

Segmentation non supervisée des images naturelles texturées : une approche statistique

A Statistical Model-Based Approach to Unsupervised Texture Segmentation

par Charles KERVRANN et Fabrice HEITZ

IRISA/INRIA, Campus Universitaire de Beaulieu,
35042 Rennes Cedex, France

Résumé

La segmentation des images texturées constitue une étape préliminaire cruciale dans de nombreuses applications en analyse d'images. Les approches par modélisation statistique conduisent à de bons résultats dans ce domaine, lorsque les paramètres des modèles statistiques et le nombre de régions à extraire sont connus *a priori*. La segmentation non supervisée d'images texturées reste, par contre, un problème délicat, auquel aucune solution complète n'a été apportée jusqu'à présent. Nous contribuons à cet effort, en proposant une méthode de segmentation ne nécessitant pas de connaissance *a priori* sur le nombre ou le type de textures présentes dans l'image. La méthode présentée ici vise en particulier à gérer le problème de la détermination du nombre de régions dans l'image. L'algorithme de segmentation s'appuie sur l'analyse des statistiques spatiales du deuxième ordre et d'ordre supérieur des images originales. Un modèle markovien est utilisé pour spécifier les propriétés d'homogénéité de la carte de segmentation recherchée. Une estimation bayésienne de cette carte est obtenue grâce à un algorithme de relaxation déterministe, qui évalue conjointement le nombre de régions dans l'image. Un seul paramètre gère l'ensemble de la procédure de segmentation. Des résultats sur des mosaïques de textures naturelles et sur des exemples d'images réelles texturées démontrent l'aptitude du modèle à produire des segmentations de bonne qualité, sans connaissance *a priori* sur les caractéristiques des composantes texturées de l'image.

Mots clés : Analyse d'images, Segmentation non supervisée, Analyse de textures, Modèles statistiques, Champs markoviens.

Abstract

The general problem of unsupervised textured segmentation remains a largely unsolved issue in image analysis. Many studies proved that statistical model based texture segmentation algorithms yield good results provided that the model parameters and the number of regions are known a priori. In this paper, the problem of determining the number of regions is addressed. The segmentation algorithm relies on the analysis of second and higher order spatial statistics of the original images. The segmentation map is represented using a Markov Random Field model and a bayesian estimate of this map is computed using a deterministic relaxation algorithm. The segmentation algorithm does only require the tuning of one parameter. Results on hand-drawn images of natural textures and real textured images show the capability of the model to yield relevant segmentations when the number of regions and the texture classes are not known a priori.

Key words : Image analysis, Non-supervised texture segmentation, Statistical models, Markov Random Field models.

1. Introduction

La segmentation reste un des problèmes fondamentaux de l'analyse d'images et se pose en préalable à la plupart des schémas d'interprétation ou de décision. Rappelons que segmenter une image consiste à produire une partition de l'image initiale en régions possédant certaines caractéristiques d'homogénéité et correspondant aux différentes composantes de la scène, [17]. Ces composantes peuvent ensuite être utilisées comme primitives en entrée

de schémas d'identification, de caractérisation et d'interprétation du contenu de l'image.

Dans cet article nous nous intéressons plus particulièrement à la segmentation des images texturées, problème notoirement plus complexe que celui de la segmentation au sens des niveaux de gris (les techniques présentées ici s'appliquent toutefois par extension à ce dernier cas). La méthode de segmentation proposée s'appuie sur des attributs statistiques locaux de texture et sur un modèle statistique global de l'image segmentée.

La segmentation des images texturées s'est posée longtemps comme un problème particulièrement difficile, d'une part parce que la notion de texture ne connaît toujours pas de définition universellement admise, d'autre part parce que cette information de texture n'est pas directement observable (contrairement à l'information de luminance par exemple). Ceci n'a néanmoins pas empêché le développement de nombreuses techniques d'analyse, de classification et de segmentation des images texturées, s'appuyant sur des approches, des descripteurs et des modèles variés, [17, 18, 19].

Depuis les travaux fondateurs d'Haralick sur les descripteurs statistiques de textures, [18], des améliorations décisives dans la qualité des segmentations produites ont été obtenues par l'introduction de modèles statistiques globaux, [8, 11, 14, 17, 16, 22]. Les modèles markoviens, [1, 11], en particulier, permettent la prise en compte des variabilités de la structure des textures naturelles analysées, et autorisent la spécification des propriétés d'homogénéité des cartes de segmentation produites.

Des résultats de qualité ont ainsi été présentés, au milieu des années 1980, dans le cas où le nombre de régions et les paramètres des modèles de texture étaient connus *a priori* ou faisaient l'objet d'un apprentissage préalable, conduisant à une segmentation *supervisée* de l'image. Les techniques *non supervisées* visent à s'affranchir de ces contraintes et on note de nombreux efforts récents dans cette direction, [9, 10, 12, 14, 20, 22, 27]. Il faut toutefois noter qu'une segmentation robuste complètement guidée par les données reste actuellement hors d'atteinte.

Dans [12, 22] les paramètres de modèles statistiques stationnaires de textures sont identifiés sur des fenêtres d'apprentissage disjointes. Les fenêtres présentant des paramétrisations voisines sont fusionnées. Cette première segmentation grossière est ensuite raffinée par des algorithmes de relaxation. Dans [14], des attributs de texture classiques sont utilisés conjointement avec un modèle markovien sur les régions. L'algorithme, s'appuyant sur une relaxation déterministe des étiquettes de région est partiellement non supervisé, mais nécessite une connaissance *a priori* du nombre approximatif de régions dans l'image ainsi qu'un ajustement fin de seuils discriminants sur les attributs de texture. Un algorithme de segmentation non supervisé en deux classes de textures, s'appuyant sur des caractéristiques de cooccurrence et un modèle markovien est présenté dans [9] comme partie intégrante d'un système d'autoguidage pour un robot d'entretien des espaces naturels. Dans [20] une technique de détection de contours dans un espace multidimensionnel d'attributs statistiques de textures est décrite. Descombes *et al.* [10] proposent une technique de segmentation de textures par seuillage d'une distance entre probabilités conditionnelles locales d'un modèle markovien non stationnaire. La topologie (structure de voisinage) ainsi que les paramètres du modèle sont estimés préalablement.

Lakshmanan et Derin proposent dans [21] un algorithme de segmentation non supervisé (au sens des niveaux de gris) s'appuyant sur des modèles markoviens dont les paramètres sont estimés simultanément avec la carte de segmentation. Une technique similaire est reprise dans les travaux de Won et Derin, [27] pour des modèles markoviens gaussiens de texture. Le nombre de régions est déterminé grâce à un critère statistique d'information s'inspirant du critère de Rissanen (une approche voisine est développée

dans [28] pour déterminer le nombre de classes dans une segmentation).

L'utilisation d'un critère statistique d'information conduit à réaliser plusieurs segmentations en faisant varier le nombre de régions. Ceci peut se révéler coûteux, en particulier lorsque le nombre de régions est important. Nous proposons ici une alternative en gérant la détermination du nombre de régions dans l'algorithme de relaxation. Cette technique avait été introduite par François et Bouthemy, [13] pour un modèle markovien de segmentation au sens du mouvement. Comme dans [14, 20], les textures sont décrites par des attributs statistiques mesurés dans des fenêtres locales, ce qui permet de réduire le nombre de paramètres à ajuster dans le modèle. Un modèle markovien est introduit pour imposer des contraintes *a priori* sur les cartes de segmentation calculées. La segmentation optimale est définie au sens du Maximum a Posteriori (MAP) et estimée grâce à un algorithme de relaxation déterministe. La méthode proposée ne nécessite, en particulier, pas de phase d'apprentissage sur les textures analysées. La méthode requiert seulement l'ajustement d'un paramètre qui correspond à une probabilité déterminant les seuils de discrimination entre textures.

La suite de l'article s'organise comme suit : la section 2 est consacrée à la présentation générale de la technique d'analyse adoptée ici (estimation bayésienne et modélisation par champs markoviens). L'approche développée est décrite en section 3. Les résultats (section 4) sur des mosaïques synthétiques de textures naturelles (extraites de l'album de Brodatz) ainsi que sur des images naturelles texturées permettent d'évaluer la robustesse de la méthode proposée sur des classes d'images texturées représentatives et variées.

2. Analyse bayésienne de l'image et champs markoviens

Nous rappelons ici les grandes lignes de la modélisation markovienne et de l'estimation bayésienne globale en traitement d'images, permettant d'introduire les principales notions utilisées par la suite.

Dans les approches markoviennes en imagerie, les observations (variables observables sur lesquelles s'appuie l'analyse de l'image) et les étiquettes (informations non observables que l'on cherche à extraire des observations) sont supposées former des champs de variables aléatoires que nous noterons respectivement O (champ d'observations) et E (champ d'étiquettes) et sont définies respectivement sur des grilles rectangulaires T et S (identiques ou non) :

$$O = \{O_s, s \in T\} \quad E = \{E_s, s \in S\}$$

où s désigne un site d'une des grilles. On notera $o = \{o_s, s \in T\}$ et $e = \{e_s, s \in S\}$ des réalisations de chacun de ces deux champs. Les étiquettes E_s prennent leurs valeurs, discrètes ou continues, dans un ensemble Λ . L'ensemble de toutes les configurations possibles de E est donc $\Omega = \Lambda^{|S|}$ où $|S|$ désigne le cardinal de S .

Une analyse bayésienne de l'image suppose que l'on dispose d'une connaissance des relations (statistiques) qu'entretiennent les variables d'observations et les étiquettes. Cette connaissance s'exprime généralement à travers la modélisation ou spécification de la distribution *a posteriori* $p(E = e|O = o)$. Le cadre de l'estimation bayésienne permet alors la définition d'estimateurs statistiques \hat{e} du champ e . Si différents estimateurs ont été considérés en analyse d'image [23] (estimateur du Maximum A Posteriori (MAP), estimateur du Mode de la Marginale a Posteriori (MPM), estimateur du champ moyen), l'estimateur le plus communément utilisé reste toutefois celui du Maximum a Posteriori.

Après application de la règle de Bayes, il conduit au problème d'estimation suivant :

$$\hat{e} = \arg \max_{e \in \Omega} p(o|e) p(e) \quad (1)$$

Le premier terme de la fonction à maximiser (probabilité d'occurrence des observations conditionnellement aux étiquettes) sera déterminé en modélisant le lien entre les étiquettes et les observations. Le second, probabilité d'occurrence de e , résultera d'une modélisation *a priori* des propriétés du champ des étiquettes.

Une modélisation simple consiste à supposer le champ E markovien relativement à un système de voisinage $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$, c'est-à-dire :

- (i) $\forall e \in \Omega, p(E = e) > 0$
- (ii) $\forall s \in S, p(E_s = e_s | E_t = e_t, t \neq s) = p(E_s = e_s | E_t = e_t, t \in \mathcal{G}_s)$.

En vertu du théorème de Hammersley et Clifford [2], E suit alors une distribution de Gibbs :

$$p(E = e) = \pi(e) \triangleq \frac{1}{Z} e^{-U_2(e)} \quad (2)$$

où

- La fonction $U_2(e)$, dite *fonction d'énergie* se décompose en une somme de fonctions locales V_c , dites *fonctions de potentiel*, définies sur les cliques $c \in \mathcal{C}$ engendrées par \mathcal{G}^1 :

$$U_2(e) \triangleq \sum_{c \in \mathcal{C}} V_c(e) \quad (3)$$

V_c n'étant en fait fonction que des étiquettes portées par les sites de c .

- Z , constante de normalisation appelée *fonction de partition*, est définie par :

$$Z \triangleq \sum_{e \in \Omega} e^{-U_2(e)} \quad (4)$$

C'est le choix des potentiels locaux V_c qui va permettre de conférer au champ d'étiquettes des propriétés *a priori* pertinentes.

Supposons que $p(o|e) > 0, \forall o, \forall e$. En posant $U_1(e, o) = -\ln(p(o|e))$ on voit que \hat{e} est un minimum global de la fonction d'énergie $U(e, o) = U_1(e, o) + U_2(e)$:

$$\hat{e} = \arg \min_{e \in \Omega} U_1(e, o) + U_2(e) . \quad (5)$$

Dans de nombreux problèmes, le terme d'énergie U_1 rendant compte des interactions entre observations et étiquettes peut également se décomposer en une somme de fonctions d'interactions locales (souvent ponctuelles) ce qui conduit à une distribution de Gibbs pour la loi *a posteriori*.

La recherche d'un minimum (global) de la fonction d'énergie $U(e, o)$ est un problème d'optimisation difficile en raison de la dimension (généralement très élevée) de l'espace des configurations Ω et de la présence de nombreux minima locaux dans le cas de modèles non linéaires. Une solution peut être obtenue soit par des algorithmes de relaxation déterministes dont l'algorithme ICM (*iterative conditional modes*) est un exemple [3], soit par des algorithmes de relaxation stochastique, [15]. Les algorithmes stochastiques convergent théoriquement vers un minimum global de la fonction d'énergie, mais avec une lenteur souvent rédhibitoire pour des espaces de configurations très grands. On leur préfère donc souvent des algorithmes déterministes. Ceux-ci, s'ils convergent beaucoup plus rapidement, présentent cependant l'inconvénient de se piéger dans d'éventuels minima locaux de la fonction d'énergie et risquent donc de converger vers des estimations de qualité inférieure. Pour pallier ce défaut on peut, lorsque c'est possible, initialiser ces algorithmes avec des configurations proches d'un minimum global. D'autres solutions efficaces consistent à utiliser des algorithmes partant d'une approximation convexe de la fonction d'énergie et introduisant une non-convexité graduelle, [4] ou encore d'avoir recours à des stratégies d'estimation multirésolution, [5, 25]. L'estimation sur des structures multirésolutions, lorsqu'elles s'accompagne d'une gestion adéquate de la paramétrisation des modèles aux différents niveaux de résolution, conduit à des gains importants en vitesse de convergence et en qualité des estimées finales, [25].

La spécification des différentes composantes de la fonction d'énergie globale décrivant les interactions entre les variables dépend bien entendu du problème particulier à résoudre. Le modèle défini dans le cas de la segmentation au sens des textures est décrit dans la section suivante.

3. Un algorithme non supervisé de segmentation au sens des textures

L'algorithme de segmentation d'images texturées décrit ici est non supervisé dans le sens où :

- aucune connaissance *a priori* sur la nature des textures présentes dans l'image n'est requise (en particulier, aucun apprentissage préalable n'est effectué),
- le nombre de régions texturées est inconnu au départ.

¹ \mathcal{C} est l'ensemble des parties non vides de S réduites à un singleton ou dont les éléments sont voisins deux à deux.

3.1. CARACTÉRISATION STATISTIQUE DES TEXTURES

Les observations sur lesquelles nous nous appuyons pour segmenter l'image originale sont constituées de différentes caractéristiques (ou attributs) statistiques de texture. Elles se présentent sous la forme d'images d'attributs calculés préalablement sur l'image à traiter. Les attributs que nous avons utilisés incluent neuf caractéristiques classiques extraites des matrices de cooccurrence définies dans [18]. Il s'agit de la moyenne, la variance, l'énergie, la corrélation, l'entropie, le contraste, l'homogénéité, les attributs «cluster shade» et «cluster prominence», [18]. Ces attributs sont calculés sur des images requantifiées sur seize niveaux de gris sur des fenêtres locales de taille 5×5 selon quatre orientations différentes. La caractéristique mesurée sur la fenêtre locale est affectée au point central de cette fenêtre. Cinq autres caractéristiques, adoptées dans [14], pour leur caractère discriminant, ont été retenues : le niveau de gris, le contraste, et trois résiduels isotropes ou directionnels définis dans [14] correspondant à l'erreur résiduelle associée à des modèles autorégressifs à paramètres fixes. Cette liste n'est pas limitative et l'approche présentée se généralise immédiatement à d'autres descripteurs de texture. Ainsi, à chaque site est assigné un vecteur de caractéristiques $o_s = (o^1(s), \dots, o^m(s))$ où $o^i(s)$ désigne la $i^{\text{ème}}$ caractéristique de texture calculée au site s .

3.2. MODÈLE DE SEGMENTATION

Comme précédemment, on désigne par e_s , la variable cachée affectée au site s qui décrit ici le numéro (l'étiquette) de la région à laquelle ce site appartient. Le champ des observations formé par les vecteurs d'attributs statistiques de textures est représenté par $o = \{o_s, s \in S\}$.

Le but est d'estimer le champ $e = \{e_s, s \in S\}$ et R (le nombre de régions) au moyen d'un estimateur du MAP. La fonction d'énergie définie se décompose en deux termes :

$$U(o, e, R) = U_1(o, e, R) + U_2(e) \quad (6)$$

La contribution U_2 à l'énergie globale correspond au modèle *a priori* qui spécifie les propriétés statistiques de la carte des étiquettes de région e . Ce modèle est posé comme étant markovien. Dans la segmentation nous privilégions les partitions en *régions spatialement homogènes* grâce à des potentiels d'interaction locale de la forme :

$$V_2(e_s, e_t) = \mu(1 - 2\delta(e_s - e_t)), \quad (7)$$

où μ est un paramètre positif, et $\delta(x) = 1$ pour $x = 0$, $\delta(x) = 0$ sinon. Le système de voisinage retenu est un huit-voisinage, et seules les cliques à deux éléments sont considérées (figure 1), [15]. L'énergie U_2 s'en déduit d'après l'équation (3).

Notre modèle d'interaction entre observations et étiquettes (énergie U_1) s'inspire d'un modèle défini par Geman *et al.* dans [14]. Le modèle de Geman *et al.* est basé sur une mesure de disparité entre descripteurs définis sur des sites voisins. La mesure de disparité s'appuie sur une distance de Kolmogorov-Smirnov entre les distributions des descripteurs associés aux deux sites voisins, [14]. Nous introduisons ici une nouvelle forme de la fonction de

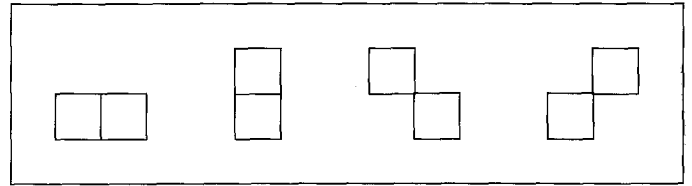


Figure 1. – Cliques à deux éléments utilisées dans le modèle *a priori*.

potentiel qui mesure la distance, en un point, entre la distribution des attributs de texture estimés sur la région à laquelle on souhaite affecter ce point (distribution globale), et celle des attributs associés à une fenêtre locale centrée au point considéré (distribution locale). Cette modélisation nous permet, en particulier, d'identifier le nombre (inconnu) de régions, comme nous le verrons par la suite (en section 3.3). La fonction d'énergie U_1 se décompose comme suit, [14] :

$$U_1(o, e, R) = \sum_{s \in S} V_1(o(B_s), o(R(e_s))) \quad (8)$$

où : $o(B_s) = \{o_t, t \in B_s\}$ sont les vecteurs des attributs associés à la fenêtre locale centrée au site s et $o(R(e_s)) = \{o_t, t \in R(e_s)\}$ est l'ensemble des vecteurs d'attributs de la région d'étiquette e_s (figure 2).

Le potentiel d'interaction V_1 est défini par :

$$V_1(o(B_s), o(R(e_s))) = \sum_{i=1}^m (2\Gamma(d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)}) - 1) \quad (9)$$

où $d(o^{(i)}(R(e_s)), o^{(i)}(B_s))$ représente la distance de Kolmogorov-Smirnov entre les fonctions de répartition de l'attribut de texture $o^{(i)}$, respectivement sur la région d'étiquette e_s et sur la fenêtre B_s centrée en s .

$c^{(i)}$ est un seuil obtenu à partir de la table de la distribution limite de Kolmogorov-Smirnov, déduit pour un choix de la probabilité p définie par :

$$p = \text{Prob}(d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)}) \quad (10)$$

qui est la probabilité que la distance de Kolmogorov dépasse le seuil $c^{(i)}$ sachant que les deux distributions sont les mêmes; p est le seul paramètre qui est ajusté, en pratique, dans la version implantée de l'algorithme. La fonction Γ rend la valeur 1 si $d(o^{(i)}(R(e_s)), o^{(i)}(B_s)) > c^{(i)}$, 0 sinon. Ce potentiel favorise l'affectation de l'étiquette $e_s = r$ au site s si la distribution globale de la région d'étiquette r est voisine de la distribution locale autour du site s .

Finalement l'estimation au sens du MAP conduit à la minimisation suivante :

$$\hat{e} = \arg \min_{e \in \Omega} U_1(o, e, R) + U_2(e) \quad (11)$$

Cette minimisation, ainsi que la détermination du nombre de régions R , est menée grâce à un algorithme de relaxation déterministe décrit dans la section suivante.

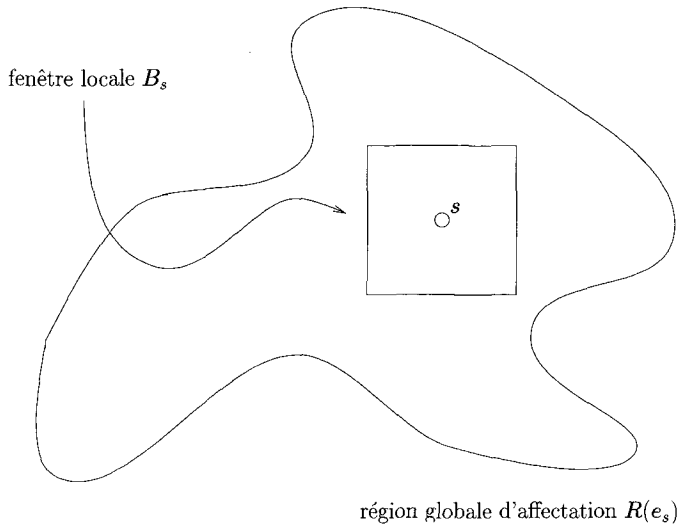


Figure 2. – Domaines supports des distributions locales et globales.

3.3. SCHÉMA DE RELAXATION

Une relaxation stochastique [15] a été écartée en raison de son coût calculatoire prohibitif. Nous adoptons une version modifiée de l'algorithme déterministe ICM, [3], qui identifie le nombre inconnu de régions simultanément à l'estimation des étiquettes e .

Estimation de e

On considère que l'estimation courante du nombre de régions est \hat{R} . Au site s on doit théoriquement proposer pour l'étiquette e_s l'ensemble des étiquettes possibles $r = 1 \dots \hat{R}$. L'étiquette qui conduit localement à la décroissance d'énergie la plus forte est alors adoptée. Pour diminuer la charge calculatoire nous avons restreint l'ensemble des étiquettes candidates pour le site s aux étiquettes voisines de ce site (l'étiquette e_s comprise), ceci sans dégradation visible des résultats. Pour gérer le nombre inconnu de régions, nous introduisons une étiquette supplémentaire, dite étiquette d'attente ρ , [13]. Cette étiquette est affectée au site s si aucune des étiquettes candidates du voisinage de s ne convient. Plus précisément, parmi les étiquettes candidates, quatre-voisines de S et l'étiquette d'attente ρ on retient l'étiquette $r = \hat{r}$ qui minimise l'énergie locale déduite de l'équation (11) et qui s'exprime comme suit :

Pour $r \neq 0$:

$$\Delta U_s(r) = \sum_{t \text{ voisins de } s} \mu(1 - 2\delta(r - e_t)) + \sum_{i=1}^m (2\Gamma(d(o^{(i)}(R(e_s = r)), o^{(i)}(B_s)) > c^{(i)}) - 1) \quad (12)$$

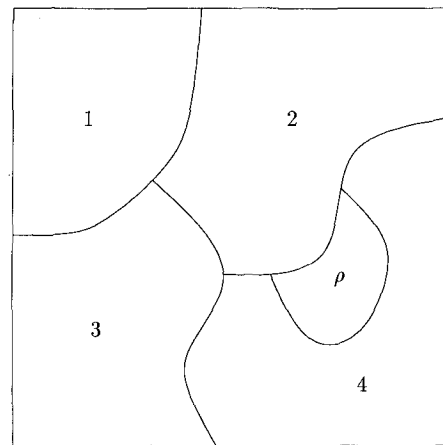
L'énergie locale associée à ρ étant définie par :

$$\Delta U_s(\rho) = \sum_{c \in C} \mu(1 - 2\delta(\rho - e_t)) + \phi \quad (13)$$

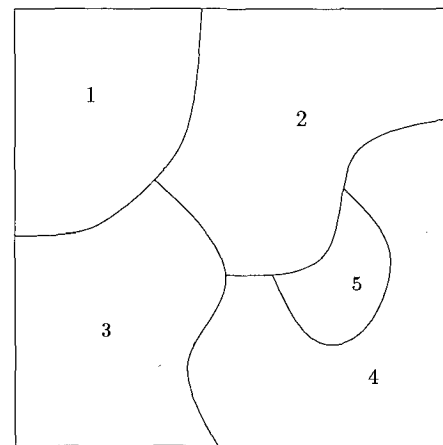
ϕ est le paramètre associé à la classe d'attente. Une valeur de ϕ faible favorise l'émergence de nouvelles régions ($\phi > 0$). Les différents sites de l'image sont visités de façon dirigée grâce à une table d'instabilité, qui permet de se focaliser prioritairement sur les sites les plus mal étiquetés. Cette technique s'inspire de la méthode HCF proposée par Chou *et al.* dans [7].

Détermination de R

Le nombre courant de régions \hat{R} évolue tout au long de la phase de relaxation. A l'initialisation, \hat{R} est fixé à 1, ce qui correspond à une seule région couvrant toute l'image. A la fin de chaque phase de relaxation (figure 3), tous les sites étiquetés en attente sont



avant renumérotation : $\hat{R} = 4$



après renumérotation : $\hat{R} = 5$

Figure 3. – Rénumérotation des sites étiquetés en attente (étiquette ρ).

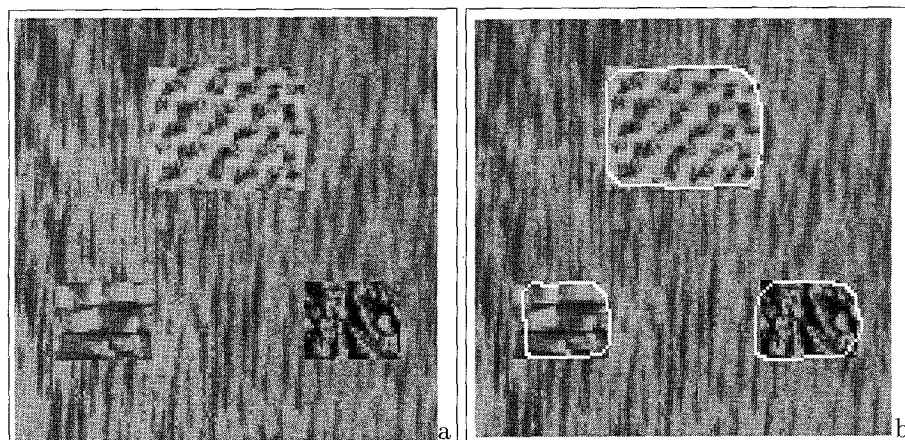


Figure 4. – Mosaïque de 4 textures (a) et segmentation finale (b) ($p = 0.5$)

renumérotés en composantes connexes, si ces composantes sont de taille suffisante (de l'ordre de 100 pixels). Une nouvelle phase de relaxation, avec la nouvelle partition est ensuite exécutée et ainsi de suite, jusqu'à stabilisation des étiquettes.

Initialisation

L'algorithme de relaxation utilisé étant déterministe, la segmentation finale dépend de la configuration initiale. Comme nous l'avons déjà précisé, \hat{R} est initialisé à 1, ce qui correspond à une seule région définie sur toute l'image. Puis une phase d'estimation au sens du maximum de vraisemblance (ne prenant pas en compte le terme d'énergie *a priori* U_2) est exécutée. La valeur de ϕ_0 retenue au cours de cette étape, est différente de celle de ϕ associée aux phases de relaxations proprement dites. ϕ_0 est fixé à une valeur élevée qui défavorise l'émergence de petites régions artefactuelles non significatives dans cette première phase du processus de segmentation. La valeur $\phi_0 = 10$ a été utilisée dans tous les cas. L'image issue de cette phase définit la partition initiale pour l'algorithme décrit précédemment.

4. Résultats expérimentaux

La technique de segmentation développée est appliquée à différents types de mosaïques composées de textures naturelles extraites de l'album de Brodatz, [6] ainsi que sur des images réelles texturées. Aucun apprentissage préalable sur les textures présentes dans une mosaïque donnée n'est effectuée, et le nombre de régions de textures est supposé inconnu au départ.

Les attributs de cooccurrence sont calculés sur une image requantifiée linéairement sur 16 niveaux de gris, dans des fenêtres 5×5 , à une distance de 3 suivant 4 orientations. Les autres paramètres (niveaux de gris, contraste et résiduels directionnels) sont calculés localement sur des fenêtres de taille 15×15 . La taille des

fenêtres B_s sur lesquelles sont établies les distributions de probabilité locales est de 13×13 (figure 2). En outre, nous adoptons la valeur $\mu = 1.5$ qui permet d'obtenir une bonne pondération entre le terme lié aux observations U_1 et le terme modélisant l'*a priori* U_2 . Un choix de ϕ égal à 2.5 permet par ailleurs de discriminer les textures sans trop de fausses alarmes. Rappelons que la configuration initiale correspond à une seule région et que ϕ_0 est fixé à 10. Les seuils $c^{(i)}$ sont choisis tous identiques et calculés à partir de la spécification de la probabilité p définie par l'équation (10). *Le seul paramètre que nous avons fait varier dans nos expérimentations est p .* Il faut toutefois noter, comme on l'observe souvent [14] lorsque l'on détermine des seuils à partir de tables statistiques, que les valeurs de p à ajuster pour obtenir des segmentations satisfaisantes peuvent être relativement éloignées de celle que préconiserait la théorie. L'ajustement de ce paramètre n'est toutefois pas très sensible (une précision de 10% étant suffisante en pratique) : un exemple des résultats obtenus pour différentes valeurs du paramètre p est présenté dans la suite (figure 6).

Nous décrivons tout d'abord des résultats de traitements obtenus sur diverses mosaïques présentant des textures et des configurations géométriques variées. La figure 4a présente la première mosaïque traitée. Elle se compose de 4 textures différentes (bois, chevrons, écailles, cuir). Les frontières de la segmentation finale sont surimposées sur l'image originale en figure 4b. La figure 5 présente à titre d'illustration les étapes intermédiaires du processus de relaxation. Le résultat de l'estimation au sens du maximum de vraisemblance, qui sert d'initialisation au processus de relaxation, est présenté figure 5a. Comme on le constate la non prise en compte du modèle markovien *a priori* conduit à une segmentation partielle et bruitée. Le résultat de la première étape de relaxation apparaît en figure 5b : le terme *a priori* permet clairement de contraindre la cohérence spatiale de l'image. La figure 5c présente un résultat intermédiaire de réétiquetage en composantes connexes. Le nombre final de régions restituées par l'algorithme est de quatre dans ce cas (figure 4b ou 5d). Les frontières des différentes textures sont restituées avec une assez bonne précision, même si le modèle markovien *a priori* tend à lisser les portions de frontières à forte courbure. Une solution à ce problème

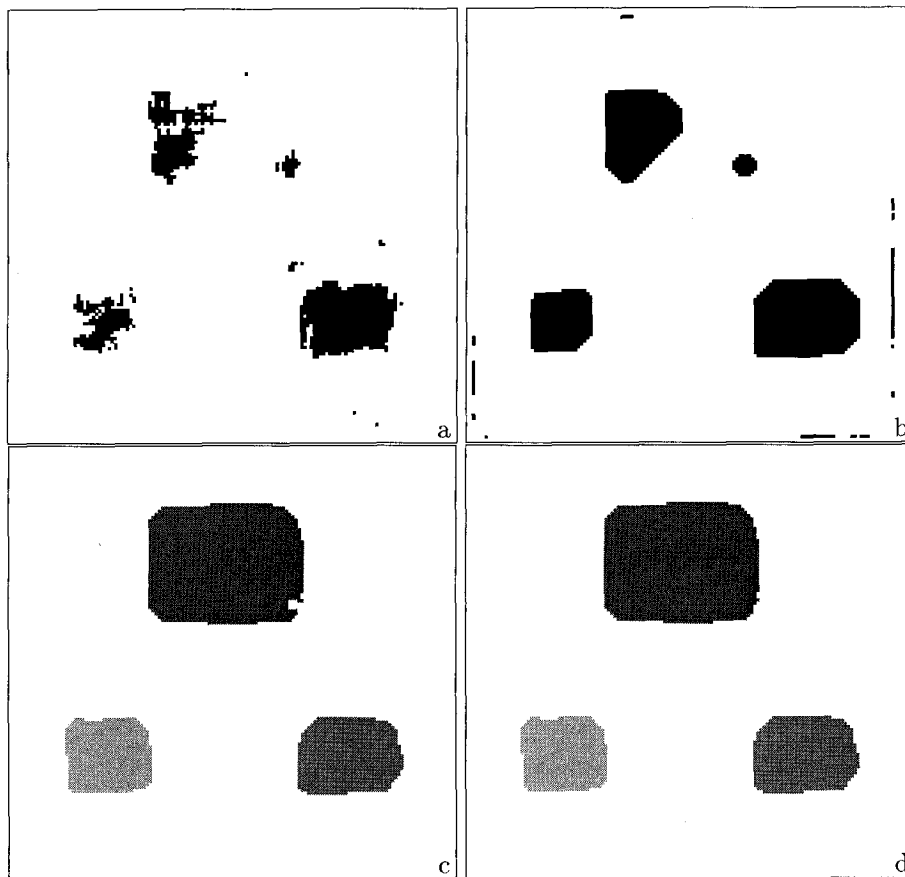


Figure 5. – Étapes intermédiaires du processus de relaxation : (a) estimation au sens du maximum de vraisemblance, (b) première étape de relaxation, (c) étape de relaxation intermédiaire (après réétiquetage en composantes connexes), (d) segmentation finale

est proposée, par exemple, par Geman *et al.* dans [15], en adjoignant au modèle des sites contours sur lesquels des contraintes géométriques peuvent être imposées. Cet élément de modélisation n'a toutefois pas été introduit ici. Le temps de calcul, obtenu sur une station SUN-4, est de l'ordre de 25 mn pour une image de taille 128×128 pixels.

La figure 6 illustre, sur le même exemple, la sensibilité du résultat de la segmentation au choix du paramètre p . Rappelons que le paramètre p est une probabilité définie par l'équation (10) ($0 \leq p \leq 1$). Comme on peut le constater, un ajustement fin de ce paramètre n'est pas nécessaire pour obtenir une segmentation pertinente (des valeurs de p entre 0,45 et 0,55 conduisent à des résultats satisfaisants dans ce cas). La dégradation de la segmentation est graduelle lorsque l'on s'éloigne de la plage de bon réglage du paramètre (figures 6d, 6e et 6f).

Des résultats qualitativement équivalents ont été obtenus sur d'autres mosaïques composées d'objets de géométries différentes (figure 7 et figure 8). Une certaine imprécision sur la localisation des frontières relève du modèle markovien *a priori* (comme nous l'avons déjà dit) mais s'explique également par la taille des fenêtres locales sur lesquelles sont calculées les attributs statistiques.

Nous présentons également quatre exemples de traitement d'images naturelles texturées (figures 9, 10, 11, et 12), qui ont été choisies pour leur représentativité et leur difficulté. Les textures qu'elles présentent sont variées et peuvent dans certains cas se révéler localement non stationnaires (figures 10, 11, et 12). Les non-stationnarités locales sont difficiles à traiter et conduisent généralement à des images sur-segmentées, si elles ne sont pas prises en compte dans la modélisation². L'approche proposée révèle une certaine robustesse par rapport à ce problème comme les résultats obtenus le démontrent.

La figure 9 correspond à une image aérienne texturée de champs cultivés : la frontière entre les deux régions est localisée avec une bonne précision, bien que la région supérieure soit inhomogène. La figure 10 est un exemple typique des situations rencontrées dans les images naturelles, pour lesquelles les textures sont non stationnaires et le nombre effectif de régions est difficile à évaluer. Ici, l'algorithme segmente l'image en quatre régions : une région a été créée dans une zone de non-stationnarité de la partie boisée.

² Des modèles non stationnaires de texture ont été proposés récemment dans [10, 26].

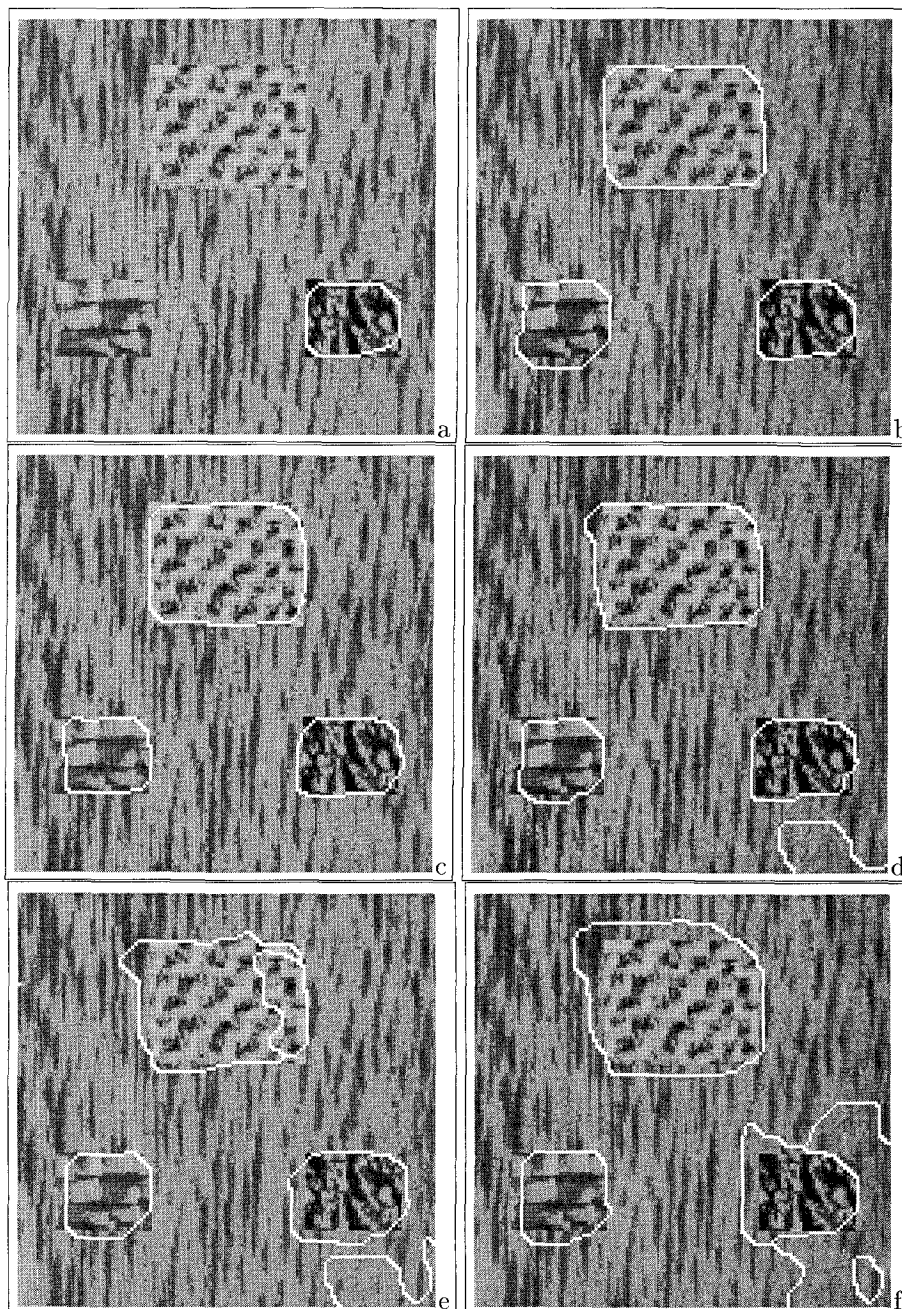


Figure 6. – Sensibilité de la segmentation à l’ajustement du paramètre p : (a) $p = 0.3$, (b) $p = 0.4$, (c) $p = 0.5$, (d) $p = 0.6$, (e) $p = 0.7$, (f) $p = 0.8$.

La figure 11 présente la segmentation d’une image satellitaire aux difficultés similaires. Nous présentons enfin un dernier exemple de segmentation sur une image échographique comportant un fort bruit de speckle (figure 12). Dans ce cas, l’algorithme réussit à extraire l’objet d’intérêt du fond, bien que la frontière de l’objet ne soit définie qu’imparfaitement sur la partie gauche de l’objet.

Ces différents exemples permettent d’illustrer l’aptitude du schéma de segmentation proposé à gérer des situations difficiles (habituelles dans les applications!) dans lesquelles peu de connaissances *a priori* existent sur la nature et le nombre de régions texturées se présentant dans l’image et où les textures observées possèdent souvent un certain degré de non stationnarité.

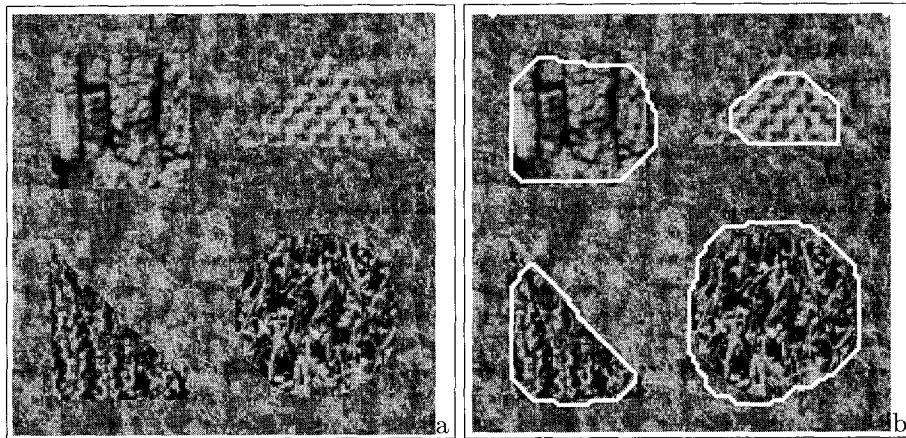


Figure 7. – (a) mosaïque de 5 textures, (b) segmentation finale ($p = 0.65$)

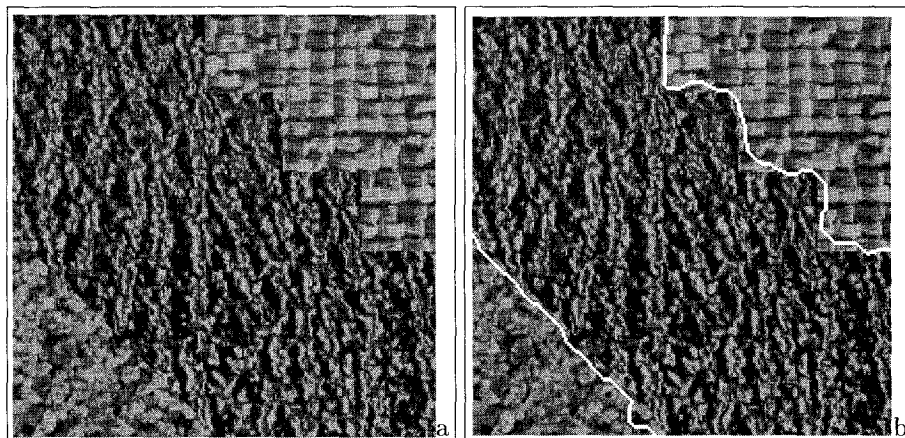


Figure 8. – (a) mosaïque de 3 textures, (b) segmentation finale ($p = 0.75$)

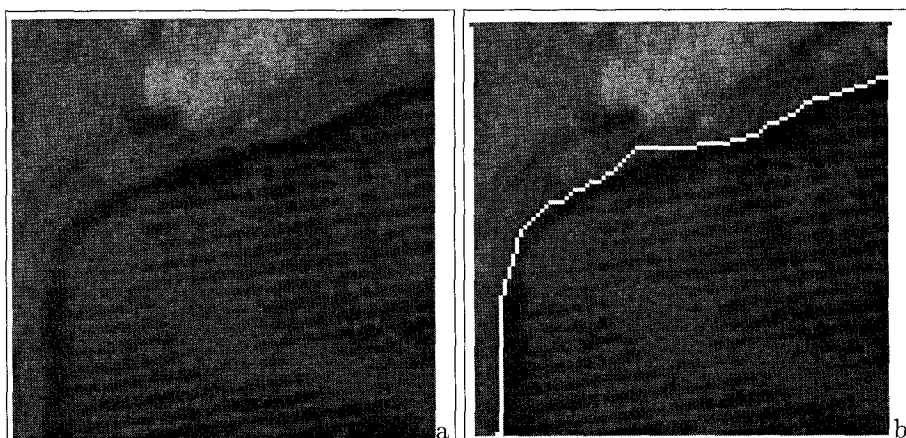


Figure 9. – (a) image aérienne texturée (Calvi – Mission MATRA 1987 – DRET), (b) segmentation finale ($p = 0.2$)

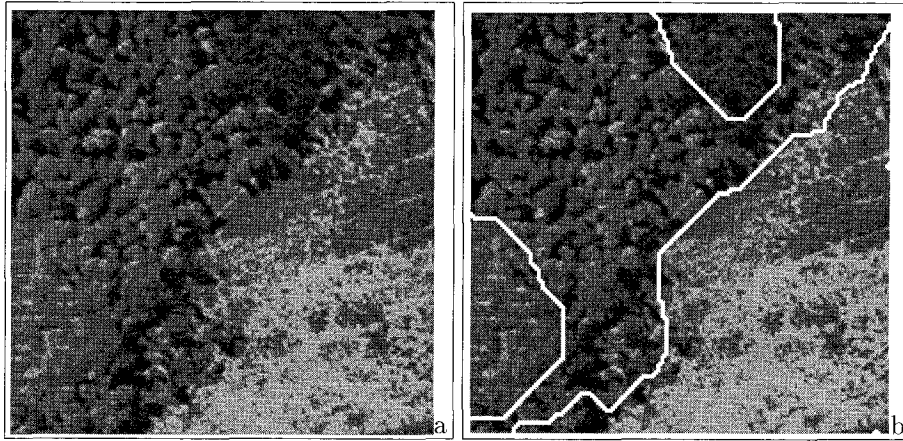


Figure 10. – (a) image aérienne texturée (Roumard – Mission INRA Télédétection 1976 – DRET), (b) segmentation finale ($p = 0.65$)

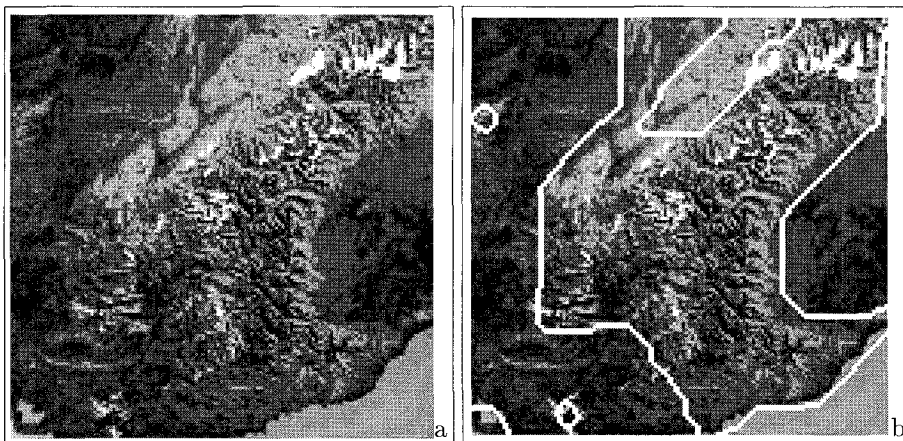


Figure 11. – (a) image aérienne des Alpes (France), (b) segmentation finale ($p = 0.25$)

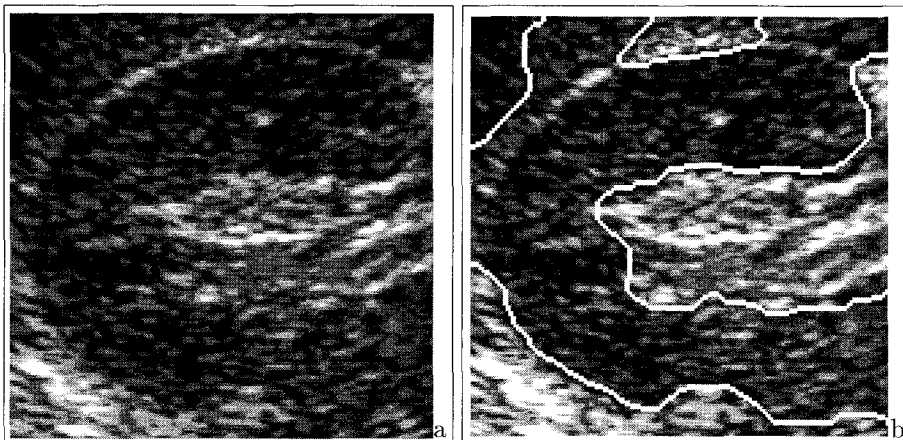


Figure 12. – (a) image échographique du rein (LEP – base de données images, GDR 134), (b) segmentation finale ($p = 0.3$)

5. Conclusion

Dans cet article, une méthode statistique de segmentation d'images par les textures a été décrite. La méthode proposée ne nécessite de connaissance *a priori* ni sur le type de texture, ni sur le nombre de régions texturées présentes dans l'image. La méthode nécessite l'ajustement d'un seul paramètre qui détermine les seuils discriminant les différentes textures. Cet algorithme constitue donc un premier pas vers une segmentation des images texturées complètement guidée par les données. Pour cela des procédures statistiques de détermination du paramètre seraient à introduire.

Soulignons que dans l'algorithme implanté, nous avons opté pour un choix particulier de paramètres de texture qui mériterait bien entendu d'être enrichi. L'approche que nous proposons s'y prête de façon naturelle, et les techniques de sélection de paramètres proposées dans [24] pourraient s'avérer fort utile.

Les résultats que nous avons obtenus sur des images réelles texturées démontrent, avec les paramètres de textures retenus ici, la bonne robustesse de l'approche proposée.

Cette étude a été soutenue par le PRC «Communication Homme-Machine» (Pôle Vision), et par le Groupement de Recherche TdSI (GDR 134).

Manuscrit reçu le 26 mars 1993.

BIBLIOGRAPHIE

- [1] R. AZENCOTT. Markov fields and image analysis. In *6^e Congrès de Rec. des Formes et Intelligence Artificielle*, Antibes, France, 1183–1191, Novembre 1987.
- [2] J. BESAG. Spatial interaction and the statistical analysis of lattice systems. *J Royal Statist. Soc.*, Série B, Vol. 36, 192–236, 1974.
- [3] J. BESAG. On the statistical analysis of dirty pictures. *J Royal Statist. Soc.*, Vol. 48, Série B, No 3, 259–302, 1986.
- [4] A. BLAKE. Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, No 1, 2–12, Jan. 1989.
- [5] C. BOUMAN and B. LIU. Multiple resolution segmentation of textured images. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No 2, 99–113, Feb. 1991.
- [6] P. BRODATZ. *Textures : A Photographic Album for Artists and Designers*. Dover Publications, 1966.
- [7] P.B. CHOU and C.M. BROWN. The theory and practice of bayesian image modeling. *Int. J. Comp. Vis.*, Vol. 4, 185–210, 1990.
- [8] H. DERIN and H. ELLIOT. Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 9, No 1, 39–55, Jan. 1987.
- [9] M. DERRAS, M. BERDUCAT, P. BONTON, J. GALLICE, and R. CANALS. Segmentation texturale originale appliquée au guidage visuel d'un robot d'entretien d'espaces naturels. In *14^e colloque GRETSI*, Juan-Les-Pins, France, 771–774, Sept. 1993.
- [10] X. DESCOMBES. *Champs markoviens en analyse d'images*. PhD thesis, Telecom Paris, Rapport de présoutenance, 1993.
- [11] R.C. DUBES and A.K. JAIN. Random field models in image analysis. *J. Applied Statistics*, Vol. 16, No 2, 131–163, 1989.
- [12] Z. FAN and F.S. COHEN. Unsupervised texture image segmentation. In *Proc. Int. Congress Intelligent Autonomous Systems*, Amsterdam, Holland, Dec. 1986.
- [13] E. FRANÇOIS and P. BOUTHEMY. Multiframe-based identification of mobile components of a scene with a moving camera. In *IEEE Int. Conf. Computer Vision Pattern Recognition*, Hawaii, 166–172, June 3–6 1991.
- [14] D. GEMAN, S. GEMAN, C. GRAFFIGNE, and D. PONG. Boundary detection by constrained optimization. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No 7, 609–628, July 1990.
- [15] S. GEMAN and D. GEMAN. Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 6, No 6, 721–741, November 1984.
- [16] S. GEMAN, D. GEMAN, and C. GRAFFIGNE. Locating texture and object boundaries. In P.A. Devijver and J. Kittler, editors, *Pattern Recognition. Theory and Applications*, volume NATO ASI Series, Vol. F30, 165–177, Heidelberg, 1987. Springer Verlag.
- [17] C. GRAFFIGNE. *Experiments in Texture Analysis and Segmentation*. PhD thesis, Brown University, 1986.
- [18] R.M. HARALICK. Statistical and structural approaches to texture. *Proc. IEEE*, Vol. 67, No 5, 786–804, May 1979.
- [19] R.M. HARALICK. Statistical image texture analysis. In T.S. Young and K.S. Fu, editors, *Handbook of Pattern Recognition*, Academic Press, 247–279, 1986.
- [20] A. KHOTANZAD and J.Y. CHEN. Unsupervised segmentation of textured images by edge detection in multidimensional features. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 1, No 4, 414–421, 1989.
- [21] S. LAKSHMANAN and H. DERIN. Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, No 8, 799–813, 1989.
- [22] B.S. MANJUNATH and R. CHELLAPPA. Unsupervised texture segmentation using Markov Random Fields. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, 478–482, 1991.
- [23] J.L. MARROQUIN, S. MITTER, and T. POGGIO. Probabilistic solution of illposed problems in computational vision. *J. American Statist. Assoc.*, Vol. 82, 76–89, 1987.
- [24] P. MIGNOT, J. LEVY VEHEL, and Y. LECHEVALLIER. Arthur : un système d'analyse de texture. *Traitement du Signal*, Vol. 9, No 6, 507–517, 1992.
- [25] P. PEREZ and F. HEITZ. Une approche multiéchelle à l'analyse d'images par champs markoviens. *Traitement du Signal*, Vol. 9, No 6, 459–466, 1992.
- [26] F. PRETEUX and X. DESCOMBES. Synthèse et analyse de textures par coopération de processus multi-échelles. In *8^e congrès AFCET/INRIA Reconnaissance des Formes et Intelligence Artificielle*, Lyon, France, 1015–1022, 1991.
- [27] C.S. WON and H. DERIN. Unsupervised segmentation of noisy and textured images using Markov Random Fields. *CVGIP : Graphical Models and Image Processing*, Vol. 54, No 4, 308–328, July 1992.
- [28] J. ZHANG and J.W. MODESTINO. Unsupervised image segmentation using a Gaussian model. In *Proc. CISS 1988*, Princeton, NJ, 7–10, 1988.

LES AUTEURS

Charles KERVANN

Charles Kervann est né en 1969. Titulaire du DEA "Signal Télécommunications Image et Radar" à l'Université de Rennes I en 1992, il prépare depuis un an un doctorat en traitement du signal à l'IRISA, à Rennes.

Ses domaines d'intérêt couvrent la modélisation statistique de l'image, l'application des modèles markoviens à l'imagerie ainsi que la reconnaissance des formes en mouvement dans des séquences d'images par optimisation globale.

Fabrice HEITZ

Fabrice Heitz est chargé de recherches INRIA. Il mène actuellement ses travaux de recherche dans le projet TEMIS ("Traitement Exploitation et Modélisation d'Images Séquentielles") à l'IRISA, à Rennes. Diplômé de Télécom Bretagne (1984), il a préparé un doctorat en automatique et traitement du signal à Télécom Paris (1988). Ses premiers travaux ont porté sur l'analyse non destructive d'œuvres d'art par des techniques d'analyse statistique de l'image, en collaboration avec le Laboratoire de Recherches des Musées de France.

Ses domaines d'intérêt actuels couvrent la vision par ordinateur, la modélisation statistique de l'image, l'analyse du mouvement dans des séquences d'images ainsi que la parallélisation des algorithmes de traitement d'images.