

Dans une approche autorégressive à moyenne ajustée (ARMA), l'image est décrite par la formule :

$$f(i,j) = \sum_{(k,l)\in\mathcal{D}_s} a_{k,l} f(i-k, j-l) + b_{0,0} B(i,j) + \sum_{(k,l)\in\mathcal{D}_e} b_{k,l} B(i-k, j-l) \tag{4.1}$$

où $\mathcal{D}_s$ représente le domaine de prédiction lié à la sortie du filtre, $\mathcal{D}_e$, celui lié à l'entrée du filtre et B décrit un processus de bruit. Si les coefficients $b_{k,l}$ sont nuls pour tout (k,l) appartenant à $\mathcal{D}_e$, alors le processus est AR. Si les coefficients $a_{k,l}$ sont nuls pour tout (k,l) appartenant à $\mathcal{D}_s$, alors le processus est MA.

Le choix des domaines $\mathcal{D}_s$ et $\mathcal{D}_e$ incombe à l'utilisateur. Dans de nombreux cas on choisit $\mathcal{D}_s = \mathcal{D}_e = \mathcal{D}$. Afin de conserver aux modèles ARMA une bonne localité, il est souhaitable que le nombre de termes dans $\mathcal{D}$ soit faible, mais pour bien représenter des signaux à longue périodicité, il vaut mieux qu'il soit grand. Ces modèles expriment naturellement une dépendance causale et suscitent donc des discussions comme celles que nous avons abordées au chapitre **??**.

Les modèles ARMA permettent de représenter très bien n'importe quelle densité spectrale de puissance. Ceci se fait en utilisant les équations normales associées de Yule Walker [Alata et Cariou, 2001]. Les représentations par modélisation autorégressive des textures est particulièrement efficace pour des textures très périodiques et de large extension.

4.4 Les approches par champs markoviens

Les champs de Markov (cf. chapitre **??**) peuvent être vus comme une sous-classe des processus ARMA, mais leur adaptation à traiter des images les distinguent. Ils se prêtent doublement à l'analyse des textures :

1. tout d'abord parce qu'ils comportent naturellement des descriptions des dépendances spatiales entre pixels par le choix des cliques et des potentiels d'interaction au sein des cliques,
2. mais aussi parce qu'ils ont, dans le terme d'énergie lié aux connaissances à priori, les éléments qui permettent de décrire les régions et leurs interactions.

Les champs de Markov sont particulièrement adaptés pour synthétiser et modéliser des textures, puisqu'il suffit de se donner les potentiels correspondant aux propriétés de dépendance, d'isotropie, de distance aux centres de classe, etc. que l'on souhaite voir représenter. Il y a donc une abondante littérature qui exploite le formalisme des champs markoviens pour segmenter des textures : [Hu et Fahmy, 1992, Derin et Elliott, 1987] [Won et Derin, 1992, Andrey et Tarroux, 1998]

très faciles d'emploi en synthèse des textures, ils sont cependant d'un usage plus difficile en analyse car il n'existe pas de formulation explicite permettant de déterminer les potentiels correspondant à une texture dont on dispose. Ce problème a été abordé au chapitre **??**, à la section **??**.

En règle générale, pour retrouver un champ de Markov à partir de la réalisation d'une texture, il faut définir les cliques significatives, puis les potentiels qui lient les sites dans ces cliques [Azencott et al., 1997].

1. La détermination des cliques utiles : on peut l'obtenir soit par l'analyse des fonctions de corrélation, soit par des techniques de croissance de voisinage, accompagnées de tests d'indépendance au fur et à mesure que l'on agrandit le voisinage.
2. La détermination des potentiels : elle passe souvent par la réduction des potentiels possibles à des classes étroites (par exemple les potentiels quadratiques ou les potentiels polynomiaux), elle peut alors se faire par des techniques de moindres carrés ou par des méthodes de filtrage (techniques de renormalisation).

4.4.1 La méthode de Manjunath et Chellappa

C'est une technique assez typique des approches markoviennes. Elle repose sur l'hypothèse de probabilités gaussiennes dans chaque classe (donc de potentiels quadratiques) [Manjunath et Chellapa, 1991]. Les énergies d'attache aux données pour une classe $L_s = l$ sont de la forme :

$$U_1(y_s|y_r, l) = \frac{1}{2\sigma_l^2}\left[y_s^2 - 2\sum_r \theta_t^l y_s y_r\right] \tag{4.2}$$

dont les inconnues sont, pour chaque texture l, les θ_i, μ et σ^2, où les θ_i correspondent aux cliques des directions S, SE, E, etc. du point s (cf. figure 4.9).

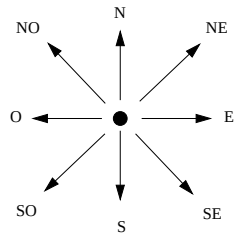

FIG. 4.9 – Le 8-voisinage utilisé par Manjunath et Chelappa. Seules les cliques d'ordre 2 sont utilisées.

Dans une première étape, il faut apprendre les paramètres des textures. L'image est arbitrairement subdivisée en petites fenêtres Ω. Prenons une seule fenêtre de taille $n \times n$ dont on va supposer la texture homogène. Sur cette fenêtre on estime $\hat{\Theta}$ aux moindres carrés par la formule[4] :

$$\hat{\Theta} = \left[\sum_\Omega Q_s Q_s^t\right]^{-1}\left[\sum_\Omega Q_s y_s\right]$$

avec $\quad Q_s = [y_{s+1},\ y_{s-1},\ y_{s+2},\ ...]^t$

et $\quad \hat{\sigma}^2 = \frac{1}{n^2}\sum_\Omega \left[y_s - \hat{\Theta}^t Q_s\right]^2$.

Une tâche délicate est l'estimation du nombre de classes qui se fait soit de façon supervisée, soit par analyse du nuage des points obtenus sur toutes les fenêtres. Supposons que l'on recherche N classes. On détermine les

[4] C'est ici que l'hypothèse gaussienne simplifie le calcul : l'estimation peut se faire par moindres carrés de façon explicite.

N meilleures classes du nuage des points obtenus sur toutes les fenêtres par exemple par un algorithme de nuées dynamiques. Cela nous fournit N vecteurs de paramètres $f_k = \theta_1^k, \theta_2^k, \theta_3^k, \theta_4^k, \mu^k, \sigma_2^k$.

On peut alors procéder à la phase de segmentation. On va désormais travailler au niveau des sites du champ de Markov. On introduit un champ de liaisons entre les pixels, décrit par la seule variable binaire (processus de bord) b qui vaut 0 si deux pixels voisins sont dans une même classe et 1 s'ils sont dans des classes différentes. On recherche la distribution des b et des $V_{s\lambda}$, variables caractéristiques de l'appartenance du pixel s à la classe λ ($V_{s\lambda} = 0$ ssi le pixel au site s appartient à la classe λ) qui minimise l'énergie :

$$U = \sum_s \sum_\lambda u(s,\lambda) V_{s\lambda} - (1-b) \sum_s \sum_\lambda \sum_{t \in N_s} V_{t\lambda} V_{s\lambda}$$

où le premier terme exprime l'attache aux données et le second l'a priori sur les régions. Les u incluent toute l'information connue sur le site s appartenant à la classe λ : par exemple $u(s,\lambda) = w(\lambda) + u_1(Ys,\lambda)$, u_1 étant comme dans la formule 4.2 et w étant un biais propre à la classe λ. L'optimisation d'une telle fonction d'énergie se fait soit par relaxation déterministe (par exemple l'ICM *Iterated Conditional Mode* qui remplace à chaque étape la classe λ par celle qui minimise l'énergie U au site considéré), soit par recuit simulé (relaxation stochastique) (cf. chapitre **??**).

4.4.2 La méthode de Kervrann et Heitz

Dans cette méthode [Kervrann et Heitz, 1995], on calcule sur des fenêtres recouvrant l'image, des matrices de cooccurrence (cf. section 4.5). En chaque fenêtre, on représente la texture par un vecteur d'attributs issus de ces matrices $O_s = \{o_s^1, o_s^2, \ldots, o_s^k\}$. Le terme d'attache aux données s'écrit sous la forme :

$$U_1 = \sum_{s \in \Omega} V(O_s, O_\lambda)$$

où O_s est le vecteur d'attributs du site s et O_λ est le vecteur d'attributs de la région d'étiquette λ affectée au site s. Le potentiel V s'écrit à partir de la distance Δ de Smirnov-Kolmogorov des composantes o_s^i et o_λ^i des vecteurs :

$$V(O_s, O_\lambda) = \sum_{i=1}^{k} \left[2\Gamma(\Delta(o_s^i, o_\lambda^i) > c^i) - 1 \right]$$

où la fonction Γ vaut 0 ou 1 selon que les valeurs de Δ sont supérieures ou inférieures à une valeur de seuil c^i issue des tables du test de Smirnov-Kolmogorov [Saporta, 1990].

Le champ évolue vers le critère du MAP (Maximum A Posteriori) par la technique de relaxation sous-optimale de l'ICM : pour chaque site on choisit la meilleure classe parmi les seules classes des voisins du site considéré (cf. chapitre **??**).

La construction des classes o_λ se fait de la façon suivante. Initialement il n'y a qu'une classe, identifiée au vecteur de la première fenêtre rencontrée. Les autres vecteurs de texture sont associés à cette classe si leur distance est inférieure au seuil donné. Afin de créer de nouvelles classes, on crée une classe « fourre-tout » dans laquelle sont rangées toutes les textures trop éloignées de celle déjà trouvée. Après un passage sur l'image, on détermine les paramètres de la classe trouvée (en calculant un vecteur O_λ sur toutes les réalisations trouvées) et on accroît de 1 le nombre de classes. On reprend le balayage de l'image. On itère ce procédé jusqu'à ce qu'un nombre pré-établi de classes soit trouvé ou qu'il n'y ait plus de candidats dans la classe fourre-tout.

Comme dans le cas précédent, la taille des fenêtres doit être :

1. assez grande pour permettre une caractérisation statistique réaliste des textures (et de leurs motifs) ;
2. assez petite pour conduire à des segmentations fines des détails.

Notons que cette méthode, au contraire de la précédente, ne permet pas de déterminer les potentiels et les cliques des textures retenues, autrement que par référence aux primitives de la matrice de cooccurrence.

4.5 Les méthodes structurales

Ces méthodes ne s'appliquent qu'aux textures construites sur une primitive bien identifiable qui se répète sur une maille régulière. L'analyse d'une texture homogène se fait en plusieurs étapes :

1. l'image est découpée en sous-fenêtres, chacune contenant plusieurs primitives (typiquement de l'ordre de 5×5 primitives),
2. sur chaque sous fenêtre on détermine les 2 vecteurs moyens de périodicité, par exemple par l'examen de la fonction d'auto-corrélation,
3. connaissant ces vecteurs, chaque sous-image est normalisée par un rééchantillonnage de façon à ce que les nouveaux vecteurs de périodicité soient tous dans une position de référence,
4. on calcule alors par TF le motif de la primitive en sommant dans l'espace de Fourier toutes les sous-fenêtres normalisées et en isolant le sous-domaine du fondamental,
5. on est alors en mesure de reconstruire une texture parfaite en répétant le motif moyen sur la maille moyenne, ou des textures plus ou moins parfaites en mélangeant des motifs réels ou moyens sur des mailles réelles ou moyennes.

Des méthodes plus structurales encore s'appuient sur des techniques de graphes ou de grammaires. Dans les approches par graphes, on représente les dépendances entre les diverses cmposantes de la texture par des graphes (arbres, cycles) mettant en évidence les dépendances entre pixels. Dans les approches par grammaires, on parcourt la texture ou des sous-ensembles de la texture par un balayage systématique de tous les sites, construisant ainsi une chaîne de descripteurs. Les relations entre les descripteurs successifs sont exprimées par des règles qui indiquent l'enchaînement des niveaux de gris le long du parcours. Les chaînes font alors l'objet d'un traitement syntaxique qui extrait ces dépendances [Lu et Fu, 1978, Lu et Fu, 1979].

4.6 Textures hétérogènes

Si l'on fait l'hypothèse que la texture sous-jacente est stationnaire, on peut déduire des variations de ses propriétés statistiques dans l'image, des éléments suffisants pour remonter à une connaissance de certains éléments de relief et d'orientation 3D des objets de la scène observée. Cela est naturellement fait par système visuel humain (cf. figure 4.2). Pour faire cela par traitement numérique, il convient de mesurer en chaque point de l'image (c'est-à-dire en des zones suffisamment petites autour de chaque point) des propriétés statistiques d'ordre 2 (prenant en compte les pixels 2 par 2, comme par exemple la décroissance de la corrélation à $1/2$), puis d'étudier les variations spatiales de ces propriétés. Ces variations sont alors reliées à l'orientation de la surface observée (généralement sous l'hypothèse que celle-ci fait un angle constant avec la direction d'observation).