

Segmentation d'images texturées par régularisation de graphes

Dans ce chapitre, nous présentons une approche permettant la prise en compte de caractéristiques de textures dans l'algorithme de segmentation par régularisation de graphe basé sur le modèle de Chan et Vese [Ta, 2009] présenté section 1.7.4.2. L'intérêt de cette contribution est de combiner deux techniques de traitement d'images qui manipulent les données relatives à l'image avec un plus haut niveau d'abstraction que les pixels qui la composent. Ces méthodes sont dites « non-locales ». La première est relative à l'utilisation d'une méthode basée sur les graphes, alors que la seconde repose sur l'utilisation de caractéristiques de textures.

Nous commencerons par établir un état de l'art des principales méthodes de caractérisation de textures, puis enchaînerons sur une présentation de notre approche. Nous illustrerons ensuite le fonctionnement de notre algorithme sur des images texturées artificielles, puis terminerons ce chapitre par une présentation des résultats obtenus sur des images échographiques 2D et 3D.

Les résultats de ces travaux ont été publiés dans [Fauchaux *et al.*, 2012b] et [Fauchaux *et al.*, 2012a].

2.1 État de l'art des méthodes de caractérisation de textures

La majorité des algorithmes de segmentation d'images travaille à partir de statistiques de premier ordre concernant la distribution des niveaux de gris : seule l'information relative à la probabilité d'obtenir un pixel d'un niveau de gris donné à un emplacement prédéfini est prise en compte.

Cependant, comme cela est décrit dans les travaux de Julesz sur le système psychovisuel humain [Julesz, 1962], et tel que l'illustre trivialement la figure 2.1, deux images de

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

textures « visuellement différentes¹ » peuvent présenter des statistiques de premier ordre identiques. Celles-ci ne sont donc pas suffisamment discriminantes, et limitent la capacité des algorithmes qui en dépendent à traiter des problèmes de vision complexes.

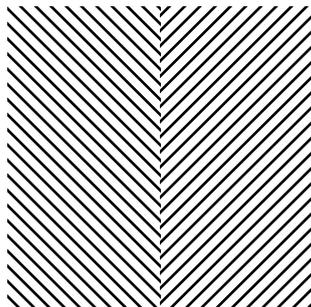


FIGURE 2.1 – Exemples de textures visuellement différentes mais dont les statistiques de premier ordre concernant les niveaux de gris sont identiques. Dans chacune des deux régions, le niveau de gris moyen vaut 192.

La notion de texture – bien qu'évidente pour notre cerveau – est relativement complexe à définir. Par ailleurs, comme l'illustre l'ensemble des définitions compilées dans [Tuceryan et Jain, 1990], celle-ci dépend souvent de l'objectif recherché par l'analyse d'une texture. En ce qui nous concerne, le point important est qu'au-delà de l'information relative à l'intensité des pixels, l'organisation spatiale de ces derniers joue un rôle extrêmement important. Ainsi, appliquée aux textures de la figure 2.1, une telle considération permet de mettre en évidence des continuités et discontinuités selon les angles de +45 et -45 degrés.

La notion de motifs permet d'offrir une description de plus haut niveau d'une texture, quantifiant ainsi son orientation, sa finesse ou encore sa périodicité. Ces concepts peuvent notamment être définis à l'aide de statistiques de second ordre, c'est à dire d'informations concernant la probabilité de trouver deux pixels de niveaux de gris donnés séparés par une distance et une orientation prédéfinies. Julesz souligne notamment que deux textures ayant des statistiques de second ordre identiques sont visuellement indiscernables.

Le rôle précis des différentes méthodes de caractérisation de textures est de fournir une description tonale et spatiale de la répartition des niveaux de gris dans une image. Dans la suite de cette section, nous allons présenter les principales méthodes de caractérisation de textures.

2.1.1 Caractéristiques d'Haralick

Haralick considère que la texture d'une image peut être décrite par le nombre et le type de ses primitives tonales ainsi que par leur organisation spatiale [Haralick *et al.*, 1973, Haralick, 1979]. Il propose donc de caractériser une texture à partir d'un ensemble de statistiques de second ordre qui permettent la mise en évidence des motifs dans la distribution et l'agencement des niveaux de gris qui la composent.

1. Julesz ne s'intéresse dans ses travaux qu'à la capacité du cerveau à discriminer sans efforts deux textures distinctes, c'est à dire sans procéder à un examen détaillé de celles-ci.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

Il s'appuie pour cela sur la notion de matrice de co-occurrences des niveaux de gris². Le but d'une matrice de co-occurrence est de dénombrer les occurrences de paires d'évènements. Dans le cadre de l'analyse de textures, Haralick considère cette paire d'évènements comme étant une transition de niveaux de gris, c'est à dire l'apparition de deux pixels d'intensités données séparés par une relation de voisinage $\Delta = (\delta_x, \delta_y)$ ³. Une telle matrice recense alors les différentes transitions de niveaux de gris dans l'image. Elle est le plus souvent utilisée sous sa forme normalisée, et représente alors une distribution de probabilités.

Soit I une image composée de N niveaux de gris. Étant donnée une relation de voisinage Δ , la matrice de co-occurrence associée à cette image est définie pour chaque paire de niveaux de gris i et j par :

$$C_{\Delta}(i, j) = \text{Card} \{ (\mathbf{p}, \mathbf{q}) \in I \text{ tel que } \mathbf{q} = \mathbf{p} + \Delta \text{ et } I(\mathbf{p}) = i \text{ et } I(\mathbf{q}) = j \} \quad , \quad (2.1)$$

ou \mathbf{p} et \mathbf{q} représentent les coordonnées des pixels, et $I(\cdot)$ l'intensité du niveau de gris aux coordonnées spécifiées.

La taille d'une matrice de co-occurrence est quadratiquement proportionnelle au nombre de niveaux de gris dans l'image. Afin de permettre l'émergence de certaines tendances ou schémas dans la matrice, il est important de ne pas considérer un nombre trop important de niveaux de gris, au risque d'obtenir des valeurs statistiquement insignifiantes. Il est ainsi courant de réduire l'étendue des valeurs possibles en ré-échantillonnant l'image initiale, projetant ainsi l'ensemble des niveaux de gris dans un espace de taille réduite. Il n'existe cependant aucune règle indiquant le nombre de niveaux de gris nécessaires à une description optimale d'une texture, mais des valeurs situées entre 16 et 32 sont relativement courantes.

Dans ses travaux, Haralick ne tient pas compte de la direction de la relation de voisinage. Lors du calcul, il est donc nécessaire de considérer la relation de voisinage initiale ainsi que son opposée :

$$C_{\Delta}(i, j) = \text{Card} \{ (\mathbf{p}, \mathbf{q}) \in I \text{ tel que } \mathbf{q} = \mathbf{p} + \Delta \text{ et } (I(\mathbf{p}) = i \text{ et } I(\mathbf{q}) = j \text{ ou } I(\mathbf{p}) = j \text{ et } I(\mathbf{q}) = i) \} \quad . \quad (2.2)$$

La matrice de co-occurrence ainsi générée est symétrique : $C_{\Delta}(i, j) = C_{\Delta}(j, i)$. La figure 2.2 présente une image synthétique et la matrice de co-occurrences obtenue pour le décalage $\Delta = (1, 0)$.

À partir de la matrice de co-occurrences normalisée, Haralick propose un ensemble de quatorze statistiques permettant d'évaluer différentes propriétés de la texture analysée, telles que son niveau de contraste, sa régularité ou encore son degré d'entropie. Le lecteur intéressé trouvera en annexe A plus d'informations concernant le calcul de ces caractéristiques. Nous y faisons notamment le point sur les différents noms et formulations que la littérature leur associe et discutons de la validité de certaines d'entre elles. Nous en profitons aussi pour compléter les travaux d'Haralick et proposons une interprétation de chacune des caractéristiques ainsi qu'une évaluation de leur domaine de valeurs.

2. Aussi appelée matrice de dépendance spatiale des niveaux de gris, ou encore GLCM (de l'anglais *Grey-Level Co-occurrence Matrix*).

3. La notation Δ sera conservée car cette notion de relation de voisinage peut être étendue à des dimensions supérieures.

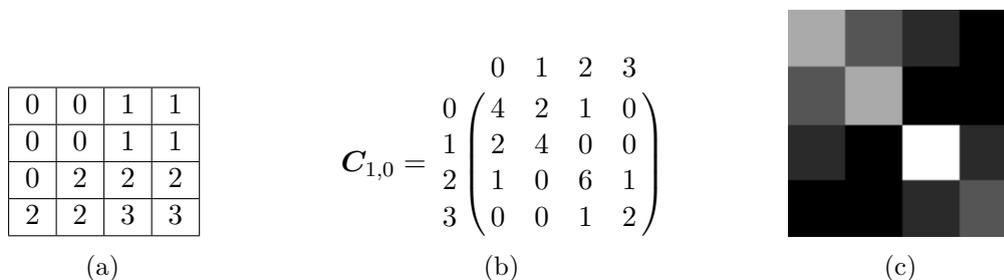


FIGURE 2.2 – Exemple de matrice de co-occurrences. (a) : une image de taille 4×4 composée de 4 niveaux de gris. (b) : la matrice de co-occurrences calculée pour le décalage $\Delta = (1, 0)$. (c) : la même matrice représentée sous forme de carte.

Étant donné que le décalage Δ définit la direction selon laquelle une texture est étudiée, une matrice de co-occurrences est un outil fortement anisotrope qui ne caractérisera pas de la même manière deux textures de motifs identiques mais d'orientations différentes. Afin d'assurer une invariance à la rotation, Haralick suggère de caractériser une texture selon quatre orientations, 0° , 45° , 90° et 135° , puis de considérer deux vecteurs composés respectivement de la moyenne et de l'écart-type de chacune des quatorze caractéristiques.

2.1.2 Modèles auto-régressifs

Un modèle auto-régressif (*AutoRegressive Model*, ou *AR model*) est un outil statistique qui vise à exprimer un processus aléatoire discret à valeurs réelles par le biais d'une combinaison linéaire de ses propres valeurs passées :

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon \quad , \quad (2.3)$$

avec X le processus aléatoire et ε un bruit blanc d'écart-type σ_ε dont le but est de donner un caractère non-déterministe au modèle. On appelle « paramètres du modèle » les variables $\{\varphi_1, \dots, \varphi_p\}$ qui peuvent être considéré comme une représentation spectrale du signal, et donc une caractéristique de textures. Lorsque le modèle fait référence à p variables passées, il est dit d'ordre p .

À partir d'un échantillon $X = \{X_1, \dots, X_n\}$, la conception d'un modèle auto-régressif consiste à estimer les paramètres φ_i ainsi que l'écart-type σ_ε du bruit en optimisant un critère d'attache aux données compte tenu de la présence d'un élément de caractère aléatoire. Plusieurs méthodes sont disponibles, la plupart étant basées sur des outils statistiques : l'estimation du maximum de vraisemblance [Parzen, 1968], les moindres carrés [Akaike, 1969] ou encore les équations de Yule-Walker (application du raisonnement de la méthode des moments à partir d'une estimation des auto-covariances des éléments de l'échantillon) [Chan et Langford, 1982]. D'autres approches, notamment optimisées pour les données 2D, ont plus récemment vu le jour, telles que l'algorithme rapide 2D des moindres carrés récursif en treillis [Alata et Olivier, 2003].

Initialement appliqués à la modélisation et à la prédiction de signaux 1D tels que les séries temporelles [Akaike, 1969], les modèles auto-régressifs ont ensuite été étendus

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

aux signaux 2D pour le traitement d'images, tout d'abord en les considérant comme des signaux 1D, puis en intégrant la structure matricielle des images à la modélisation [Haralick, 1979, Maragos *et al.*, 1984]. Les modèles auto-régressifs sont utilisés dans des applications telles que la synthèse de textures [McCormick et Jayaramamurthy, 1974], la segmentation [Deguchi et Morishita, 1978, Bouman et Liu, 1991, Ergen, 2012], ou encore la compression d'image [Maragos *et al.*, 1984, Debure et Kubato, 1998].

Lorsqu'un tel modèle est appliqué au traitement d'images, excepté pour certains cas de synthèse de textures, la notion de « valeur passée » peut ne plus être pertinente. Le modèle peut ainsi considérer l'ensemble de pixels situés dans un voisinage donné – il est alors qualifié de modèle auto-régressif symétrique (*SAR*, pour *Symmetric AutoRegressive*). Dans cette configuration, il est alors possible de rendre le modèle invariant à la rotation : *CSAR* (*Circular Symmetric AutoRegressive model*) [Kashyap et Khotanzad, 1986], *RISAR* (*Rotation-Invariant Symmetric AutoRegressive model*) [Mao et Jain, 1992].

Enfin, des travaux ont permis d'appliquer cette approche dans un contexte de multi-résolution : *MR-SAR*, *MR-RISAR* [Mao et Jain, 1992].

2.1.3 Caractéristiques de textures dérivées de méthodes de traitement du signal

Ces méthodes sont basées sur un principe commun : mesurer la similarité du signal à étudier avec un second signal, connu, possédant des propriétés particulières (fréquence, forme...). La mesure ainsi obtenue permet de quantifier la présence de ces propriétés dans le signal étudié.

Mathématiquement, cela consiste à calculer la corrélation croisée entre ces deux signaux. Dans le cadre du traitement de signaux continus à une dimension, si $f(\cdot)$ désigne la fonction à étudier, et $g(\cdot)$ la fonction à laquelle f est comparée, la corrélation croisée ($f \star g$) entre f et g est obtenue par convolution de f avec le conjugué complexe \bar{g} de g :

$$(f \star g) = \int_{-\infty}^{\infty} f(\tau) \cdot \bar{g}(\tau) \, d\tau \quad (2.4)$$

Cette formule est transposable aux signaux discrets, définis sur un support compact, ou encore à plusieurs dimensions. En particulier, lorsque appliquée au traitement d'images 2D, la formule associée à ce calcul de corrélation croisée est :

$$(f \star g) = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cdot \bar{g}(x, y) , \quad (2.5)$$

une image de résolution $M \times N$ étant ici considérée.

2.1.3.1 Transformée de Fourier

La transformée de Fourier est une transformation mathématique qui vise à décrire les composantes d'un signal quelconque non plus dans un contexte temporel ou spatial – c'est à dire en exprimant la valeur du signal en chacune de ses coordonnées – mais fréquentiel.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

Elle repose pour cela sur le théorème de Fourier, qui énonce que tout signal réel continu peut être exprimé sous la forme d'un polynôme trigonométrique (de taille potentiellement infinie), c'est à dire comme une somme de fonctions sinus et cosinus d'amplitudes et de fréquences déterminées.

Dans sa version 2D discrète, la transformée de Fourier s'appelle la DFT^4 et s'exprime à partir de l'équation 2.5 en considérant la fonction g comme une sinusoïde complexe :

$$\mathcal{F}(u, v) = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \exp\left(-2i\pi\left(\frac{xu}{N} + \frac{yv}{M}\right)\right), \quad (2.6)$$

u et v étant les paramètres qui caractérisent la période de la sinusoïde suivant les deux axes du plan, et $\mathcal{F}(u, v)$ le coefficient de Fourier associé.

La transformée de Fourier peut être vue comme une décomposition⁵ en composantes fréquentielles, l'amplitude d'un coefficient de Fourier $\mathcal{F}(u, v)$ étant proportionnelle à la contribution de la composante de paramètres u et v dans l'image d'origine. Le spectre de puissance d'une image est ainsi défini comme étant la magnitude (aussi nommé valeur absolue ou encore module) de sa transformée de Fourier. Il est ainsi envisageable de caractériser l'aspect fréquentiel d'une texture, en particulier sa période et son orientation, par une analyse des coefficients de Fourier de forte amplitude (voir figure 2.3).

Plus particulièrement, la présence de couronnes ainsi que de secteurs angulaires présentant une forte énergie permet de quantifier la granularité d'une texture ainsi que son orientation (voir figure 2.4).

Cette approche a notamment été employée pour de la reconnaissance de terrains par classification de textures [Weszka, 1976].

2.1.3.2 Transformée en ondelettes

L'un des reproches souvent adressé à la transformée de Fourier est son incapacité à fournir des informations concernant la localisation des composantes fréquentielles dans le temps. Par analogie avec la musique, cela signifie qu'il est possible de connaître l'ensemble des fréquences jouées, mais il est en revanche impossible de savoir si une note précise a été jouée à un instant donné.

Afin de combler cette lacune, Gabor initia la théorie des ondelettes en proposant de limiter la portée spatiale de la sinusoïde présente dans la transformée de Fourier traditionnelle [Gabor, 1946]. Il employa pour cela une transformée de Fourier fenêtrée, dans laquelle le signal à analyser est convolué avec une sinusoïde modulée par une fonction Gaussienne. La sinusoïde a désormais la forme d'une petite onde, ce qui donnera naissance à la notion d'ondelette : un signal oscillant défini sur un support compact. Ce type de signal est aussi connu sous le nom d'« atome de Gabor ».

Tout comme la transformée de Fourier traditionnelle, celle de Gabor permet la décomposition ainsi que la reconstruction d'un signal selon ses composantes fréquentielles.

4. Pour *Discrete Fourier Transform*.

5. La transformée de Fourier est inversible : il est possible de reconstruire l'image à partir de ses composantes fréquentielles.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

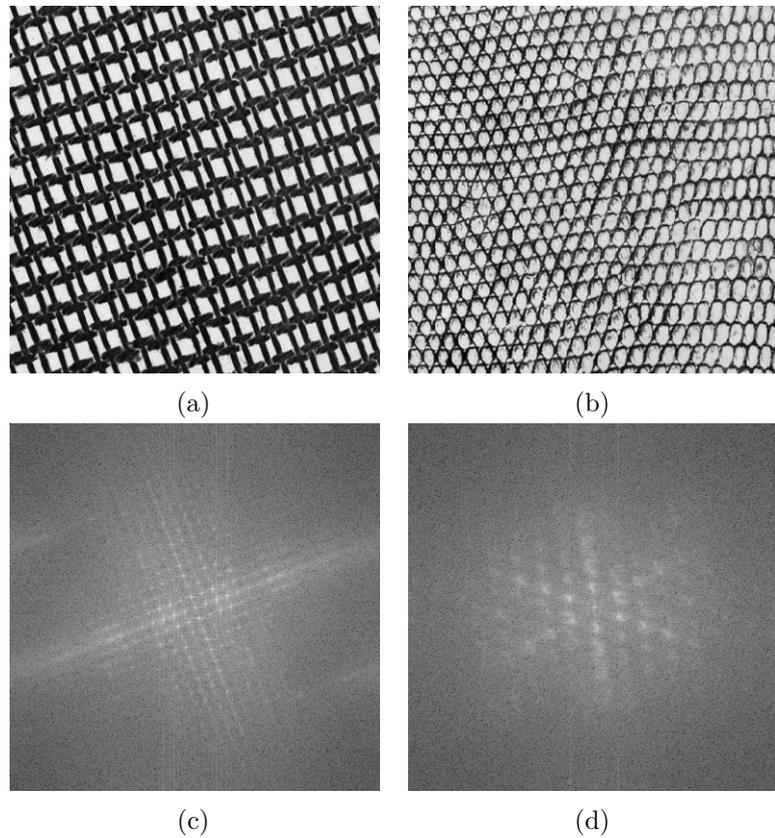


FIGURE 2.3 – (a), (b) : Exemples de textures issues de la base d'images Brodatz [Brodatz, 1966]. (c), (d) : Leurs spectres de puissance respectifs.

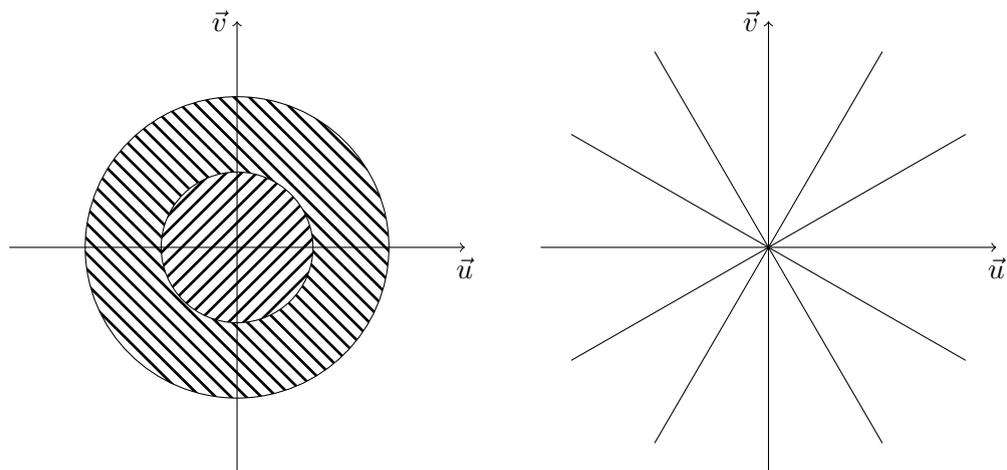


FIGURE 2.4 – Régions du spectre de puissance permettant l'identification de la fréquence (couronnes) et de l'orientation (secteurs angulaires) des composantes fréquentielles d'une texture.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

La décomposition ainsi obtenue est cependant redondante, car l'ensemble formé par les translations en temps et en fréquence de l'ondelette mère ne constitue pas une base de l'espace à analyser. Cette sous-optimalité est sans aucun doute une des raisons du peu d'intérêt porté à cette transformée. Une version parcimonieuse de cette approche, les filtres de Gabor (voir section 2.1.3.3), aura cependant un plus grand succès.

La théorie des ondelettes a connu un essor important dans les années 80. Alors qu'il travaillait sur l'étude de signaux sismiques, Morlet proposa de remplacer les atomes de Gabor par d'autres fonctions oscillantes : les ondelettes [Grossmann et Morlet, 1984]. À partir d'une ondelette mère $\psi(x)$, et plus particulièrement de ses translations et dilatations $\psi_{s,t}(x) = \sqrt{s} \psi(sx + t)$, $(s, t) \in \mathbb{R}^+ \times \mathbb{R}$, il devient possible de décomposer tout signal continu réel de carré intégrable.

Dans [Meyer, 1985], il a été prouvé l'existence d'ondelettes mères $\psi(x)$ telles que l'ensemble des translations et dilatations $\psi_{j,k}(x) = \sqrt{2^j} \psi(2^j x + k)$, $(j, k) \in \mathbb{Z}^2$ d'une ondelette forme une base de l'espace des fonctions réelles de carré intégrable. Ce résultat est particulièrement important, puisqu'il permet d'assurer l'absence de redondance dans les éléments issus d'une décomposition en ondelettes.

Les méthodes de décomposition basées sur les ondelettes ont la particularité de fournir une analyse multi-échelle d'un signal. En effet, si l'on considère une ondelette $\psi()$ et la base $\{\psi_{j,k}\}_{(j,k) \in \mathbb{Z}^2}$ générée par ses translations et dilatations, les coefficients de la transformée associés à un niveau de dilatation j sont représentatifs du détail du signal étudié à l'échelle 2^j , tandis que les coefficients liés à des échelles supérieures fournissent des informations liées à des niveaux de détails plus fins.

Cependant, la décomposition complète d'un signal requiert de considérer l'ensemble des échelles possibles. À partir de ce constat, Mallat développe une méthode de décomposition multi-résolution de signal [Mallat, 1989]. L'idée est d'associer à une ondelette mère $\psi()$ une fonction d'échelle $\phi()$, elle aussi utilisée pour générer une famille de fonctions $\phi_{j,k}(x) = \sqrt{2^j} \phi(2^j x + k)$, $(j, k) \in \mathbb{Z}^2$, et dont le but est de fournir une série d'approximations d'un signal à des échelles successives⁶.

Ces deux fonctions sont mathématiquement liées de façon à ce qu'à une échelle 2^j et une translation k données, les fonctions $\phi_{j,k}()$ et $\psi_{j,k}()$ soient complémentaires. En d'autres termes, l'ondelette fait office de filtre passe-haut et permet l'extraction des composantes de haute fréquence, tandis que la fonction d'échelle joue le rôle de filtre passe-bas et permet d'approximer le signal, codant ainsi ce qui n'a pas pu l'être avec l'ondelette.

Ce processus de décomposition est en réalité implémenté à l'aide de deux filtres de convolution, G et H , lesquels sont respectivement générés à partir de l'ondelette et de la fonction d'échelle. Un signal initial S est ainsi convolué avec chacun de ces deux filtres, les réponses obtenues étant ensuite rééchantillonnées avec un facteur de 2. Les valeurs obtenues à partir des filtres G et H forment respectivement ce que l'on appelle les coefficients de détails et d'approximation. Ce processus peut ensuite être réappliqué récursivement en considérant les coefficients d'approximation en lieu et place du signal initial. Ces coefficients représentant une approximation du signal à une échelle inférieure, chaque itération permet

6. À ce titre, cette famille de fonctions ne constitue pas une base de l'espace des fonctions de carré intégrable car toute approximation à une échelle 2^j donnée contient une partie de l'information contenue à l'échelle 2^{j-1}

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

d'obtenir une nouvelle description du signal à une nouvelle échelle, et donc pour de nouvelles bandes de fréquence. Grâce à l'utilisation de la fonction d'approximation, ce processus peut être arrêté à tout moment, une décomposition en n niveaux étant au final constituée de n vecteurs de coefficients de détail et d'un unique vecteur de coefficients d'approximation. La figure 2.5 illustre le principe d'une décomposition en ondelettes 1D.

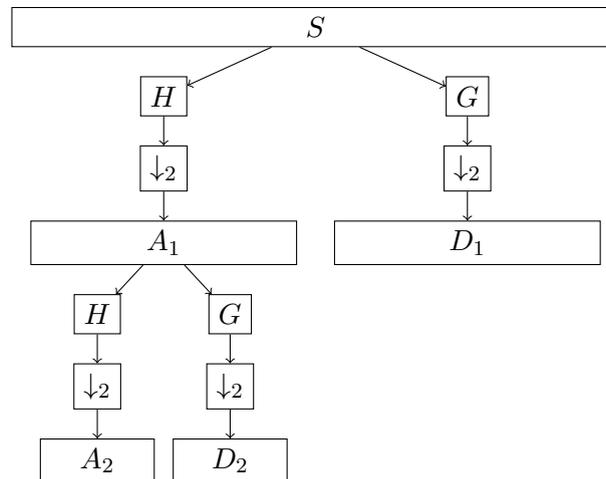


FIGURE 2.5 – Décomposition en ondelettes d'un signal 1D. S correspond au signal initial, A_x et D_x respectivement aux coefficients d'approximation et de détails associés à l'échelle x , G et H aux filtres passe-haut et passe-bas associés respectivement à l'ondelette et à la fonction d'échelle. \downarrow_2 dénote un rééchantillonnage d'un facteur 2.

Le passage à des signaux 2D discrets tels que les images est relativement simple à mettre en place. Comme illustré figure 2.6, chaque étape de l'algorithme de décomposition se passe désormais en deux temps. Chaque ligne du signal est tout d'abord décomposée, ce qui génère deux jeux de coefficients :

- D : détail horizontal,
- A : approximation horizontale.

Les colonnes de ces deux jeux de coefficients sont ensuite décomposées de la même façon. Nous obtenons ainsi quatre nouveaux jeux de coefficients :

- DD : détail,
- DA : détail horizontal et approximation verticale,
- AD : détail verticale et approximation horizontale,
- AA : approximation.

Le processus de décomposition peut ensuite être poursuivi à partir des coefficients d'approximation AA. La figure 2.6 illustre le principe d'une décomposition en ondelettes 2D.

Lors d'une transformée en ondelettes, la décomposition d'un signal est obtenue en traitant récursivement les coefficients d'approximation globale, ce qui permet une représentation de plus en plus fine des basses fréquences. Afin de permettre une meilleure représentation des hautes fréquences, la méthode des paquets d'ondelettes propose de décomposer aussi les coefficients de détails obtenus pour chaque échelle [Coifman *et al.*, 1992]. La décomposition passe ainsi d'une décomposition de type pyramidal en une décomposition

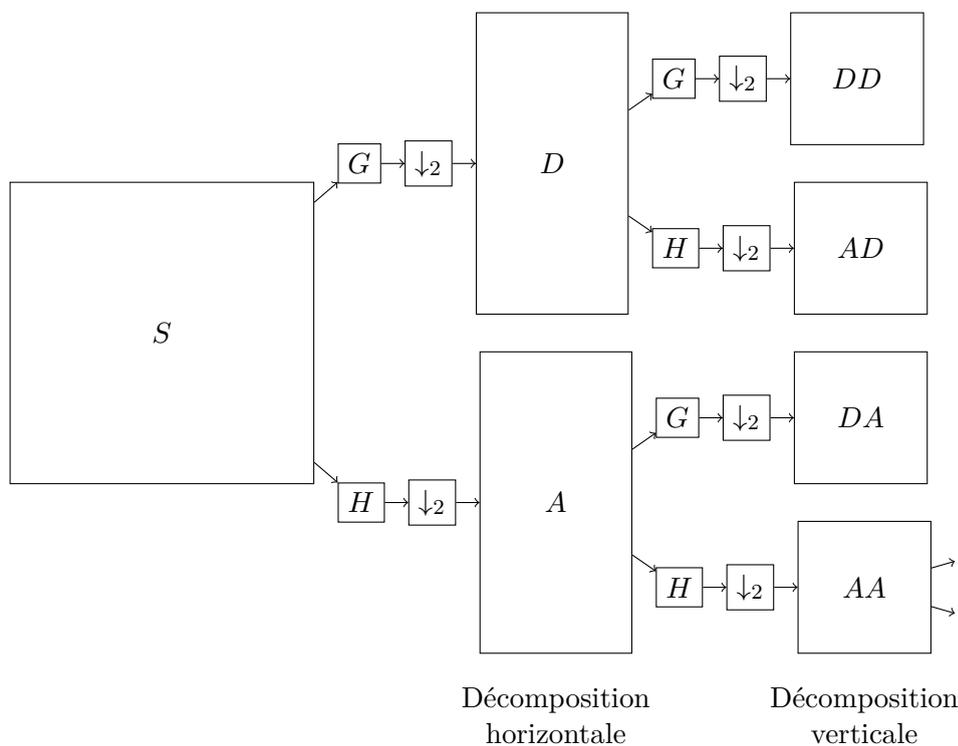


FIGURE 2.6 – Décomposition en ondelettes d'un signal 2D. Le signal est tout d'abord décomposé horizontalement, puis verticalement.

de type arborescente.

Coifman et Wickerhauser introduisent également un algorithme de sélection de la meilleure base [Coifman et Wickerhauser, 1992]. Il permet d'optimiser la représentation d'un signal à partir de ses coefficients d'ondelettes en mesurant la « quantité d'information » contenue par chaque décomposition. Sa mise en œuvre consiste à réaliser une décomposition par paquets complète du signal puis, en partant du niveau le plus bas de l'arborescence, à comparer l'énergie⁷ de chaque décomposition à la somme des énergies des décompositions filles. Ainsi, si l'une d'entre elles présente une énergie largement supérieure à celle de ses filles, cela signifie qu'il n'est pas nécessaire de la décomposer car aucune information complémentaire ne pourra être obtenue. Les décompositions filles qui ne satisfont pas cette condition sont donc supprimées de l'arbre de décomposition. Lorsque cet algorithme a été appliqué à la totalité de l'arbre, les décompositions associées à l'ensemble des feuilles forment une base qui représente de façon optimale le signal initial.

C'est précisément cette approche qui a été retenue par les premiers travaux de caractérisation de textures basés sur la décomposition en ondelettes. Dans [Laine et Fan, 1993], cette mesure d'énergie sert à caractériser les différentes bandes de fréquences que représente la décomposition. Une signature de la texture peut ainsi être créée, puis identifiée à l'aide d'un classificateur basé sur une mesure de distance ou encore un réseau de neurones. Selon un principe similaire, il a aussi été proposé de réduire la signature d'une texture aux énergies

7. Une mesure alternative basée sur l'entropie est aussi proposée.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

des k décompositions les plus pertinentes [Chang et Kuo, 1993].

2.1.3.3 Bancs de filtres

À la différence de la transformée de Fourier qui réalise une décomposition exhaustive d'une texture en ses composantes fréquentielles, les méthodes à bancs de filtres visent à caractériser une texture par une mesure de la corrélation de celle-ci avec un ensemble prédéfini d'éléments de textures ayant des propriétés (forme, orientation, échelle...) particulières. De par l'utilisation de filtres localisés dans l'espace – alors que la transformée de Fourier ne s'intéresse qu'à la composante fréquentielle – cette approche permet de caractériser une texture en un point précis d'une image.

Les mesures de corrélations obtenues en un pixel donné par l'ensemble des filtres composant un banc forment ce que l'on appelle couramment la « réponse du banc ». Celle-ci peut être vue comme la signature de la texture présente en ce point vis-à-vis de ce banc et peut, associée à une mesure de distance, être utilisée pour classifier des textures.

Filtres de Gabor [Gabor, 1946]

Les filtres de Gabor sont les premiers à avoir été appliqués à la caractérisation de textures [Turner, 1986]. Ils correspondent à un cas particulier de la transformée de Fourier, le contrôle de la phase de la sinusoïde ayant été ajouté. Par ailleurs, afin de limiter la portée de la transformation, et ainsi d'assurer la localisation spatiale de la transformée, celle-ci est fenêtrée à l'aide d'une enveloppe gaussienne. Un filtre de Gabor s'exprime alors par :

$$g(x, y, u, v, \phi, \sigma) = \exp\left(2i\pi(ux + vy) + \phi\right) \cdot \exp\left(-\left(\frac{x^2 + y^2}{2\sigma^2}\right)\right), \quad (2.7)$$

avec u et v les paramètres définissant la fréquence et l'orientation de la sinusoïde, ϕ sa phase et σ l'écart-type associé à l'enveloppe gaussienne.

La réponse obtenue par convolution d'un filtre en chacun des pixels composant la texture permet de caractériser une texture. Un banc composé de 16, 24, 32 ou 48 filtres combinant plusieurs orientations, échelles et phases est généralement utilisé (voir figure 2.7).

Dans [Manjunath et Ma, 1996a], il est proposé de définir une signature d'une texture à l'aide de statistiques de premier ordre (moyenne, écart-type) pour chacun des filtres du banc. Des algorithmes de partitionnement, tels que les k -moyennes ou encore les arbres-kd peuvent aussi être utilisés afin de fournir un ensemble de « réponses types » d'une texture [Jain et Farrokhnia, 1991, Rubner et Tomasi, 1999].

Filtres de Schmid [Schmid, 2001]

L'auteur s'inspire des filtres de Gabor pour proposer un banc de treize filtres dit « génériques » qui résultent de la composition d'une sinusoïde exprimée sous forme radiale et d'une fenêtre gaussienne :

$$g(x, y, \sigma, \tau) = F_0(\sigma, \tau) + \cos\left(\frac{\pi\tau\sqrt{x^2 + y^2}}{\sigma}\right) \cdot \exp\left(\frac{x^2 + y^2}{2\sigma^2}\right), \quad (2.8)$$

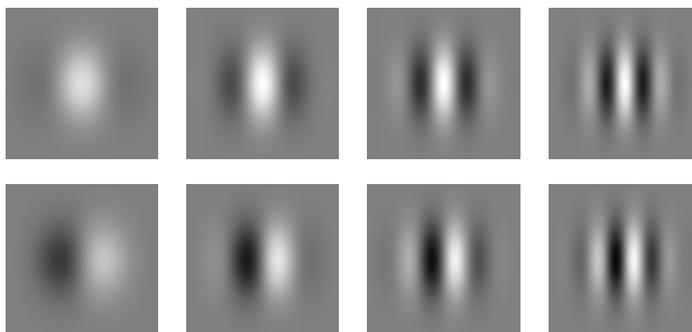
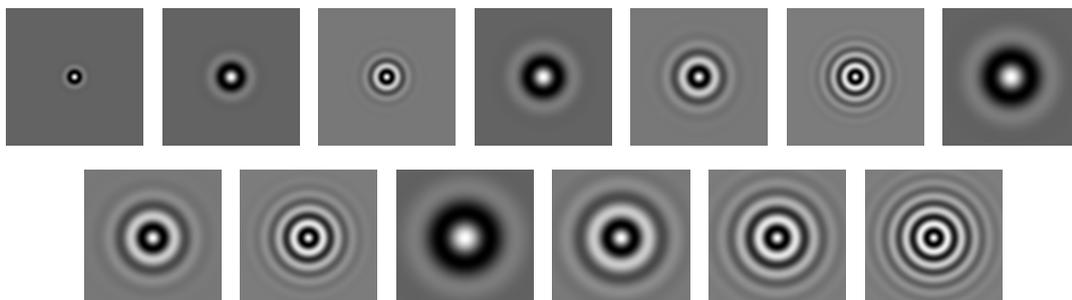


FIGURE 2.7 – Exemples de filtres de Gabor.

Exemples de filtres de Gabor pour différentes valeurs de phase, d'échelles et de fréquences. Le niveau de gris correspond à la valeur du filtre en chaque pixels.

τ représentant le nombre de cycle de la fonction harmonique, σ l'écart-type associé à la fenêtre gaussienne, et F_0 une compensation permettant de donner au filtre une moyenne nulle (les filtres sont par ailleurs généralement normalisés).

Le banc proposé est généré pour (σ, τ) prenant les valeurs de $(2, 1)$, $(4, 1)$, $(4, 2)$, $(6, 1)$, $(6, 2)$, $(6, 3)$, $(8, 1)$, $(8, 2)$, $(8, 3)$, $(10, 1)$, $(10, 2)$, $(10, 3)$, $(10, 4)$ (voir figure 2.8).

FIGURE 2.8 – Filtres de Schmid obtenus pour les valeurs (σ, τ) de $(2, 1)$, $(4, 1)$, $(4, 2)$, $(6, 1)$, $(6, 2)$, $(6, 3)$, $(8, 1)$, $(8, 2)$, $(8, 3)$, $(10, 1)$, $(10, 2)$, $(10, 3)$, $(10, 4)$.

L'utilisation de filtres présentant une symétrie radiale permet à l'auteur de générer un ensemble de caractéristiques de textures invariantes à la rotation. De la même façon que les approches appliquées aux filtres de Gabor, Schmid a recours à l'algorithme des k-moyennes pour générer un ensemble de signatures types. Il justifie alors l'intérêt d'un banc de taille réduite (comparé aux filtres de Gabor de taille 32 ou 48 régulièrement utilisés) par une réduction des risques liés à l'emploi de descripteurs de trop grande dimensionnalité.

Filtres de Leung-Malik (LM) [Leung et Malik, 2001]

Le banc LM est composé de 48 filtres générés à partir des dérivées premières et secondes de Gaussiennes selon 6 orientations et 3 échelles, 4 Gaussiennes et 8 Laplaciens de Gaussiennes (voir figure 2.9).

Ce banc de filtres est là encore couplé à l'algorithme des k-moyennes afin d'établir un dictionnaire de réponses types. L'originalité de cette approche réside dans l'application de

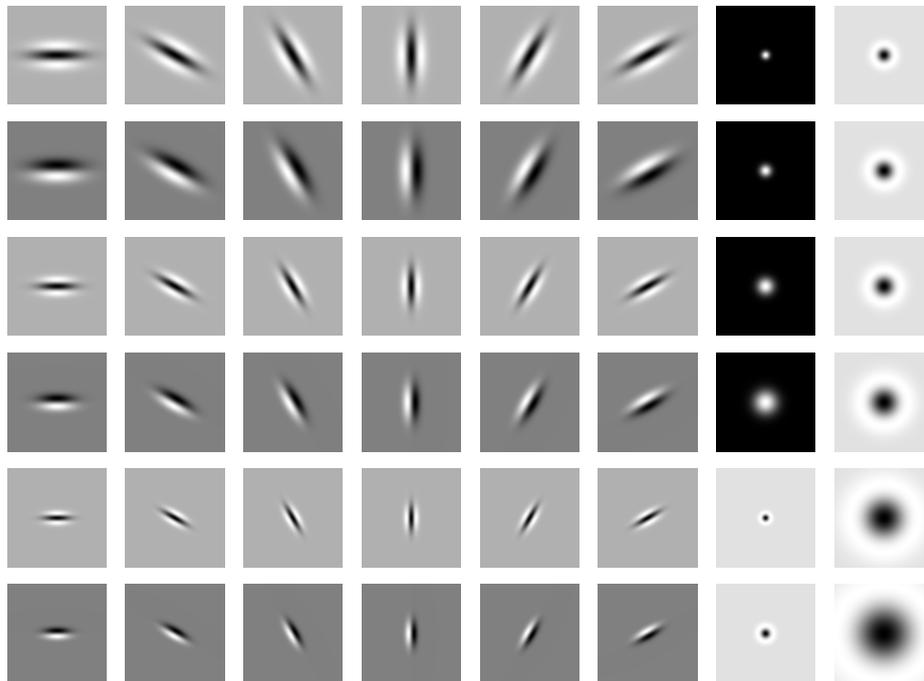


FIGURE 2.9 – Filtres de Leung-Malik.

ces filtres à la reconnaissance de textures 3D, composées des acquisitions photographiques de l'état de surface de matériaux réalisées sous différentes combinaisons d'angles d'incidence et de conditions d'éclairage.

Filtres de réponse maximum (MR) [Varma et Zisserman, 2002]

Les bancs MR⁸ sont de conception similaire au banc LM, puisque composés d'une Gaussienne, de son Laplacien, ainsi que des dérivées premières et secondes de Gaussiennes selon plusieurs échelles et orientations. Au-delà de la réduction du nombre de Gaussiennes et de Laplaciens, la différence vient du fait que seules les réponses des filtres directionnels présentant le plus grand intérêt seront par la suite considérées. En particulier, afin de permettre l'obtention d'une signature de texture invariante à la rotation, seule la réponse maximum d'un filtre parmi toutes ses orientations sera prise en compte.

Le banc MR4, illustré figure 2.10, ne considère qu'une unique échelle pour 6 orientations. 14 filtres sont ainsi générés, pour seulement 4 réponses conservées. Le banc MR8 considère quant à lui 3 échelles, portant le nombre de filtres à 38, et à 8 le nombre de réponses. Le banc MR4S, composé des mêmes filtres que le banc MR8, assure enfin l'invariance à la rotation ainsi qu'au changement d'échelle en ne conservant qu'une unique réponse parmi l'ensemble des orientations et échelles possibles d'une même famille de filtres.

Dans [Varma et Zisserman, 2002, Varma et Zisserman, 2005], ces filtres ont été comparés à ceux de Leung-Malik ainsi qu'à ceux de Schmid dans le cadre de la classification de textures 3D. Les tests réalisés montrent que les bancs MR8 et de Schmid obtiennent de

8. Pour *Maximum Response*.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

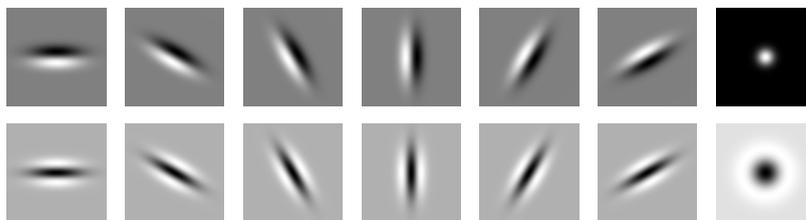


FIGURE 2.10 – Banc de filtres MR4. Pour chaque filtre directionnel, une unique réponse sera conservée parmi toutes les orientations possibles, permettant ainsi l'obtention d'un descripteur de texture invariant à la rotation.

meilleurs résultats que le banc LM, soulignant non seulement l'intérêt de travailler avec des descripteurs invariants à la rotation, mais aussi la possibilité de travailler avec des descripteurs de taille réduite. En revanche, les mauvaises performances du banc MR4 indiquent, quant à elles, la pertinence des approches multi-échelles.

2.1.4 Modèle binaire local

Plus connu sous son nom anglais, *local binary patterns* (LBP), cette méthode s'inspire des travaux réalisés sur les spectres de textures [Wang et He, 1990], dans lesquels un code est associé à chaque pixel en fonction de la différence de contraste de ce pixel avec chacun de ses voisins.

Wang et He définissent la texture d'un pixel à l'aide d'un code ternaire. Si l'on considère un pixel p , chacun de ses huit plus proches voisins se voit associé une valeur 0, 1 ou 2 indiquant si le niveau de gris de celui-ci est respectivement inférieur, égal ou supérieur au niveau de gris du pixel p . Selon un ordre de parcours du voisinage prédéfini, ces différentes valeurs permettent de construire un code ternaire qui identifie la texture localement associée au pixel p . L'ensemble de ces codes définit ce que l'on appelle un dictionnaire d'« unités de texture ».

Afin de réduire la taille de ce dictionnaire, il a été proposé dans [Ojala *et al.*, 1996] d'utiliser un code binaire en ne considérant que les valeurs 0 et 1 pour les états « inférieur ou égal » et « supérieur » (voir figure 2.11). Le dictionnaire passe ainsi de $3^8 = 6561$ éléments pour les spectres de textures à $2^8 = 256$ pour la méthode des motifs binaires locaux. À partir des codes obtenus pour l'ensemble des pixels d'une région représentative d'une texture, un histogramme des motifs présents est construit. Ces histogrammes peuvent alors être comparés par le biais de mesures de similarités, permettant ainsi la reconnaissance et la classification des textures associées.

Étant basée sur une mesure du contraste, cette méthode est invariante aux changements d'intensité. Afin de permettre une analyse multiéchelle, il est proposé dans [Ojala *et al.*, 2002] de recourir à un voisinage constitué de N points équirépartis sur un cercle de rayon R , le choix des valeurs associées à ces deux paramètres étant laissé à l'utilisateur. Une méthode permettant de rendre les motifs binaires locaux invariants à la rotation y est également proposée. Des extensions à la 3D ont par ailleurs été proposées [Fehr et Burkhardt, 2008, Paulhac *et al.*, 2008].

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

61	70	68
45	42	42
36	35	37

(a)

1	1	1
1		0
0	0	0

(b)

1	2	4
16		8
32	64	128

(c)

FIGURE 2.11 – Méthode de calcul d'un motif binaire local. (a) : un pixel et son voisinage, (b) : motif binaire local, (c) : pondération associée aux différents pixels du voisinage. Le motif de valeur $1 + 2 + 4 + 16 = 23$ est attribué au pixel.

Les différentes invariances que propose cette approche font des motifs binaires locaux une méthode robuste, capable de fournir des caractéristiques à la fois statistiques et structurelles sur une texture. Sa simplicité en fait par ailleurs un outil de choix pour les applications temps réel, comme par exemple la reconnaissance de visages [Zhao et Pietikäinen, 2006, Lei *et al.*, 2013].

2.1.5 Champs de Markov aléatoires

Un champs de Markov est un outil statistique qui permet de modéliser un phénomène stochastique par une analyse de ses dépendances spatiales. D'un point de vue mathématique, un tel phénomène correspond à un ensemble de variables aléatoires mises en relation par un système de voisinage. Sa modélisation à l'aide d'un champ de Markov aléatoire revient à exprimer la probabilité de réalisation de chacune des variables uniquement en fonction des valeurs courantes des variables aléatoires voisines.

La notion de texture étant intimement liée à la notion de dépendance spatiale, la modélisation d'une image par champs de Markov constitue une approche relativement populaire qui trouve des applications telles que la synthèse et la compression d'images [Cross et Jain, 1983, Chellappa *et al.*, 1985, Efros et Leung, 1999], le débruitage d'images [Malfait et Roose, 1997, Pierce et Ulaby, 2002] ainsi que la classification et la segmentation d'images [Manjunath et Chellappa, 1991, Wang et Liu, 1999, Hiremath *et al.*, 2012, Voisin, 2012].

Dans la terminologie des champs de Markov, une image composée de N pixels est représentée par un ensemble de N sites $S = \{S_1, \dots, S_N\}$. S est muni d'un système de voisinage V qui associe à chaque site $x \in S$ un ensemble de sites voisins $V_x \subset S$ et qui doit vérifier les propriétés suivantes :

$$x \notin V_x \quad , \quad (2.9)$$

$$\forall y \in S, x \in V_y \Leftrightarrow y \in V_x \quad . \quad (2.10)$$

En traitement d'images, le voisinage d'un site s correspond généralement à l'ensemble des pixels situés à une distance de s inférieure à un certain seuil. La taille du voisinage n'est cependant pas donnée par la valeur de ce seuil mais par un ordre de grandeur (voir figure 2.12). On parle alors de l'ordre du voisinage.

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

5	4	3	4	5
4	2	1	2	4
3	1	s	1	3
4	2	1	2	4
5	4	3	4	5

FIGURE 2.12 – Ordre (de grandeur) du voisinage pour un site s d'un champ de Markov. Le voisinage d'ordre n du site s est composé des sites dont la valeur associée est inférieure ou égale à n .

Des sous-ensembles de sites appelés cliques sont définis dans ce voisinage. Ils représentent les supports pour lesquels des motifs caractéristiques des textures seront par la suite recherchés. Ce sont les probabilités d'apparition de ces motifs qui seront ensuite utilisées afin d'exprimer la probabilité conditionnelle de réalisation de chacun des sites (voir figure 2.13).

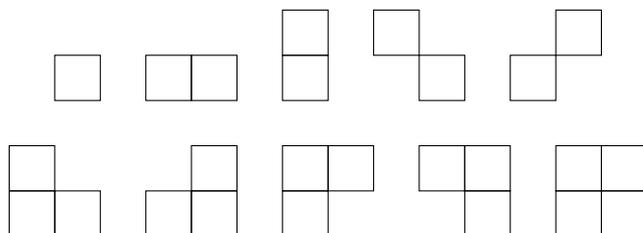


FIGURE 2.13 – Cliques de second ordre utilisées lors de la modélisation d'un phénomène par champs de Markov.

Chaque pixel de l'image étant vu comme une variable aléatoire, le fait qu'un pixel soit d'un niveau de gris donné peut être vu comme une réalisation de la variable associée. Chaque variable $X_s \in S$ prend donc ses valeurs dans un ensemble d'états Ω correspondant à l'ensemble des niveaux de gris. Dans le cas des images, nous avons généralement $|\Omega| = 256$. La probabilité de réalisation $x_s \in \Omega$ d'une variable aléatoire $X_s \in S$ se note alors $P(X_s = x_s)$.

La modélisation d'une texture par champs de Markov revient à exprimer la probabilité conditionnelle de réalisation de chaque site par rapport à son voisinage comme une loi de probabilité f dont la formulation aura préalablement été définie :

$$P(X_s = x_s | V_s) = f(X_s, x_s, V_s) \quad \forall X_s \in S, x_s \in \Omega \quad . \quad (2.11)$$

Cette modélisation s'appuie pour cela sur le théorème d'Hammersley-Clifford, qui établit que tout champs de Markov peut être associé à un champ de Gibbs [Besag, 1974]. À ce titre, la probabilité conditionnelle de réalisation d'un site s étant donné son voisinage V_s peut s'exprimer par la formule suivante :

$$P(X_s = x_s | V_s) = \frac{\exp(-U(X_s, x_s, V_s))}{\sum_{x \in \Omega} \exp(-U(X_s, x, V_s))} \quad , \quad (2.12)$$

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

avec $U(X_s, x, V_s)$ une fonction qui mesure l'énergie liée à la réalisation x d'un site X_s par rapport à son voisinage V_s . Cette énergie s'exprime généralement à l'aide d'une fonction δ_c qui mesure le potentiel de chaque clique c présente dans le voisinage V_s à caractériser la réalisation x de X_s :

$$U(X_s, x, V_s) = \sum_{c \in \text{Cliques}(V_s)} \delta_c(X_s, x) \quad . \quad (2.13)$$

Parmi les différentes méthodes de modélisation d'une texture par champs de Markov, nous pouvons citer le modèle de Derin-Elliott [Derin et Elliott, 1987] ainsi que le modèle auto-binomial [Besag, 1974]. Ces deux méthodes expriment de façon similaire le potentiel des cliques présentes dans le voisinage d'un site. Elles se basent sur la formulation suivante :

$$U(X_s, x, V_s) = \theta_1 \cdot w_1(X_s, x) + \theta_2 \cdot w_2(X_s, x) + \theta_3 \cdot w_3(X_s, x) + \theta_4 \cdot w_4(X_s, x) \quad , \quad (2.14)$$

le rôle de la fonction $w_r(X_s, x)$ étant de lier la présence de motifs (par le biais de cliques) à la probabilité de réalisation x du site X_s . Nous avons ainsi :

$$w_r(X_s, x) = \begin{cases} I(x, x_{s+r}) + I(x, x_{s-r}) & \text{pour le modèle de Durin-Elliott} \\ x \cdot (x_{s-r} + x_{s+r}) & \text{pour le modèle auto-binomial} \end{cases} \quad \forall r \in [1; 4] \quad , \quad (2.15)$$

avec :

$$I(a, b) = \begin{cases} -1 & \text{si } a = b \\ 1 & \text{sinon} \end{cases} \quad . \quad (2.16)$$

Les variables $\theta_1, \theta_2, \theta_3, \theta_4$ constituent les paramètres du modèle. Les probabilités $P(X_s = x_s | V_s)$ étant connues, la modélisation d'une texture par champs de Markov revient à estimer ces paramètres. Plusieurs méthodes permettent leur estimation, telles que les moindres-carrés [Zou *et al.*, 1995], le maximum de vraisemblance [Descombes *et al.*, 1997], le maximum *a posteriori* [Wu et Noé, 2010] ou encore le recuit simulé [Jeng et Woods, 1990]. Ces paramètres constituent une signature de la texture sur laquelle des algorithmes de classification ou de segmentation pourront se baser.

2.1.6 Bilan

Après cet état de l'art sur les méthodes de caractérisation de textures, il convient de se demander quel descripteur choisir. La littérature, relativement abondante à ce sujet, expose de nombreux avis contradictoires qui ne permettent pas d'identifier une méthode universelle.

D'après la littérature, les méthodes à base de champs de Markov semblent être d'assez mauvais candidats [Ohanian et Dubes, 1992, Clausi, 2001], tandis que les bonnes performances des caractéristiques d'Haralick sont régulièrement soulignées [Clausi, 2001, Howarth et Stefan, 2004]. Dans [Tang, 1998], les caractéristiques d'Haralick sont notamment reconnues comme plus performantes que les méthodes à base d'ondelettes en ce qui concerne la classification de textures. La conclusion inverse est cependant formulée dans [Bharati *et al.*, 2004]. Comparés à la transformée de Fourier, les filtres de Gabor se révèlent être plus

2.1. ÉTAT DE L'ART DES MÉTHODES DE CARACTÉRISATION DE TEXTURES

performants lorsqu'appliqués à une problématique d'indexation et de recherche d'images [Ursani *et al.*, 2007]. Néanmoins, la tendance est inversée en présence d'images bruitées. Les avis s'opposent concernant les performances des méthodes empruntées au traitement du signal [Clausi, 2001, Pichler *et al.*, 1996, Chen et Chen, 1999], mais s'accordent autour de l'importante complexité algorithmique de celles à base de bancs de filtres [Randen et Husoy, 1999, Manjunath et Ma, 1996b, Chen et Chen, 1999].

Dans [Stachowiak *et al.*, 2005], une étude visant à évaluer les performances de différents descripteurs de textures (caractéristiques d'Haralick, motifs binaires locaux, transformées en ondelettes et champs de Markov) sur différents types de textures est menée. Dans un contexte de classification, chaque descripteur se révèle être performant sur un type de texture, mais aucun ne l'est sur l'ensemble du panel de test. Les caractéristiques d'Haralick et la transformée en ondelettes offrent cependant une bonne complémentarité et leur combinaison permet d'assurer une indépendance vis à vis du contenu des images. Ces conclusions confirment celles déjà formulées dans le passé [Randen et Husoy, 1999, Sharma et Singh, 2001] : il n'existe pas de méthode de caractérisation de textures universelle et la combinaison de plusieurs descripteurs est nécessaire afin de permettre le traitement de tous types d'images.

La littérature s'accorde enfin sur un point, l'existence de deux types de textures : les microtextures, composées de motifs dont l'échelle est de l'ordre de quelques pixels, et les macrotextures, généralement composées de motifs plus larges [Davis *et al.*, 1979, Tuceryan et Jain, 1993, Unser, 1995, Wang *et al.*, 2013]. Les méthodes de caractérisation de textures qui reposent sur des supports de petite taille (motifs binaires locaux, champs de Markov) sont généralement reconnues comme peu efficaces sur les macrotextures car elles ne permettent pas d'appréhender correctement l'échelle de la texture. Pour la même raison, les méthodes empruntées au traitement du signal sont souvent peu adaptées aux microtextures.

Les méthodes que nous développons dans cette thèse sont indépendantes du descripteur de textures utilisé. L'objet de nos travaux n'étant pas l'évaluation des performances de ces descripteurs, nous avons choisi de nous limiter à seulement l'un d'entre eux. Notre choix s'est porté vers les caractéristiques d'Haralick pour plusieurs raisons. Tout d'abord, il s'agit d'un descripteur dont les bonnes performances sont régulièrement évoquées dans la littérature. C'est notamment l'un des rares descripteurs capable de traiter à la fois des micro-textures et des macro-textures. Il offre en effet la possibilité de régler précisément la taille de la région sur laquelle une matrice de co-occurrences est calculée ainsi que l'échelle à laquelle nous voulons qu'une texture soit analysée. Ensuite, il s'agit d'une méthode permettant d'obtenir directement un ensemble d'attributs numériques caractérisant une texture. Nous voulions en effet éviter d'avoir recours à des algorithmes de partitionnement de données, ce qui est requis par certaines méthodes qui imposent l'utilisation de dictionnaires de signatures. Enfin, l'un de nos objectifs étant de pouvoir traiter des images 3D, nous avons besoin d'une caractéristique de textures qui soit extensible à des dimensions supérieures, ce que les caractéristiques d'Haralick permettent nativement.

2.2 Contribution

L'algorithme de segmentation par régularisation de graphe sur lequel nous nous appuyons fait partie de la classe des méthodes d'optimisation globale pour la segmentation de données [Ta, 2009] présentée section 1.7.4.

Soit $G = (V, E, w)$ un graphe de similarité, et $f_0 \in \mathcal{H}(V)$ la fonction représentant les données à segmenter. Pour rappel, cette classe de modèles s'exprime sous la forme du problème de minimisation suivant :

$$\min_{f:V \rightarrow \{0,1\}} \left\{ \mathcal{E}_2(f, f_0, \lambda) = \mathcal{R}_{w,1} + \sum_{u \in V} \lambda g(f_0)(u) f(u) \right\} , \quad (2.17)$$

avec f la fonction indicatrice de la segmentation, $\mathcal{R}_{w,1}$ le terme de régularisation, $g(\cdot)$ le terme d'attache aux données et λ le coefficient exprimant le compromis entre la précision et la régularité de la solution.

Nous nous intéressons plus particulièrement à l'implémentation basée sur le critère de Chan et Vese [Chan et Vese, 2001], présenté section 1.7.4.2, dans laquelle le terme d'attache aux données s'exprime par :

$$g(f_0) = (c_1 - f_0)^2 - (c_2 - f_0)^2 \quad , \quad (2.18)$$

avec $c_1, c_2 \in \mathbb{R}$ deux paramètres représentant la valeur moyenne de la fonction f_0 respectivement à l'intérieur et à l'extérieur de la segmentation.

Cette méthode permet nativement de tirer parti d'informations non-locales caractérisant les pixels. À l'aide d'une mesure de similarité adaptée aux données manipulées, il est en effet possible de pondérer les relations relativement à ces données. Elles peuvent ainsi influencer indirectement sur le déroulement de l'algorithme. Ta applique notamment cette approche à la segmentation d'images texturées et bruitées, où la mesure de similarité est calculée dans l'espace Euclidien des patches d'images [Tschumperlé et Brun, 2009].

Les patches d'images ne sont pas des caractéristiques de textures à proprement parler, mais une construction permettant de comparer deux pixels en tenant compte des régions au milieu desquelles ils se trouvent. Soit $I : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}^m$ une image, potentiellement multivaluée. Le patch associé à un pixel de I est composé du pixel et de son voisinage, dont la taille est préalablement fixée. Il est ainsi possible de calculer la similarité des patches associés à deux pixels $p_1 = (x_1, y_1)$ et $p_2 = (x_2, y_2)$ en calculant la distance Euclidienne entre leurs patches respectifs :

$$\mu(p_1, p_2)^2 = \sum_{x=-v}^v \sum_{y=-v}^v G_a(x, y) \cdot \|I(x_1 + x, y_1 + y) - I(x_2 + x, y_2 + y)\|_2^2 \quad , \quad (2.19)$$

avec $v \in \mathbb{N}^+$ le rayon du voisinage, et G_a un noyau Gaussien de variance a , permettant de donner plus d'importance aux pixels centraux.

Les patches d'images ont été utilisés pour de la synthèse de textures [Efros et Leung, 1999] et plus récemment dans le cadre du filtrage et de la restauration d'images [Ta *et al.*, 2007a, Gilboa et Osher, 2007, Buades *et al.*, 2005, Bougleux *et al.*, 2008, Elmoataz *et al.*,

2008]. L'ensemble de ces travaux a mis en évidence l'importance des approches non-locales, qui permettent une meilleure prise en compte de la présence de motifs et de structures complexes dans les images.

Le critère d'attache aux données de Chan et Vese sur lequel repose la méthode de Ta ne profite cependant pas de ces informations non-locales, car il travaille uniquement à partir du niveau de gris de chacun des pixels. Plus particulièrement, son rôle est de minimiser l'écart-type de niveaux de gris de chacune des régions. Il s'agit donc de statistiques d'ordre 1, ce qui le rend incapable de discriminer deux textures visuellement différentes dont les niveaux de gris moyens seraient égaux (voir figure 2.1).

Cette limitation a cependant été étudiée, et il a été démontré qu'il est possible de remplacer l'intensité du niveau de gris de chaque pixel par une caractéristique non-locale évaluant l'agencement des niveaux de gris dans l'image, telle qu'une mesure de la courbure ou encore de l'orientation [Chan et Vese, 2001].

Cependant, la caractéristique de texture utilisée est unique. Il est donc nécessaire de disposer d'un certain *a priori* concernant l'image, afin de savoir quelle caractéristique employer. Afin de palier cela, nous avons opté pour une modification du critère d'attache aux données afin de profiter d'informations de plus haut niveau, disponibles sous la forme d'un vecteur de caractéristiques.

Nous nous sommes pour cela appuyés sur une extension du modèle de Chan et Vese qui permet le traitement d'images multivaluées [Chan *et al.*, 2000]. Le critère d'attache aux données qui y est défini tient compte de l'information présente dans l'ensemble des canaux de l'image, permettant ainsi la segmentation d'images couleurs ou multi-spectrales.

Soit $f_0 : \Omega \subset \mathbb{R}^m \rightarrow [0, 1]^n$ la fonction initiale représentant l'image. La fonctionnelle d'énergie sur laquelle se base le problème de minimisation s'exprime dans le domaine continu par :

$$\begin{aligned} \mathcal{J}_{CV}(f, f_0, \lambda_1, \lambda_2) = \int_{\Omega} \|\nabla f\|_1 + \sum_{i=1}^n \int_{\Omega} \left(\lambda_{1,i} \cdot f(x)(c_{1,i} - f_{0,i}(x))^2 \right. \\ \left. + \lambda_{2,i} \cdot (1 - f(x))(c_{2,i} - f_{0,i}(x))^2 \right) dx \quad , \end{aligned} \quad (2.20)$$

avec $c_{1,i}$ et $c_{2,i}$ les variables représentant les valeurs moyennes de chacun des n canaux de l'image f_0 respectivement à l'intérieur et à l'extérieur de la segmentation, et $\lambda_{1,i}, \lambda_{2,i} \in \mathbb{R}^+$ les paramètres permettant de moduler l'importance des différents termes de l'équation.

Cette formulation sous-entend que lors de l'utilisation d'un vecteur composé de k caractéristiques, le problème de minimisation possède $2k$ paramètres. De nombreuses caractéristiques de textures étant composées d'un nombre important de valeurs, un tel paramétrage du problème n'est pas envisageable, et une réduction du nombre de paramètres est nécessaire. À l'image de l'algorithme originel de Ta, nous avons tout d'abord opté pour la conservation d'un unique paramètre λ , et ainsi n'avoir à réaliser qu'un compromis entre régularité et précision de la solution. Néanmoins, les tests que nous avons réalisés nous ont permis de constater qu'un contrôle plus fin de l'algorithme était nécessaire, au risque de ne pas obtenir de résultats pertinents pour certaines images. D'après les tests réalisés dans [Chan *et al.*, 2000], deux options nous sont offertes.

La première consiste à utiliser un unique λ_i pour chaque canal i de l'image, avec

2.2. CONTRIBUTION

$\lambda_{1,i} = \lambda_{2,i} = \lambda_i$, ce qui permet de donner plus ou moins d'importance à certains canaux de l'image. Cependant, cela nécessite de nouveau de posséder un *a priori* concernant l'importance des différents canaux, information complexe à obtenir (voire inexistante) lors de l'utilisation de caractéristiques de textures. Par ailleurs, cette approche ne simplifie pas suffisamment le paramétrage du problème, puisque qu'il reste nécessaire de fixer autant de paramètres que de canaux.

La seconde option consiste à n'utiliser que deux paramètres, λ_1 et λ_2 , avec $\lambda_{1,1} = \dots = \lambda_{1,n} = \lambda_1$ et $\lambda_{2,1} = \dots = \lambda_{2,n} = \lambda_2$. À l'image de ce qui est fait dans l'algorithme de Chan et Vese scalaire, cela permet de donner plus de poids au terme d'attache aux données associé à l'une des deux régions. Nos tests nous ont permis de vérifier que la présence de ces deux paramètres était suffisante pour corriger les problèmes de pertinence des résultats liés à l'utilisation d'un unique paramètre. La fonctionnelle de régularisation sur laquelle nous nous basons peut donc être reformulée :

$$\begin{aligned} \mathcal{J}_{CV}(f, f_0, \lambda_1, \lambda_2) = \int_{\Omega} \|\nabla f\|_1 + \sum_{i=1}^n \int_{\Omega} & \left(\lambda_1 \cdot f(x) \cdot (c_{1,i} - f_{0,i}(x))^2 \right. \\ & \left. + \lambda_2 \cdot (1 - f(x)) \cdot (c_{2,i} - f_{0,i}(x))^2 \right) dx \quad , \end{aligned} \quad (2.21)$$

Dans cette formulation vectorielle du problème, le critère de Chan et Vese offre toujours une évaluation de l'homogénéité intra-région, mais cette mesure est désormais globalement calculée sur l'ensemble des canaux de l'image.

La procédure permettant de transformer le problème de minimisation représenté par la fonctionnelle d'énergie (2.21) en un problème d'optimisation globale s'exprimant par la seconde classe de modèles variationnels est identique à celle proposée par Ta (exposée section 1.7.4.2). En réécrivant le terme d'attache aux données présent dans l'équation (2.21) :

$$\begin{aligned} & \sum_{i=1}^n \left(\lambda_1 \cdot f \cdot (c_{1,i} - f_{0,i})^2 + \lambda_2 \cdot (1 - f)(c_{2,i} - f_{0,i})^2 \right) \\ & = f \cdot \sum_{i=1}^n \left(\lambda_1 \cdot (c_{1,i} - f_{0,i})^2 - \lambda_2 \cdot (c_{2,i} - f_{0,i})^2 \right) + \lambda_2 \sum_{i=1}^n (c_{2,i} - f_{0,i})^2 \quad , \end{aligned} \quad (2.22)$$

nous pouvons mettre en évidence le dernier terme qui, étant donné que les variables $c_{1,i}$ et $c_{2,i}$ sont calculées avant chaque itération de l'algorithme, est constant et n'influe donc pas sur le processus de minimisation de la fonctionnelle (2.21). Dans ces conditions, nous pouvons formuler un problème de minimisation équivalent à partir de la fonctionnelle suivante :

$$\begin{aligned} \mathcal{J}_{CV}(f, f_0, \lambda_1, \lambda_2) = \int_{\Omega} \|\nabla f\|_1 + \sum_{i=1}^n \int_{\Omega} & f(x) \cdot \left(\lambda_1 \cdot (c_{1,i} - f_{0,i}(x))^2 \right. \\ & \left. - \lambda_2 \cdot (c_{2,i} - f_{0,i}(x))^2 \right) dx \quad . \end{aligned} \quad (2.23)$$

Cette dernière constitue un cas particulier défini dans le domaine continu de la seconde classe de modèles variationnels (1.41) ; il est donc naturel de considérer la version discrète

2.3. RÉSULTATS

définie sur les graphes de ce modèle de Chan et Vese vectoriel :

$$\min_{f:V \rightarrow \{0,1\}} \left\{ \mathcal{E}_{CV}(f, f_0, \lambda_1, \lambda_2) = \mathcal{R}_{w,1} + \sum_{u \in V} g(f_0, \lambda_1, \lambda_2)(u) f(u) \right\}, \quad (2.24)$$

avec :

$$g(f_0, \lambda_1, \lambda_2) = \sum_{i=1}^n \left(\lambda_1 \cdot (c_{1,i} - f_{0,i})^2 - \lambda_2 \cdot (c_{2,i} - f_{0,i})^2 \right). \quad (2.25)$$

Algorithme 1 : Algorithme d'approximation du modèle de Chan et Vese pour images multi-valuées.

Entrées : un graphe de similarité $G = (V, E, w)$, une fonction $f_0 \in \mathcal{H}^m(V)$ représentant les caractéristiques des pixels, $m = V \rightarrow \{0, 1\}$ l'initialisation de la segmentation, N un nombre d'itérations à réaliser, $\lambda_1, \lambda_2 \in R^+$ les paramètres du problème de minimisation.

Sorties : une fonction $f : V \rightarrow \{0, 1\}$ indicatrice de la segmentation.

pour tous les $u \in V$ **faire** $f^0(u) = m(u)$;

pour $n = 0$ **a** $N - 1$ **faire**

pour $i = 1$ **a** m **faire**

┌ $c_{1,i} =$ moyenne de $f_{0,i}$ lorsque $f^n \geq 0.5$

└ $c_{2,i} =$ moyenne de $f_{0,i}$ lorsque $f^n < 0.5$

pour tous les $(u, v) \in E$ **faire**

┌ $\beta_{uv} = \sqrt{w_{uv}} |f^n(u) - f^n(v)|^{-1}$

pour tous les $u \in V$ **faire**

$$\gamma = \frac{\sum_{v \sim u} \beta_{uv} \cdot f^n(v) - \sum_{i=1}^m \left(\lambda_1 \cdot (c_{1,i} - f_{0,i}(u))^2 - \lambda_2 \cdot (c_{2,i} - f_{0,i}(u))^2 \right)}{\sum_{v \sim u} \beta_{uv}}$$

└ $f^{n+1}(u) = \max(\min(\gamma, 1), 0)$

pour tous les $u \in V$ **faire**

┌ **si** $f^N(u) > 0.5$ **alors** $f(u) = 1$ **sinon** $f(u) = 0$;

L'algorithme 1 résume la méthode de régularisation basée sur le modèle de Chan et Vese vectoriel.

Si l'on considère une image scalaire comme une image multivaluée composée d'un unique canal, notre méthode est compatible avec celle de Ta.

2.3 Résultats

2.3.1 Segmentation d'images multispectrales

L'une des applications de l'algorithme de Chan et Vese vectoriel est l'imagerie multispectrale [Chan *et al.*, 2000]. Une image multispectrale correspond à une image dans laquelle

2.3. RÉSULTATS

une scène est acquise à l'aide de différents capteurs, chacun étant conçu pour quantifier une propriété particulière. L'image est alors composée de plusieurs canaux.

La technique d'imagerie multispectrale la plus commune est la photographie couleur, pour laquelle chaque canal de l'image caractérise une gamme de longueurs d'ondes à laquelle la vision humaine est sensible. De nombreux appareils de télémétrie tels que les satellites de télédétection ou encore les télescopes spatiaux sont capables d'acquérir des images à des longueurs d'ondes auxquelles l'œil humain n'est pas sensible, comme l'infrarouge ou l'ultraviolet. Certaines technologies d'imagerie médicale peuvent aussi caractériser simultanément plusieurs propriétés physiques. La microscopie à fluorescence permet par exemple d'exciter certains composés chimiques afin de leur faire émettre des photons et ainsi révéler leur présence.

L'intérêt du critère vectoriel d'attache aux données de Chan et Vese est de pouvoir combiner l'information présente dans chacun des canaux et ainsi de permettre le traitement de l'image malgré la présence de bruit ou d'occlusions dans certains d'entre eux.

Afin d'illustrer la capacité de notre méthode à traiter ce type d'image ainsi que le rôle des différents paramètres, nous allons dans un premier temps traiter l'image artificielle présentée figure 2.14.

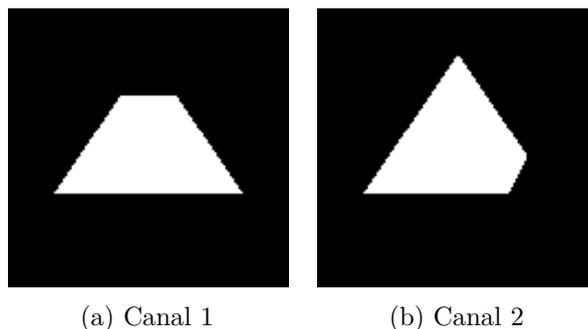


FIGURE 2.14 – Exemple d'image multispectrale. Les occlusions présentes dans chacun des canaux ne permettent pas de distinguer l'objet dans sa totalité.

Le critère d'attache aux données que nous employons (voir équation (2.25)) peut être reformulé ainsi :

$$\begin{aligned} g(f_0, \lambda_1, \lambda_2) &= \lambda_1 \cdot \sum_{i=1}^2 (c_{1,i} - f_{0,i})^2 - \lambda_2 \cdot \sum_{i=1}^2 (c_{2,i} - f_{0,i})^2 \quad , \\ &= \lambda_1 \cdot d(c_1, f_0)^2 - \lambda_2 \cdot d(c_2, f_0)^2 \quad , \end{aligned} \quad (2.26)$$

avec f_0 le vecteur de caractéristiques associé à un nœud, c_1 et c_2 les vecteurs composés des valeurs moyennes des différents canaux de l'image pour chacune des deux régions identifiées par la segmentation et $d()$ l'opérateur distance euclidienne. Il s'agit donc d'une mise en opposition des distances entre les caractéristiques d'un nœud et les caractéristiques moyennes des deux régions. Pour un nœud particulier, lorsque ce critère prend une valeur négative (respectivement positive), cela signifie que du point de vue du critère d'attache aux données, ce nœud appartient à la région intérieure (respectivement extérieure).

2.3. RÉSULTATS

Les paramètres λ_1 et λ_2 permettent de pondérer cette mise en opposition en autorisant une des deux régions à présenter un écart à la moyenne plus important. Ainsi, si $\lambda_1 > \lambda_2$ (respectivement $\lambda_1 < \lambda_2$), l'homogénéité, relativement à la distance euclidienne, de la région intérieure (respectivement extérieure) sera privilégiée.

Dans l'algorithme 1, le terme d'attache aux données contrebalance le terme de régularisation. Ce dernier n'étant associé à aucun paramètre permettant de le pondérer, ce sont les valeurs de λ_1 et λ_2 qui jouent ce rôle : plus λ_1 et λ_2 seront faibles, plus la régularité de la solution sera privilégiée.

Afin de traiter l'image 2.14, nous avons construit un graphe grille en 4-voisinage. Aucune mesure de similarité n'a été utilisée : la valeur constante 1 a été affectée à l'ensemble des arêtes du graphe pour permettre au processus de diffusion de s'exécuter. Dans chacun de nos tests, l'algorithme sera initialisé dans l'état présenté figure 2.15a.

Afin d'identifier le triangle complet, il convient de régler les paramètres λ_1 et λ_2 de façon adéquate. Pour cette image, cette tâche se révèle être relativement simple. En effet, nous savons que dans la segmentation que nous souhaitons obtenir, la région extérieure ne devra être composée que de pixels noirs, tandis que – du fait de la présence d'occlusions – la région représentant le triangle devra contenir à la fois des pixels de faible et de forte intensité. L'homogénéité de la région extérieure doit donc être privilégiée : λ_2 devra avoir une valeur plus importante que λ_1 .

Concernant la régularité de la solution, la forme que nous cherchons à identifier présentant des angles francs, il est là encore évident que l'attache aux données devra être privilégiée. λ_1 et λ_2 ne devront donc pas être fixés à des valeurs trop faibles.

La segmentation du triangle est accomplie avec succès en fixant $\lambda_1 = 1$ et $\lambda_2 = 2$ (voir figure 2.15b). À l'inverse, en privilégiant l'attache aux données de la région intérieure ($\lambda_1 = 2$, $\lambda_2 = 1$), seule la région du triangle commune aux deux canaux sera identifiée (voir figure 2.15c). Si en revanche nous choisissons de ne pas privilégier une région par rapport à une autre, le terme de régularisation se charge d'établir un compromis entre la régularité et la précision de la solution (voir figure 2.15d). L'obtention d'un résultat stable nécessite alors un plus grand nombre d'itérations car la régularisation s'appuie sur un principe de diffusion. Enfin, si nous accordons moins d'importance au terme d'attache aux données, les contours du triangle perdent en précision et gagnent en régularité (voir figures 2.15e et 2.15f).

2.3.2 Segmentation d'images texturées artificielles

Dans ce second test, nous avons souhaité mettre en évidence la capacité de notre méthode à segmenter des images à partir de caractéristiques de textures. Le principe est similaire au traitement des images multispectrales, à la différence que chaque canal représente désormais une caractéristique de texture. Afin d'illustrer cela, nous nous intéressons à l'image de test présentée figure 2.16a.

Cette image résulte de l'addition d'un bruit blanc Gaussien d'écart-type $\sigma = 48$ à l'image 2.16b. Cette dernière est composée de deux motifs qui ne diffèrent que par leur orientation : hachures orientées à -45° pour l'arrière-plan et à $+45^\circ$ pour l'objet central. Ce bruit étant appliqué de façon homogène à toute l'image, la distribution des niveaux de gris

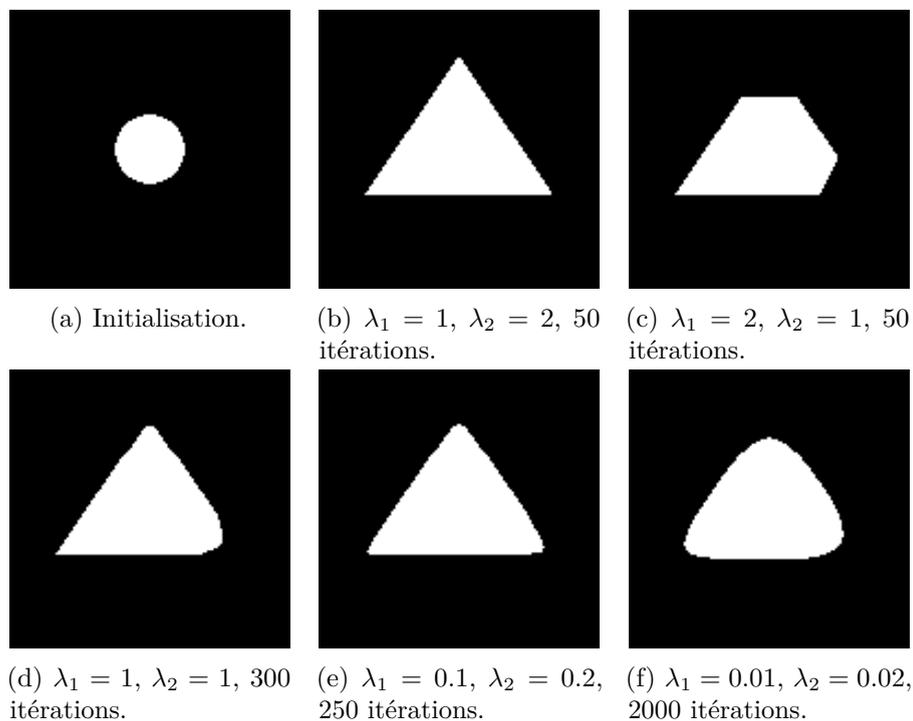


FIGURE 2.15 – (a) : initialisation de l’algorithme. (b, c, d, e, f) : segmentations obtenues après stabilisation du résultat.

dans chacune des deux régions reste homogène.

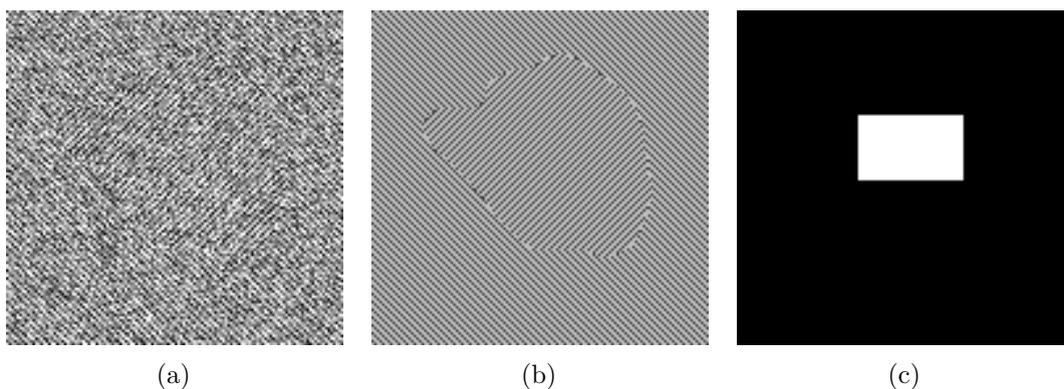


FIGURE 2.16 – Exemple d’image texturée artificielle. L’image (a) résulte de l’addition d’un bruit blanc Gaussien d’écart-type $\sigma = 48$ à l’image (b). Les statistiques de premier ordre concernant la distribution des niveaux de gris de l’objet central et de l’arrière-plan sont identiques. L’image (c) correspond à la segmentation initiale qui servira de point de départ de l’algorithme dans chacun des tests.

En choisissant de travailler avec des textures dont les statistiques de premier ordre sont identiques, nous excluons tout algorithme qui ne se baserait que sur une analyse brute

de la distribution des niveaux de gris. Ceux basés sur le critère d'attache aux données de Chan et Vese (dans sa version scalaire) sont les premiers visés car incapables de différencier les deux régions à partir de leurs niveaux de gris moyens respectifs. Il est donc nécessaire d'avoir recours à des caractéristiques de textures.

Les caractéristiques d'Haralick [Haralick *et al.*, 1973] ont été calculées en chaque pixel de cette image dans une région de 11 pixels de diamètre et selon le décalage (1, 1), l'image ayant été préalablement réduite à 16 niveaux de gris. Ce décalage a été choisi car il permet de cibler une propriété évidente de la texture : son orientation. Le décalage correspond en effet à l'orientation du motif de la région intérieure. Comme illustré figure 2.17, la matrice de co-occurrences associée à un pixel de cette région est majoritairement composée de co-occurrences de niveaux de gris similaires, ce qui se traduit par la présence de fortes valeurs à proximité de sa première diagonale. La matrice associée à un pixel de la région extérieure présente en revanche un nombre important de co-occurrences de niveaux de gris d'intensité différentes, ce qui contribue à la présence de valeurs importantes à proximité de la seconde diagonale. Ces deux structures de matrices sont reflétées par plusieurs caractéristiques d'Haralick, nous permettant de différencier les deux régions de l'image.

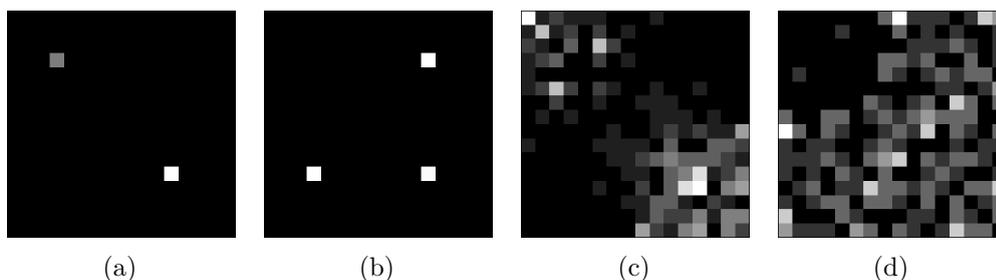


FIGURE 2.17 – Structures typiques des matrices de co-occurrences des deux régions de l'image texturée artificielle. (a) : région intérieure de l'image non bruitée. (b) : région extérieure de l'image non bruitée. (c) : région intérieure de l'image bruitée. (d) : région extérieure de l'image bruitée. La région intérieure est essentiellement composée de co-occurrences d'intensités similaires, tandis que la région extérieure présente un nombre important de co-occurrences d'intensités différentes.

Une représentation des caractéristiques d'Haralick normalisées correspondant à cette image est présentée figure 2.18.

Influence des paramètres λ_1 et λ_2 sur la qualité du résultat

Dans un premier temps, nous avons souhaité mettre en évidence le rôle important des deux paramètres λ_1 et λ_2 . Nous avons donc segmenté l'image 2.16a pour différentes valeurs de ces paramètres tout en suivant l'évolution du résultat au cours de l'exécution du processus de régularisation. Les valeurs de λ_1 et λ_2 considérées ont été prises dans l'ensemble $\{0.5, 0.75, \dots, 3\}$, lequel échantillonne un intervalle raisonnable de valeurs de compromis. Le graphe utilisé pour représenter l'image correspond à un 4-voisinage, et la constante 1 fait office de mesure de similarité.

Les tableaux 2.1 et 2.2 résument les résultats obtenus après convergence de l'algorithme

2.3. RÉSULTATS

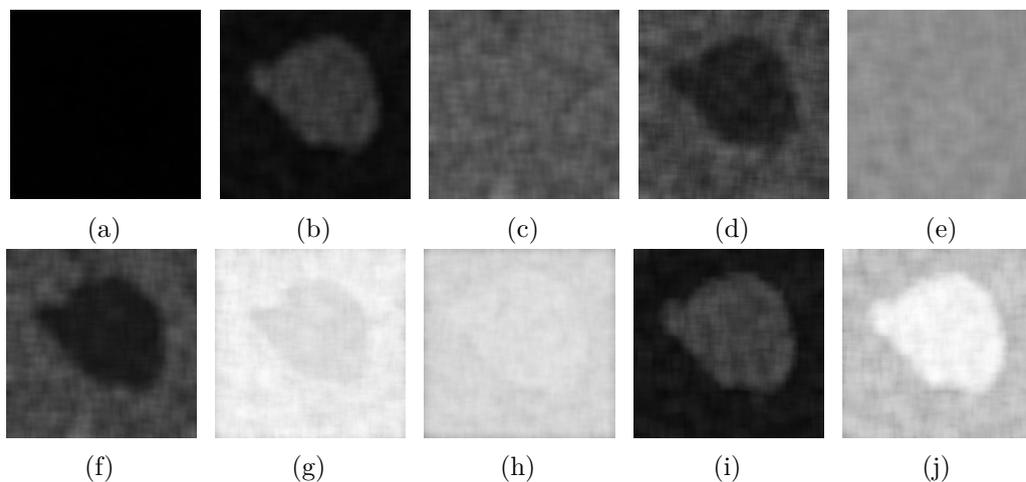


FIGURE 2.18 – Représentation normalisée des caractéristiques d’Haralick obtenues pour chaque pixel de l’image 2.16a. Paramètres de calcul : nombre de niveaux de gris = 16, diamètre de la fenêtre = 11 pixels, décalage = (1, 1).

pour l’ensemble des configurations. L’axe horizontal représente la valeur attribuée à λ_1 et l’axe vertical celle attribuée à λ_2 . Chaque tableau présente une mesure d’erreur qui évalue la qualité de la segmentation calculée par l’algorithme vis à vis de la vérité terrain.

La mesure d’erreur utilisée dans le premier tableau est la distance générique d’anomalie, qui est définie pour les problèmes de segmentation binaire comme le nombre de pixels mal classés [Cardoso et Corte-Real, 2005]. Elle est ici exprimée sous forme normalisée, et représente alors le pourcentage de pixels mal classés dans l’image.

Le second tableau est quant à lui dédié à une mesure d’erreur basée sur la distance de Hausdorff. Celle-ci évalue la distance maximum entre le contour de l’objet détecté par l’algorithme et celui défini par la vérité terrain. Cette mesure d’erreur peut être vue comme une évaluation de la plus forte erreur commise par l’algorithme de segmentation, c’est à dire la distance à laquelle se trouve le pixel mal classé le plus éloigné du contour. Cette mesure d’erreur est relativement sensible, car il suffit d’un unique pixel mal classé pour attribuer une erreur importante au résultat.

Ces deux tableaux nous montrent que pour cette image, les bons résultats sont obtenus lorsque le rapport λ_1/λ_2 est légèrement supérieur à 1. Cela signifie que le critère d’attache aux données autorise une plus grande variabilité en ce qui concerne les textures de la région extérieure, privilégiant ainsi l’homogénéité de la région intérieure.

En revanche, nous pouvons voir que, selon la distance de Hausdorff, les meilleures solutions sont obtenues lorsque le rapport λ_1/λ_2 est légèrement plus élevé que celui des meilleurs solutions selon la distance générique d’anomalie.

Ce comportement est à attribuer au processus de régularisation, dont l’une des conséquences est de lisser le contour des régions. Or, lorsqu’un contour présente une forte courbure, sa régularisation entraîne une augmentation de l’écart avec le contour défini par les données (voir figure 2.19). L’effet est d’autant plus prononcé que la courbure est importante. Cet écart correspond précisément à ce que mesure la distance de Hausdorff.

2.3. RÉSULTATS

$\lambda_2 \downarrow / \lambda_1 \rightarrow$	0,50	0,75	1,00	1,25	1,50	1,75	2,00	2,25	2,50	2,75	3,00
0,50	2,37	4,14	5,22	6,07	6,81	7,39	7,98	8,44	8,86	9,36	9,74
0,75	1,84	1,52	2,34	3,09	3,69	4,16	4,57	4,89	5,29	5,59	5,85
1,00	3,35	1,54	1,23	1,49	2,11	2,47	2,94	3,20	3,43	3,83	3,97
1,25	4,71	2,40	1,46	1,22	1,29	1,45	1,82	2,16	2,36	2,61	2,86
1,50	5,87	3,25	1,97	1,41	1,16	1,09	1,25	1,40	1,62	1,86	2,12
1,75	6,79	4,00	2,66	1,74	1,36	1,15	1,07	1,11	1,20	1,34	1,53
2,00	7,88	4,69	3,14	2,26	1,62	1,35	1,12	1,06	1,06	1,12	1,17
2,25	9,05	5,22	3,70	2,69	2,00	1,57	1,36	1,13	1,04	1,01	1,06
2,50	10,37	5,82	4,15	3,05	2,41	1,83	1,53	1,34	1,14	1,02	0,99
2,75	12,62	6,23	4,52	3,47	2,73	2,22	1,73	1,53	1,34	1,14	1,06
3,00	18,98	6,75	4,91	3,80	3,03	2,48	2,07	1,68	1,46	1,34	1,15

TABLE 2.1 – Évaluation de l’influence des paramètres λ_1 et λ_2 . Les valeurs exposées correspondent à l’erreur mesurée selon la distance générique d’anomalie normalisée pour l’ensemble des combinaisons testées. Ces résultats sont calculés à partir des segmentations obtenues après convergence de l’algorithme.

$\lambda_2 \downarrow / \lambda_1 \rightarrow$	0,50	0,75	1,00	1,25	1,50	1,75	2,00	2,25	2,50	2,75	3,00
0,50	6,3	8,2	9,5	10,8	11,7	12,4	13,0	13,6	14,6	14,9	15,5
0,75	5,4	4,2	5,4	6,1	6,7	7,3	7,8	8,2	8,6	9,2	9,5
1,00	6,3	5,0	4,1	4,0	4,5	5,0	5,4	5,7	5,8	6,4	6,4
1,25	7,3	5,4	5,0	4,1	3,6	3,6	4,0	4,2	4,2	4,5	5,0
1,50	7,6	6,1	5,4	5,0	4,1	3,6	3,6	3,2	3,6	3,6	4,0
1,75	8,1	6,3	5,4	5,0	4,5	4,1	3,6	3,6	3,6	3,0	3,0
2,00	8,9	6,3	6,1	5,4	5,0	4,5	4,1	3,6	3,6	3,6	3,2
2,25	9,0	7,3	6,1	5,4	5,0	5,0	4,1	4,1	3,6	3,6	3,6
2,50	10,8	7,3	6,3	5,4	5,4	5,0	4,5	4,1	4,1	3,6	3,6
2,75	14,8	7,6	6,3	6,1	5,4	5,0	5,0	4,5	4,1	4,1	3,6
3,00	25,0	7,6	6,3	6,1	5,4	5,4	5,0	5,0	4,1	4,1	4,1

TABLE 2.2 – Évaluation de l’influence des paramètres λ_1 et λ_2 . Les valeurs exposées correspondent à l’erreur mesurée selon la distance de Hausdorff pour l’ensemble des combinaisons testées. Ces résultats sont calculés à partir des segmentations obtenues après convergence de l’algorithme.

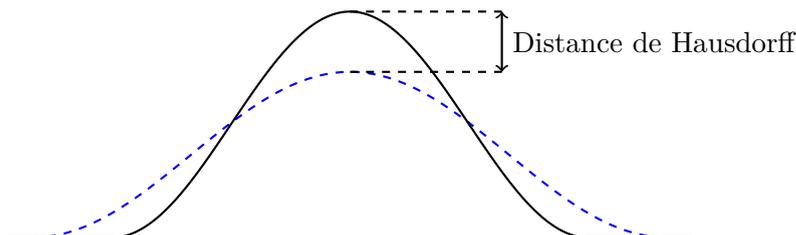


FIGURE 2.19 – Régularisation d’un contour de forte courbure. La courbe noire représente le contour initial (tel que défini par les données) et la courbe bleue en pointillés sa version régularisée.

2.3. RÉSULTATS

Nous pouvons illustrer cela sur la région supérieure gauche de l'image 2.16, dont la vérité terrain est représentée figure 2.20a. Les figures 2.20b et 2.20c représentent les meilleures solutions obtenues pour $\lambda_1 = 3$ (dernières colonnes des tableaux de résultats) avec respectivement la distance générique d'anomalie ($\lambda_2 = 2.5$) et la distance de Hausdorff ($\lambda_2 = 1.75$). Les figures 2.20d et 2.20e représentent les différences entre chacune des solutions et la vérité terrain. Les pixels noirs correspondent à ceux qui ont été ajoutés par la régularisation, et les blancs à ceux qui ont été retirés.

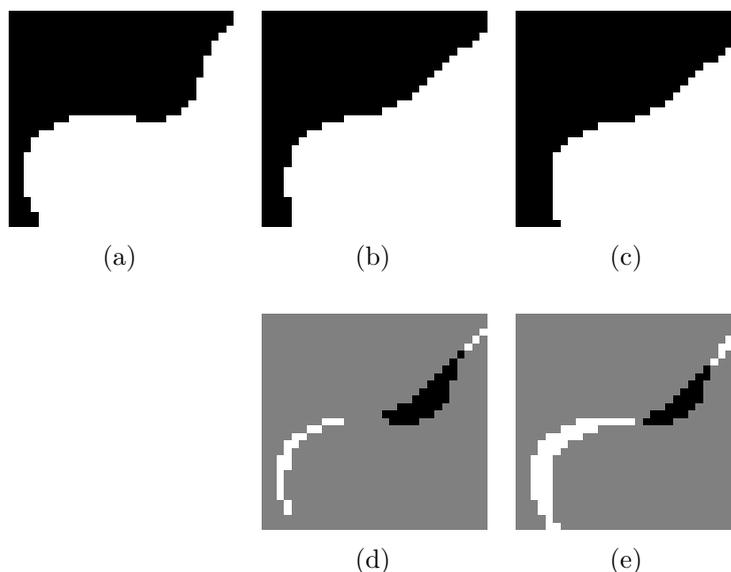


FIGURE 2.20 – Illustration de l'influence de la régularisation sur les contours de forte courbure. Les images présentées ici se focalisent sur la région supérieure gauche de l'image 2.16. La figure (a) représente la vérité terrain et les figures (b) et (c) les segmentations obtenues par notre méthode pour $\lambda_1 = 3$ et respectivement $\lambda_2 = 2.5$ et $\lambda_2 = 1.75$. Les images (d) et (e) correspondent à la différence entre la solution calculée et la vérité terrain : les pixels noirs représentent ceux qui ont été supprimés, et les blancs ceux qui ont été ajoutés.

Lorsque $\lambda_2 = 2.5$, le rapport λ_1/λ_2 est proche de 1, ce qui se traduit par une régularisation relativement importante. Cette solution étant la meilleure du point de vue de la distance générique d'anomalie, les pixels sont globalement bien classés, exceptés dans la région concave (les pixels en noir dans la figure 2.20d) qui subit une forte régularisation.

Lorsque $\lambda_2 = 1.75$, nous autorisons l'algorithme à mal classer un plus grand nombre de pixels de l'arrière-plan (les pixels blancs dans l'image 2.20e) et à moins régulariser la solution. Comparé à la solution précédente, cela se traduit par une augmentation globale du nombre de pixels mal classés, sauf dans la région concave où des pixels avaient été ajoutés.

Dans ces deux configurations, les pixels les plus éloignés du vrai contour se trouvent précisément dans cette région concave. Les paramètres λ_1 et λ_2 jouent donc un rôle important en ce qui concerne la qualité du résultat. Cette notion de qualité dépend cependant du critère d'évaluation, et un même jeu de paramètres ne permettra pas nécessairement de satisfaire plusieurs critères.

Influence des paramètres λ_1 et λ_2 sur la vitesse de convergence

Nous pouvons aussi noter que ces paramètres influent sur la vitesse de convergence de l'algorithme. Les valeurs contenues dans le tableau 2.3 nous indiquent que la méthode converge d'autant plus rapidement que les paramètres sont fixés à des valeurs fortes.

$\lambda_2 \downarrow / \lambda_1 \rightarrow$	1	1,25	1,5	1,75	2	2,25	2,5	2,75	3
1	1200	1200	1200	1200	1200	1200	1200	1200	1200
1,25	1200	1000	1000	1000	1000	1000	1000	1000	1000
1,5	1200	1000	900	900	900	900	900	900	900
1,75	1200	1000	900	900	900	900	900	900	900
2	1200	1000	900	900	700	700	700	700	700
2,25	1200	1000	900	900	700	700	700	700	700
2,5	1200	1000	900	900	700	700	600	600	600
2,75	1200	1000	900	900	700	700	600	600	600
3	1200	1000	900	900	700	700	600	600	600

TABLE 2.3 – Influence des paramètres λ_1 et λ_2 sur la vitesse de convergence. Les valeurs indiquées représentent le nombre d'itérations nécessaires à l'obtention d'une solution stable.

Le terme d'attache aux données réalise une classification qui se base sur des critères d'homogénéité. Ceux-ci sont calculés de façon globale par rapport à l'image. En une itération de l'algorithme, il est donc possible d'attribuer une classe à chacun des nœuds du graphe. La régularisation est en revanche réalisée par un processus de diffusion dont le principe est d'attribuer à chaque nœud une valeur indiquant le potentiel d'appartenance à chacune des deux classes. Ce potentiel est calculé en un nœud comme étant la moyenne du potentiel de chacun des nœuds voisins. La régularisation se déroule ainsi de proche en proche : il est donc nécessaire d'effectuer un certain nombre d'itérations afin de permettre au potentiel d'un nœud d'influer sur celui d'un autre nœud qui ne serait pas un voisin direct.

Afin d'illustrer la raison pour laquelle cette méthode nécessite un plus grand nombre d'itérations lorsque l'un des paramètres est de faible valeur, considérons $G = (V, E, w)$ un graphe de similarité, $f_0 \in \mathcal{H}(V)$ une fonction caractérisant les nœuds et u et v deux nœuds dont les caractéristiques sont typiques respectivement des régions 1 et 2 ($f_0(u) = c_1$ et $f_0(v) = c_2$). Le critère d'attache aux données :

$$g(f_0, \lambda_1, \lambda_2) = \lambda_1(c_1 - f_0)^2 - \lambda_2(c_2 - f_0)^2 \quad , \quad (2.27)$$

vaut alors pour chacun de ces nœuds :

$$g(f_0(u), \lambda_1, \lambda_2) = -\lambda_2(c_2 - c_1)^2 \quad , \quad (2.28)$$

$$g(f_0(v), \lambda_1, \lambda_2) = \lambda_1(c_1 - c_2)^2 \quad . \quad (2.29)$$

Ces formules nous indiquent que si $\lambda_1 < \lambda_2$, alors le critère d'attache aux données aura plus de poids pour le nœud u que pour le nœud v . Les nœuds qui, du point de vue des données, appartiennent à la première classe subissent ainsi plus fortement les effets de la régularisation que ceux qui appartiennent à la seconde. Un plus grand nombre d'itérations est alors nécessaire afin de stabiliser le potentiel de ces nœuds.

2.3. RÉSULTATS

Influence de la pertinence des caractéristiques

Nous pouvons voir dans la figure 2.18 que parmi les dix caractéristiques d'Haralick dont nous disposons, seulement la moitié d'entre elles contient une information utile à la segmentation de l'image. Or, nous savons que le critère de Chan et Vese considère équitablement chacune des caractéristiques de textures. Il est donc envisageable que des caractéristiques non pertinentes ou encore bruitées perturbent l'identification des textures et rendent la méthode moins précise.

Nous avons donc souhaité évaluer la capacité de notre méthode à traiter des données non pertinentes. Nous avons pour cela réduit à cinq le nombre de caractéristiques en ne conservant que celles qui, visuellement, sont les plus pertinentes. Elles correspondent à celles labellisées respectivement (b), (d), (f), (i) et (j) dans la figure 2.18. Le même test que précédemment a ensuite été réalisé à partir de ces données.

Le tableau 2.4 présente les résultats obtenus, évalués selon la distance de générique d'anomalie.

$\lambda_2 \downarrow / \lambda_1 \rightarrow$	0,50	0,75	1,00	1,25	1,50	1,75	2,00	2,25	2,50	2,75	3,00
0,50	1,26	2,03	2,70	3,30	3,84	4,16	4,55	4,86	5,08	5,33	5,62
0,75	1,97	1,19	1,25	1,53	1,93	2,31	2,61	2,85	3,11	3,36	3,52
1,00	3,06	1,65	1,14	1,04	1,15	1,37	1,57	1,80	2,08	2,25	2,46
1,25	4,01	2,38	1,56	1,15	1,03	1,01	1,13	1,20	1,40	1,53	1,70
1,50	4,72	2,94	2,06	1,50	1,18	1,04	0,99	1,02	1,10	1,14	1,29
1,75	5,29	3,52	2,54	1,83	1,44	1,21	1,07	1,00	0,96	0,99	1,07
2,00	5,87	3,97	2,93	2,25	1,73	1,45	1,26	1,10	1,04	0,99	0,99
2,25	6,30	4,38	3,35	2,62	2,09	1,68	1,44	1,26	1,15	1,04	0,99
2,50	6,81	4,73	3,58	2,92	2,42	1,98	1,68	1,44	1,26	1,15	1,10
2,75	7,18	5,04	3,89	3,20	2,67	2,21	1,84	1,66	1,43	1,28	1,17
3,00	7,50	5,27	4,17	3,45	2,94	2,51	2,10	1,81	1,61	1,45	1,29

TABLE 2.4 – Évaluation de la capacité à gérer des caractéristiques de faible pertinence. Les valeurs exposées correspondent à l'erreur mesurée selon la distance générique d'anomalie pour l'ensemble des combinaisons testées. Ces résultats sont calculés à partir des segmentations obtenues après convergence de l'algorithme.

Ces résultats nous montrent une augmentation sensible de la qualité de segmentation pour l'ensemble des jeux de paramètres. À titre indicatif, 30 résultats sont désormais sous le seuil de 1.20, contre seulement 21 lorsque l'ensemble des caractéristiques était utilisé.

Comme l'indique les résultats présentés dans le tableau 2.5, la vitesse de convergence est, elle aussi, améliorée : le nombre d'itérations est globalement réduit de 40 %.

Une étape de sélection des caractéristiques les plus pertinentes est donc nécessaire afin d'optimiser la modélisation du problème et permettre une meilleure identification des textures. Nous reviendrons dans la conclusion de ce chapitre sur une méthode permettant de la rendre automatique.

2.3. RÉSULTATS

$\lambda_2 \downarrow / \lambda_1 \rightarrow$	1	1,25	1,5	1,75	2	2,25	2,5	2,75	3
1	800	700	700	700	700	700	700	700	700
1,25	800	700	700	700	700	700	700	700	700
1,5	700	600	600	500	500	500	500	500	500
1,75	700	600	600	500	500	500	500	500	500
2	700	600	600	500	500	500	500	500	500
2,25	700	600	600	500	500	400	400	400	400
2,5	700	600	600	500	500	400	400	400	400
2,75	700	600	600	500	500	400	400	400	400
3	700	600	600	500	500	400	400	400	400

TABLE 2.5 – Évaluation de la capacité à gérer des caractéristiques de faible pertinence. Les valeurs indiquées représentent le nombre d’itérations nécessaires à l’obtention d’une solution stable.

2.3.3 Segmentation d’images échographiques 2D et 3D

Le dernier test concerne le traitement d’images échographiques. L’échographie est une technique d’imagerie médicale réputée pour produire des images particulièrement bruitées.

Ce bruit est en réalité ce que l’on appelle un motif d’interférence. Le fonctionnement d’un échographe est basé sur l’émission d’un signal sonore et l’analyse de ses échos. L’amplitude de ces échos est représentatrice de l’échogénicité des milieux que le signal traverse, c’est à dire leur capacité à renvoyer un écho. Chaque tissu ou organe possède un niveau d’échogénicité qui lui est propre, ce qui permet de les distinguer. En fonction des différents éléments que rencontre l’onde sur son passage, plusieurs échos peuvent se superposer et interférer, produisant ainsi des pics de fortes ou de faibles amplitudes, propres à la composition du milieu. Ils ne représentent donc pas un bruit mais forment des motifs dont l’analyse permet une meilleure caractérisation des différents milieux. Ces images sont donc de parfaites candidates à une analyse basée texture.

Nous avons validé notre méthode sur un jeu de six images échographiques. Ces acquisitions ont été réalisées sur l’épiderme et permettent l’analyse de névus, plus communément appelés grains de beauté. Nous disposons pour chaque image d’une vérité terrain, définie par un expert, qui nous permet d’évaluer la précision des segmentations produites.

Chaque image est représentée par un graphe de similarité. Sa structure correspond à un 4-voisinage, et la constante 1 a été utilisée en tant que mesure de similarité. Les caractéristiques d’Haralick ont été calculées en réduisant le nombre de niveaux de gris à 16, pour une fenêtre de diamètre 11 et pour le décalage (1, 0). Seules les trois caractéristiques d’Haralick les plus pertinentes ont été conservées.

Nous avons comparé notre méthode à l’algorithme de segmentation par régularisation de graphe originel [Ta, 2009] ainsi qu’à l’approche par contours actifs de Chan et Vese, dans sa version scalaire [Chan et Vese, 2001] ainsi que vectorielle [Chan *et al.*, 2000]. Pour cette dernière, nous avons employé les mêmes caractéristiques d’Haralick que pour notre méthode.

Les paramètres optimaux des différentes méthodes ont été déterminés selon un principe d’optimisation aléatoire [Matyas, 1965] visant à obtenir la distance générique d’anomalie la plus faible.

2.3. RÉSULTATS

Le tableau 2.6 résume les résultats obtenus pour chacun de ces tests. Ceux-ci montrent une amélioration notable de la précision de segmentation, en particulier par rapport à la méthode de Ta dont nous nous inspirons et qui se révèle parfois incapable de produire un résultat exploitable.

Image	Notre méthode	Méthode de C&V scalaire	Méthode de C&V vectorielle (basée textures)	Méthode de Ta
echo1	1,82	2,41	3,45	8,31
echo2	2,02	1,59	1,25	N.A.
echo3	1,86	7,16	3,26	2,45
echo4	2,47	3,77	3,34	4,90
echo5	2,62	3,39	3,62	10,69
echo6	2,16	3,07	3,14	3,40
Moyenne	2.2	6.0	3.6	3.0
Écart type	0.3	3.1	1.8	0.8

TABLE 2.6 – Comparaison de la précision de segmentation des différentes méthodes sur un ensemble d’images échographiques 2D : distance générique d’anomalie.

Nous constatons aussi que notre méthode présente de meilleurs résultats que les méthodes de segmentation par contours actifs, y compris celle basée textures. Nous pouvons donc affirmer que la combinaison d’une méthode de segmentation par régularisation de graphes et d’un critère d’attache aux données basé textures permet une meilleure identifications des contours des objets à segmenter.

Les figures 2.21 et 2.22 illustrent les résultats obtenus pour deux des images de tests pour chacune des méthodes évaluées.

Comme évoqué dans le chapitre 1, l’un des avantages des approches par graphes est de permettre à un même algorithme de traiter différents types de données, notamment de dimensions différentes. Afin d’illustrer cela, nous avons appliqué notre méthode à la segmentation d’une image échographique 3D.

L’implémentation de cette méthode reste identique, seul le graphe doit être adapté aux données manipulées. Nous avons donc étendu à la 3D la notion de grille que nous utilisions dans les précédents tests, permettant de créer des 6-voisinages. Le calcul des caractéristiques d’Haralick a lui aussi été adapté afin de supporter des fenêtre volumiques ainsi que des décalages exprimés dans les trois dimensions de l’espace.

L’image *echo6* dont il est question dans le tableau 2.6 est une coupe 2D extraite de cette image 3D. Il n’a cependant pas été possible de valider numériquement la qualité de l’intégralité du résultat, la création de vérités terrains pour des données volumiques étant un processus relativement laborieux. Seule une confirmation visuelle de la précision du résultat global a donc pû être obtenue.

La figure 2.23 illustre l’image sur laquelle notre méthode a été appliquée ainsi que des représentations volumiques du maillage ainsi calculé.

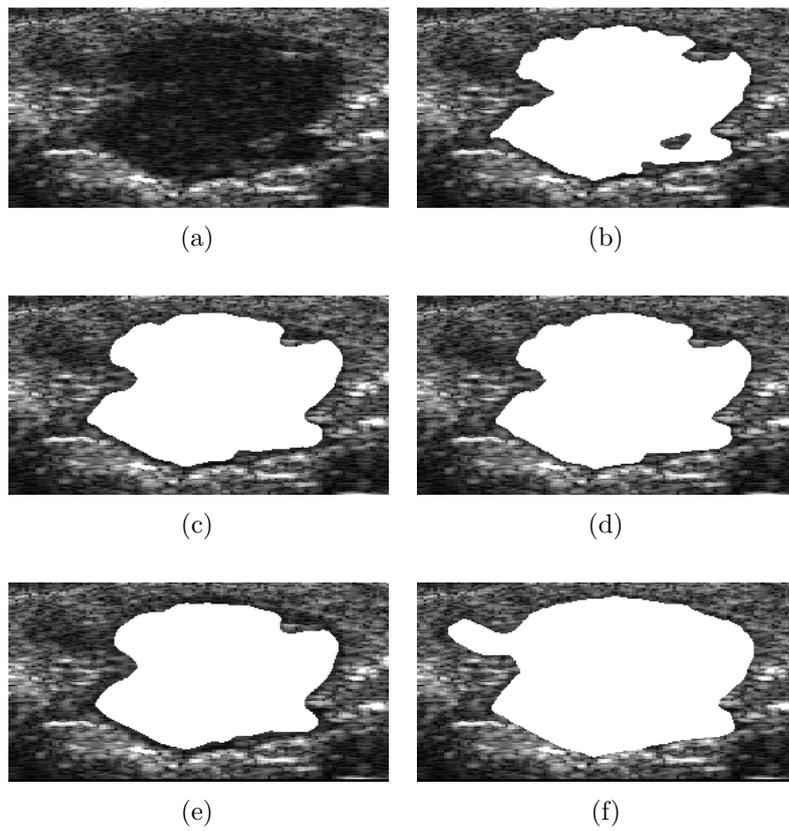


FIGURE 2.21 – Image *echo1*. (a) : image originale. (b) : vérité terrain. Résultats obtenus avec : (c) notre méthode, (d) la méthode de Chan & Vese scalaire, (e) la méthode de Chan & Vese vectorielle, (f) la méthode de Ta.

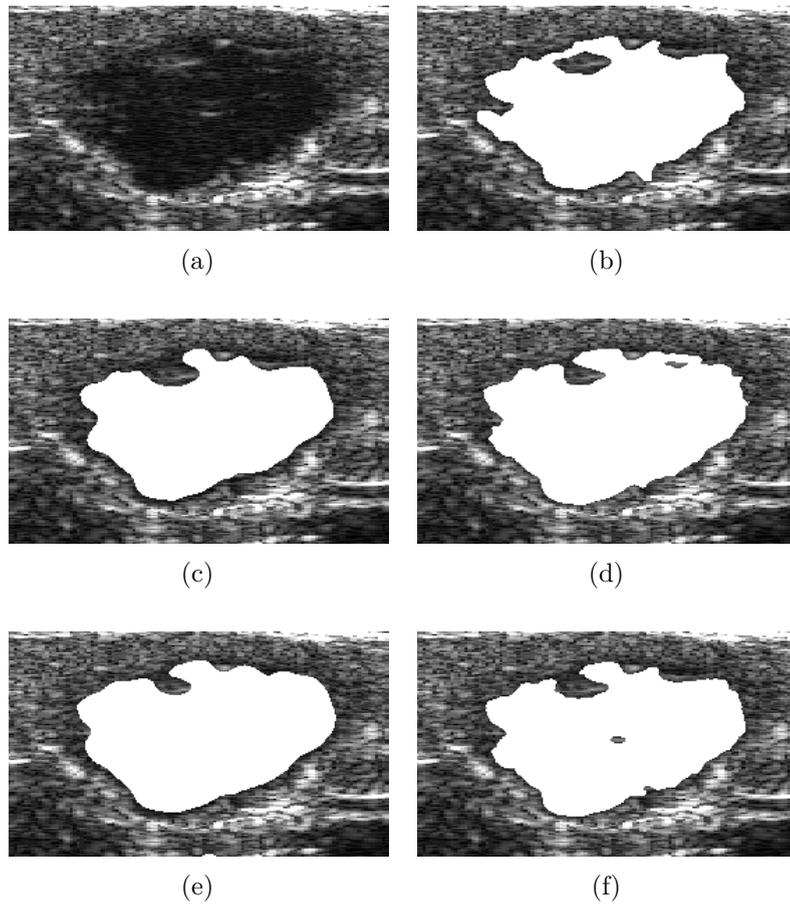
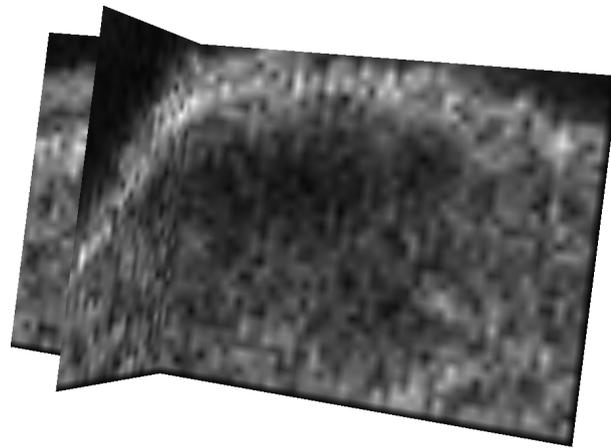
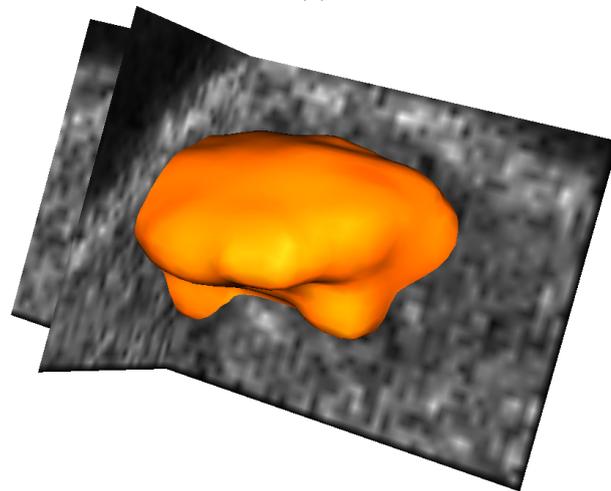


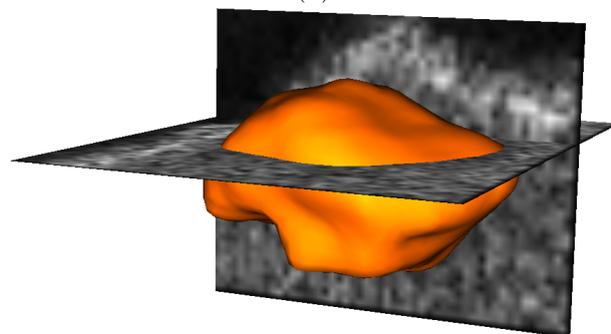
FIGURE 2.22 – Image *echo5*. (a) : image originale. (b) : vérité terrain. Résultats obtenus avec : (c) notre méthode, (d) la méthode de Chan & Vese scalaire, (e) la méthode de Chan & Vese vectorielle, (f) la méthode de Ta.



(a)



(b)



(c)

FIGURE 2.23 – Illustration des résultats obtenus sur une image échographique 3D. (a) : image initiale. (b) et (c) : représentation volumique du maillage calculé.

2.4 Conclusion

La méthode que nous avons présentée dans ce chapitre apporte à un récent algorithme de segmentation d'images par régularisation de graphes la capacité de tirer parti de caractéristiques de textures. Les différents tests réalisés ont permis d'illustrer le fonctionnement de l'algorithme ainsi que l'intérêt des approches non locales. Nous avons tout de même entrevu certaines limites de notre méthode.

La première concerne la détermination des paramètres λ_1 et λ_2 , dont les valeurs peuvent influencer grandement sur la qualité des résultats. Mis à part dans les cas extrêmement simples, il est relativement complexe de prédire quelle région doit voir son homogénéité privilégiée. Il est donc inenvisageable de les fixer automatiquement.

La seconde limite découle d'une simplification que nous avons choisi d'opérer sur le critère d'attache aux données. Afin de conserver un nombre raisonnable de paramètres, nous ne permettons pas de pondérer individuellement chaque caractéristique de textures. Nous avons montré que notre méthode est capable de fonctionner malgré un certain nombre de caractéristiques non pertinentes, mais cela se répercute sur les performances de l'algorithme, tant en précision qu'en vitesse d'exécution.

Comme nous avons pu le voir dans l'état de l'art des méthodes de caractérisation de textures réalisé en début de ce chapitre, aucune méthode existante ne permet d'obtenir une caractérisation « universelle » qui serait pertinente sur tous types de textures. Afin de remédier à cela, l'une des solutions est de combiner plusieurs descripteurs. Nous ne pouvons cependant pas avoir d'assurance concernant la pertinence de chacun d'entre eux, et un trop grand nombre de descripteurs non pertinents pourrait avoir un effet désastreux sur la capacité de notre méthode à résoudre certains problèmes de segmentation. Une sélection manuelle des caractéristiques les plus pertinentes reste donc nécessaire, ce qui réduit une éventuelle automatisation de la méthode.

Ces deux problèmes découlent de l'incapacité du terme d'attache aux données à considérer les caractéristiques de textures en fonction de leur pertinence. Afin de remédier à cela, nous proposons dans le prochain chapitre d'intégrer un classificateur supervisé au processus de segmentation. Celui-ci aura pour rôle de fournir une modélisation plus aboutie des textures à identifier, permettant ainsi l'expression d'un terme d'attache aux données plus précis.

2.4. CONCLUSION
