

Quelques améliorations à la segmentation d'images bayésienne

Première partie : modélisation

Some Improvements to Bayesian Image Segmentation Part One : Modelling

par Xavier DESCOMBES, Robin MORRIS, Josiane ZERUBIA

INRIA, 2004, route des Lucioles, BP 93
06902 Sophia Antipolis, Cedex
email : Prénom.Nom@inria.fr

résumé et mots clés

Nous nous plaçons dans le cadre de la segmentation bayésienne. Parmi les trois étapes (modélisation, estimation, optimisation), nous considérons la modélisation et l'optimisation. La modélisation est appréhendée sous l'angle des champs de Markov. Nous montrons les limites du modèle de Potts couramment employé et proposons un nouveau modèle (le chien-modèle) permettant de contrôler la longueur des contours et des lignes dans l'image segmentée. Nous préservons ainsi les structures fines présentes dans les données. Nous comparons ensuite les critères MPM et MAP conjointement aux algorithmes qui permettent de les optimiser. Les différents résultats sont obtenus sur des images synthétiques et des images SPOT. Le problème de la classification fait l'objet d'une seconde partie.

Segmentation d'images, champs de Markov, modèle de Potts, chien-modèle, critères bayésiens.

abstract and key words

In this paper we address the segmentation problem in a Bayesian framework. Of the three stages (modelling, estimation, optimisation), we consider modelling and optimisation. We consider modelling by Markov random fields. We demonstrate the limitations of the Potts model currently employed, and propose a new model (the chien model) which allows us to control the boundary length and lines in the segmented images. We also preserve fine structures in the data. Then, we compare the MPM and MAP criteria when used with the algorithms discussed above. Results are presented on synthetic images and SPOT data. The classification problem is tackled in a second part.

Image segmentation, Markov random fields, potts model, chien-model, bayesian framework.

1. Introduction

Nous nous plaçons dans un contexte bayésien. Considérons une distribution de probabilité quelconque $P(\omega)$ sur un espace des configurations noté Ω . Nous définissons alors une fonction de coût entre deux configurations $R(\omega, \omega')$. L'estimation bayésienne consiste alors à minimiser le risque bayésien, c'est-à-dire à rechercher la configuration qui minimise l'espérance de la fonction de coût :

$$\hat{\omega} = \arg \min_{\omega'} \int_{\omega} R(\omega, \omega') P(\omega) d\omega. \quad (1)$$

Traiter un problème dans le cadre de l'estimation bayésienne revient donc à définir :

- Une distribution $P(\omega)$ sur l'espace des configurations Ω
- Une fonction de coût $R(\omega, \omega')$
- Un algorithme de minimisation

Considérons maintenant le contexte de la segmentation d'images. Le champ des données est noté $X = (x_s)_{s \in S}$ où S représente la trame de l'image (ensemble des sites) et les différents x_s sont à valeur dans l'espace d'état L (ensemble des niveaux de gris). L'image segmentée est notée $Y = (y_s)_{s \in S}$, les différents y_s étant à valeur dans Λ (ensemble des classes). La modélisation markovienne permet de définir une loi *a priori* sur l'espace des

configurations Λ^S , notée $P(Y)$. D'autre part, une modélisation des données (texture, bruit, ...) permet de définir une loi d'attache aux données sur l'espace des configurations L^S , notée $P(X|Y)$. L'inconnue du problème étant Y , nous cherchons à « optimiser » la loi *a posteriori* définie par :

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}. \quad (2)$$

X étant donné, nous optimisons en réalité :

$$P(X|Y)P(Y). \quad (3)$$

Le choix de l'algorithme d'optimisation et celui de la fonction de coût sont bien évidemment liés. Nous rappelons deux critères couramment utilisés que nous comparerons dans la suite de ce travail. Le MAP (Maximum *A Posteriori*) correspond à la fonction de coût définie par :

$$R_{MAP}(\omega, \omega') = 1 - \delta_{\omega'}(\omega), \quad (4)$$

où δ représente le symbole de Kronecker. L'optimum est alors atteint pour la (une) configuration qui maximise la loi. Un tel critère est difficile à atteindre et nécessite très souvent un algorithme coûteux en temps de calcul comme le recuit simulé par exemple [8]. Des approximations de ce critère peuvent néanmoins être obtenues par des algorithmes déterministes [2], [4], [5]. Le MPM (Modes Marginales *a Posteriori*) correspond à la fonction de coût définie par :

$$R_{MPM}(\omega, \omega') = \sum_{s \in S} 1 - \delta_{\omega'}(\omega_s). \quad (5)$$

L'optimum est atteint lorsque les différentes lois marginales en un site sont maximisées. Du point de vue algorithmique, les méthodes de type Metropolis-Hasting (Markov Chain Monte Carlo) sont utilisées pour optimiser un tel critère [11], [10]. Le critère MPM minimise alors le nombre de pixels mal classés et est donc adapté aux tâches de segmentation. En revanche lorsque, pour des contraintes de temps de calcul, une approximation du critère est recherchée par des algorithmes déterministes, le critère MAP est souvent préféré. Le critère MPM est plus difficilement approché par des algorithmes déterministes.

Le cadre bayésien fait intervenir deux termes : modèle *a priori* et terme d'attache aux données. Le résultat d'une segmentation dépend bien évidemment du terme d'attache aux données et de son couplage avec le modèle *a priori* [14]. Néanmoins, le modèle *a priori* choisi doit lui-même refléter les propriétés recherchées dans l'image segmentée indépendamment de l'attache aux données pour obtenir des résultats précis et robustes. Le paragraphe 2 montre en premier lieu les limites du modèle de Potts, souvent utilisé comme loi *a priori* pour des problèmes de segmentation [8], [7]. Ces limites motivent la recherche de modèles plus complexes rendant mieux compte des différentes propriétés de l'image segmentée recherchée. Un premier modèle, fondé sur des cliques 3×3 est proposé dans [13] alors qu'un modèle sur une trame hexagonale est décrit dans [1]. Ces modèles considèrent les différentes configurations indépendamment les unes des autres. Dans le paragraphe 3, nous définissons un modèle à partir des énergies locales, pour relier les différentes configurations formant un contour ou une ligne [15]. Une version 3D du modèle est présentée dans [6].

Ce modèle est testé puis comparé au modèle de Potts à partir d'images de synthèse et d'images SPOT dans le paragraphe 4. Nous montrons ensuite les raisons qui nous font préférer le critère bayésien MPM au critère MAP pour rechercher la solution. Le problème de l'estimation du nombre de classes et des paramètres qui leur sont associés est traitée dans une seconde partie [17].

2. limites du modèle de Potts

Dans ce paragraphe, nous considérons le modèle markovien le plus souvent choisi comme loi *a priori* pour un problème de segmentation : le modèle de Potts. Nous montrons les limites de ce modèle et son manque d'adéquation avec le résultat recherché.

Nous considérons un modèle de Potts sans champ externe. Les cliques d'un tel modèle sont alors constituées de paires de pixels voisins. La loi d'un tel modèle s'écrit alors :

$$P(Y) = \frac{1}{Z(\beta)} \exp \left(+\beta \sum_{c=\{y_s, y'_s\}, c \in \mathcal{C}} \delta_{y_s=y'_s} - \delta_{y_s \neq y'_s} \right), \quad (6)$$

où \mathcal{C} désigne l'ensemble des cliques.

Le modèle de Potts dépend donc d'un paramètre : β . La probabilité d'une configuration est d'autant plus élevée que le nombre de cliques homogènes (paires de voisins ayant la même étiquette) est important. En outre, la loi du modèle peut s'écrire sous la forme :

$$P(Y) = \frac{1}{Z(\beta)} \exp +2\beta \#_Y, \quad (7)$$

où $\#_Y$ représente le nombre de cliques homogènes de la configuration Y .

Avec un tel modèle, nous pouvons donc espérer contrôler la longueur des contours ou encore la taille des régions des réalisations obtenues. En simulant différentes (un grand nombre!) réalisations du modèle, nous pouvons alors estimer la moyenne du nombre de cliques homogènes, notée $\langle \#_Y \rangle$, en fonction de la valeur du paramètre β . Cette courbe dépend de la taille de la réalisation et du nombre de classes considérées. Néanmoins, le comportement de la courbe reste identique. Un exemple de cette courbe est illustré sur la figure 1. La courbe présente une forte pente autour de la valeur critique du paramètre [16].

Sur la figure 2, des réalisations d'un modèle de Potts à 4 classes sont présentées pour différentes valeurs de β . Le modèle de Potts ne permet donc pas de modéliser des images comportant des zones homogènes du type des images segmentées désirées. En effet, pour des valeurs de β inférieures à la valeur critique, les réalisations obtenues sont fortement bruitées. Pour des valeurs supérieures à ce point critique, un des états devient rapidement prédominant et les réalisations sont constituées d'un fond homogène avec plus ou moins de bruit. Cette rupture de symétrie entre les différents états correspond à une transition de phase du modèle de Potts et a été abondamment étudiée en physique statistique [9]. Du point de vue du traitement des images, elle signifie que le modèle de Potts est mal adapté pour générer des images formées de régions homogènes.

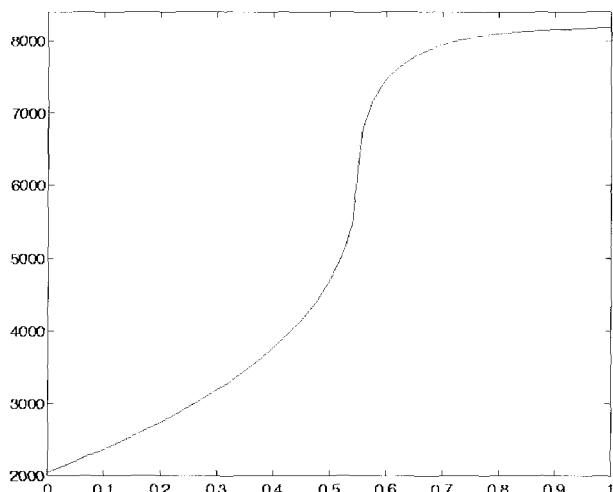


Figure 1. – Nombre de cliques homogènes en fonction de β pour un modèle de Potts à 4 classes sur une trame 64×64 .

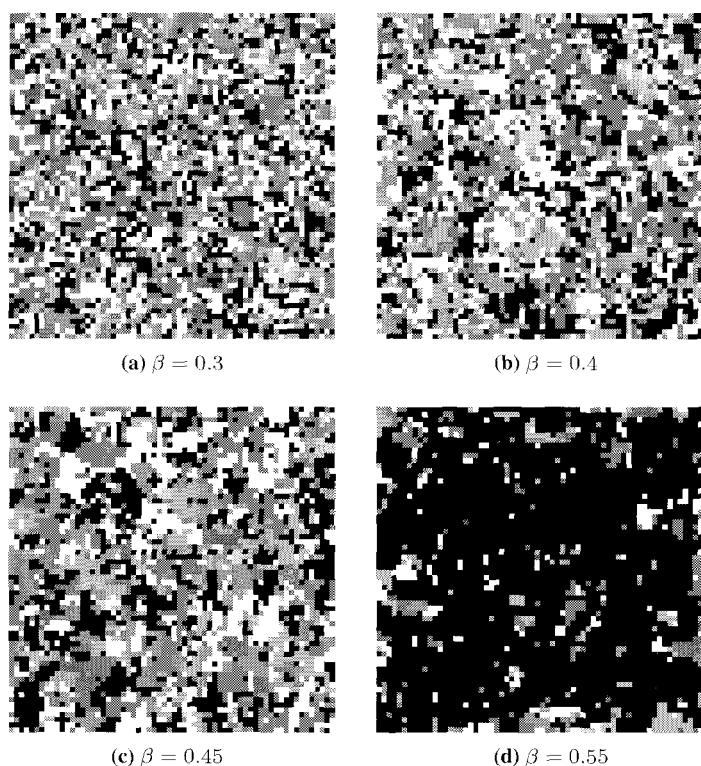


Figure 2. – Réalisations du modèle de Potts.

En pratique, ce modèle est malgré tout très utilisé comme terme régularisateur. Dans certains cas, les résultats sont acceptables car le terme d'attache aux données est suffisamment prépondérant et permet d'obtenir des régions homogènes. Ceci reste néanmoins valable uniquement si les données sont peu bruitées. Dans ce cas, l'utilisation d'une modélisation markovienne peut alors paraître inutilement sophistiquée. En présence de données bruitées, l'utilisateur est amené à utiliser une valeur de β nettement supérieure à la valeur critique, dès lors les structures fines de l'image sont perdues et la segmentation résultante est relativement grossière. Si

nous considérons la segmentation recherchée comme la réalisation d'un modèle de Potts et estimons le paramètre associé par un critère de maximum de vraisemblance, la valeur estimée s'avère bien inférieure à la valeur utilisée en pratique pour obtenir une segmentation exempte de bruit. La figure 3 représente une photographie aérienne bruitée. Une segmentation manuelle est effectuée (voir figure 3.b). Cette segmentation est alors considérée comme la réalisation d'un modèle de Potts, ce qui conduit, par une estimation au sens du maximum de vraisemblance, à une valeur du paramètre β égale à 0.544. La figure 3.c représente une simulation du modèle de Potts avec le paramètre estimé. Nous pouvons remarquer que nous sommes loin d'une image formée de régions homogènes. La présence de bruit se retrouve alors sur le résultat de la segmentation obtenue avec une valeur de β effectivement égale à 0.544 (voir la figure 3.d). En prenant une valeur du paramètre β plus forte, nous obtenons alors une segmentation grossière où de nombreux détails ont été perdus (figure 3.e). En résumé, le modèle de Potts ne semble pas bien adapté à des tâches

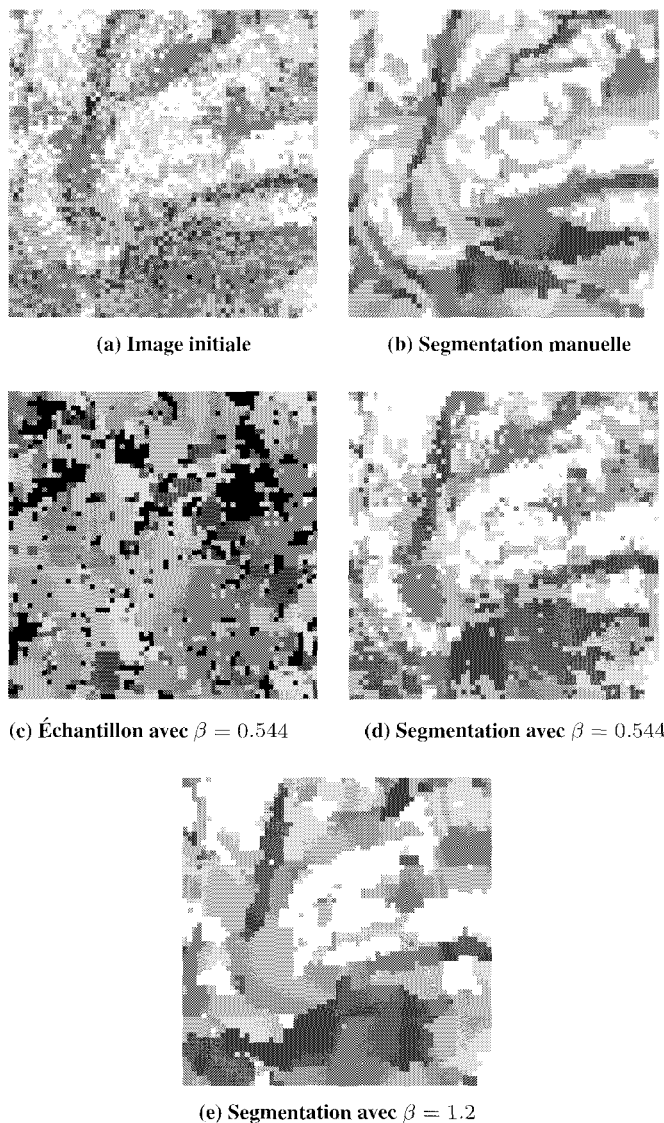


Figure 3. – Réalisations du modèle de Potts.

de segmentation d'images [16]. En réalité, ce modèle ne fait pas la distinction entre contour et bruit. Avec ce modèle, régulariser signifie minimiser le nombre de cliques inhomogènes, c'est-à-dire minimiser la longueur des contours, ce qui n'est pas identique à minimiser le bruit. D'autre part, le comportement du nombre moyen de cliques homogènes en fonction du paramètre β autour du point critique est un inconvénient majeur du modèle. En effet, les images « intéressantes », c'est-à-dire les images formées de régions homogènes, correspondent au sens de l'estimation de β , à des valeurs de β très proches les unes des autres (proches de la valeur du point critique). Le modèle de Potts permet donc en pratique de répartir les images en trois classes : le bruit, les images quasi-uniformes et « le reste du monde ». La valeur du paramètre ne reflète donc aucune propriété intéressante de l'image. Pour affiner les résultats des segmentations d'image, il semble donc nécessaire de développer des modèles plus sophistiqués pour lesquels les différents paramètres permettent de classer les images suivant des propriétés déterminées comme la longueur des contours (interfaces entre plusieurs régions homogènes), la présence et la longueur des lignes (régions homogènes d'épaisseur 1 pixel) ou encore la taille moyenne des régions homogènes. Ces modèles nécessitent de considérer des cliques contenant plus de pixels pour lesquelles les différentes configurations sont réparties en plusieurs classes.

3. le Chien-modèle

Le modèle présenté dans ce paragraphe dépend de trois paramètres respectivement notés e , l et n qui font référence aux contours (« edge »), aux lignes (« line ») et au bruit (« noise »). Le principe de construction du modèle se décompose en plusieurs étapes. Nous proposons en premier lieu un modèle binaire, étendu au cas m -aire par la suite. Nous choisissons un système de cliques : dans le cas présent, les cliques sont constituées des blocs carrés de trois pixels sur trois. Cette taille correspond à la taille minimale nécessaire à la définition de contours et de lignes. D'autre part, une taille plus grande accroît considérablement la complexité du modèle. Nous effectuons alors une classification des différentes configurations binaires possibles pour les cliques choisies. A chaque classe est associé un paramètre, valeur du potentiel correspondant. Le nombre des paramètres obtenus est encore trop grand pour être utilisable en pratique. En outre, le lien entre ces paramètres et les propriétés des réalisations n'est pas immédiat. La seconde étape consiste donc à relier ces paramètres avec des quantités facilement interprétables de l'image par l'intermédiaire d'une analyse des énergies locales.

3.1. construction du modèle binaire

Considérons donc les $2^9 = 512$ configurations binaires possibles sur une clique de 3 pixels par 3. Une première classification est obtenue en considérant que deux configurations sont équivalentes si elles se déduisent l'une de l'autre soit par la symétrie noir-blanc, soit par une rotation. Nous obtenons alors 51 classes d'équivalence

dont un représentant, et le nombre d'éléments de chaque classe, sont présentés sur la figure 4.

A chaque classe est associé un paramètre définissant la valeur du potentiel des configurations de la classe. Ainsi, sous les hypothèses d'isotropie du modèle et de symétrie noir-blanc, nous avons pour une telle topologie (cliques de 3×3 pixels) cinquante et un degrés de liberté. L'énergie étant définie à une constante près, nous imposons un potentiel nul aux configurations uniformes ($C(1) = 0$), ce qui ramène à cinquante le nombre de degrés de liberté. La construction du modèle consiste alors à imposer des contraintes en reliant les différents paramètres entre eux. Nous considérons que les états fondamentaux (états d'énergie minimale) sont donnés par (au moins) les configurations uniformes. De sorte que la méthodologie de construction du modèle peut se résumer ainsi :

- Modéliser les contraintes sur les réalisations globales par les énergies locales et en déduire les équations sur les paramètres définis par la figure 4.
- Vérifier la cohérence du système d'équations et le résoudre.
- Vérifier qu'aucune réalisation n'est d'énergie négative.

Le premier type de contraintes concerne la pénalisation des contours. Nous imposons le coût d'un contour par une énergie par

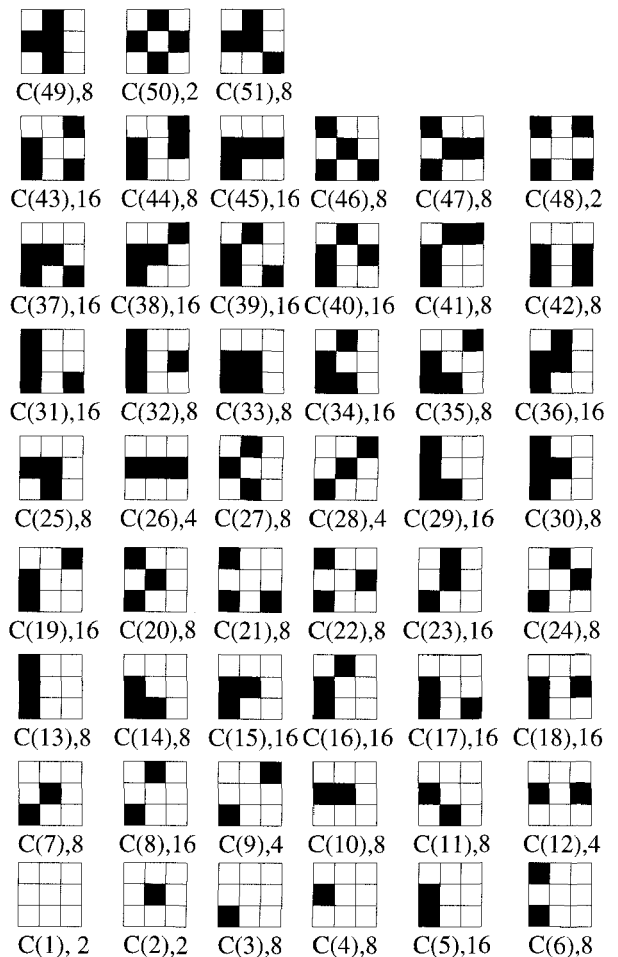


Figure 4. – Classes d'équivalence des configurations binaires 3×3 et nombre d'éléments dans chaque classe.

unité de longueur, notée e . Avec la taille élémentaire choisie (taille des cliques), nous avons huit directions possibles (voir figure 5). En considérant le modèle symétrique par rotation, à ces directions correspondent trois types de contours, à savoir vertical, oblique et diagonal. Nous imposons à chaque direction une même énergie par unité de longueur pour obtenir un modèle isotrope. Aux trois types de contours sont alors associées les trois équations suivantes, représentées sur la figure 6 :

$$2C(13) = e \quad (8)$$

$$2C(3) + 2C(14) = \sqrt{2}e \quad (9)$$

$$C(5) + C(29) + C(3) = \frac{\sqrt{5}}{2}e \quad (10)$$

Ces contraintes sont définies pour les objets de largeur au moins égale à 3. Pour les autres objets, nous définissons des contraintes similaires. Pour les objets d'épaisseur 1, nous introduisons un paramètre supplémentaire l déterminant l'énergie par unité de longueur des lignes, ce qui induit trois nouvelles équations (cf figure 7). Les objets d'épaisseur 2 et 3 sont définis comme ayant une énergie $2e$ par unité de longueur correspondant aux bords droit et gauche. Les équations obtenues s'écrivent alors :

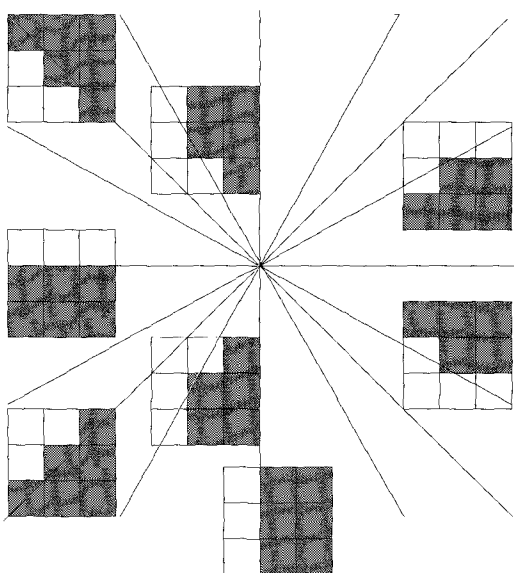


Figure 5. - Orientations définies dans le chien-modèle.

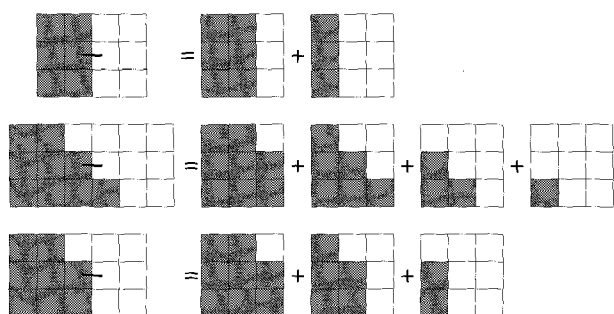


Figure 6. - Équations correspondant aux contraintes de contours.

- pour les objets verticaux (et horizontaux) :

$$2C(13) + C(26) = l \quad (11)$$

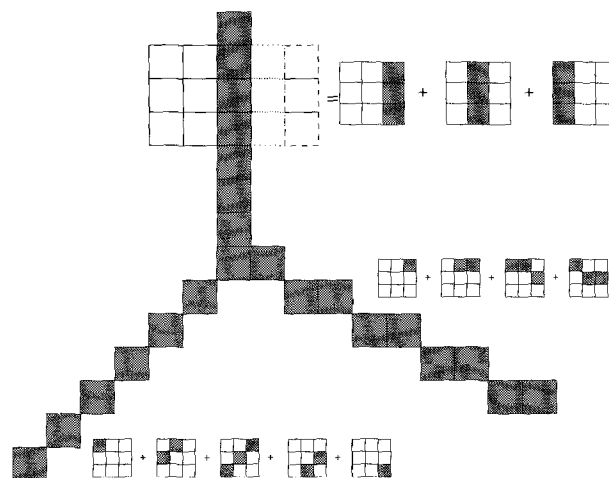


Figure 7. - Définition des lignes dans le chien-modèle.

$$4C(13) = 2e \quad (12)$$

- pour les objets diagonaux :

$$2C(3) + 2C(11) + C(28) = \sqrt{2}l \quad (13)$$

$$2C(3) + 2C(14) + 2C(35) = 2\sqrt{2}e \quad (14)$$

$$2C(3) + 4C(14) + C(9) = 2\sqrt{2}e \quad (15)$$

- pour les objets obliques :

$$C(3) + C(16) + C(23) + C(5) = \frac{\sqrt{5}}{2}l \quad (16)$$

$$C(3) + 2C(29) + C(19) + C(5) = \sqrt{5}e \quad (17)$$

$$2C(5) + 2C(29) + 2C(3) = \sqrt{5}e \quad (18)$$

Ces différentes contraintes se ramènent à onze équations faisant intervenir quatorze paramètres. Il nous faut maintenant vérifier que le système possède bien des solutions. Deux paramètres ($C(3)$ et $C(5)$) jouent un rôle particulier dans ce système puisqu'ils interviennent dans la plupart des équations. Nous exprimons alors les équations en fonction de ces deux paramètres pour obtenir le système suivant :

$$C(13) = \frac{e}{2} \quad (19)$$

$$C(14) = \frac{\sqrt{2}}{2}e - C(3) \quad (20)$$

$$C(26) = l - e \quad (21)$$

$$C(29) = \frac{\sqrt{5}}{2}e - C(3) - C(5) \quad (22)$$

$$2C(11) + C(28) = \sqrt{2}l - 2C(3) \quad (23)$$

$$C(35) = \frac{\sqrt{2}e}{2} \quad (24)$$

$$C(9) = 2C(3) \quad (25)$$

$$C(16) + C(23) = \frac{\sqrt{5}}{2}l - C(3) - C(5) \quad (26)$$

$$C(19) = C(3) + C(5) \quad (27)$$

Nous résolvons ce système en nous assurant qu'aucune configuration globale n'est d'énergie négative.

Pour ce faire, nous considérons une solution pour laquelle aucune configuration locale n'est négative. Les configurations n'apparaissant pas dans le système d'équations sont considérées comme définissant le bruit et ont donc une énergie égale à n qui définit le troisième paramètre du modèle.

Pour des valeurs strictement positives des paramètres, nous avons uniquement les deux configurations uniformes comme états fondamentaux. En revanche si $e = l = 0$, toute configuration composée de bandes parallèles ou d'objets circulaires est un état fondamental.

L'inconvénient de ce modèle par rapport au modèle d'Ising réside essentiellement dans sa complexité. L'utilisation de cliques 3×3 entraîne un voisinage de 5×5 pixels. Le temps de calcul nécessaire à l'optimisation du modèle est donc accru par rapport à un simple modèle d'Ising 4 ou 8 connexe. En revanche, le modèle proposé est beaucoup plus souple puisqu'il fait la distinction entre bruit, lignes et contours. Il représente donc un *a priori* plus fin sur la solution de la segmentation, ce qui sera montré sur les résultats expérimentaux.

3.2. extension à un modèle m-aire

Dans ce paragraphe, nous étendons le chien-modèle binaire à un modèle m-aire. De même que pour le cas binaire, nous définissons un coût énergétique pour les contours, les lignes et le bruit respectivement. Afin de généraliser le cas binaire, deux approches peuvent être envisagées. La première consiste à pénaliser le désordre, la seconde à favoriser l'ordre. Explicitons ces deux approches sur l'exemple de la figure 8. Dans la première approche, cette configuration sera pénalisée du fait des pixels de la partie droite. En revanche, elle sera favorisée dans la seconde approche en vertu de l'uniformité de la partie gauche. Dans la première approche, certaines configurations de forte énergie (phases désordonnées) peuvent représenter de profonds minimums locaux et poser des problèmes pour l'algorithme d'optimisation. En effet, certaines configurations seraient de forte énergie mais nécessiteraient le changement de labels de plusieurs pixels pour décroître en énergie. Pour éviter d'avoir des minimums locaux trop profonds, nous adoptons donc ici la seconde approche.

Soient X une configuration sur une clique définie au paragraphe 3.1 et $\lambda_i, i \in \{1, \dots, n\}$ les différentes étiquettes présentes dans la configuration. Considérons les différentes configurations binaires X_i obtenues en associant l'état 1 à tous les sites ayant pour étiquette λ_i et l'état 0 aux autres sites. Le potentiel $V_b(X_i)$ associé à la configuration X_i est alors celui correspondant au cas du modèle

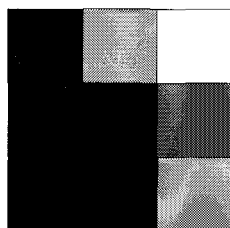


Figure 8. – Configuration m-aire contenant une partie ordonnée et une partie désordonnée.

binaire. Le potentiel associé à la configuration X est alors donné par la formule (voir figure 9) :

$$V(X) = \sum_{i \in \{1, \dots, n\}} V_b(X_i). \quad (28)$$

Sur la figure 10, nous présentons des simulations du chien-modèle pour différentes valeurs des paramètres. Bien que différentes des images réelles segmentées, ces images reflètent certaines propriétés structurelles (longueur des contours, nombre de lignes) permettant d'adapter les paramètres du modèle aux structures de l'image traitée, ce qui n'est pas le cas pour le modèle de Potts.



Figure 9. – Généralisation du modèle binaire au cas m-aire.

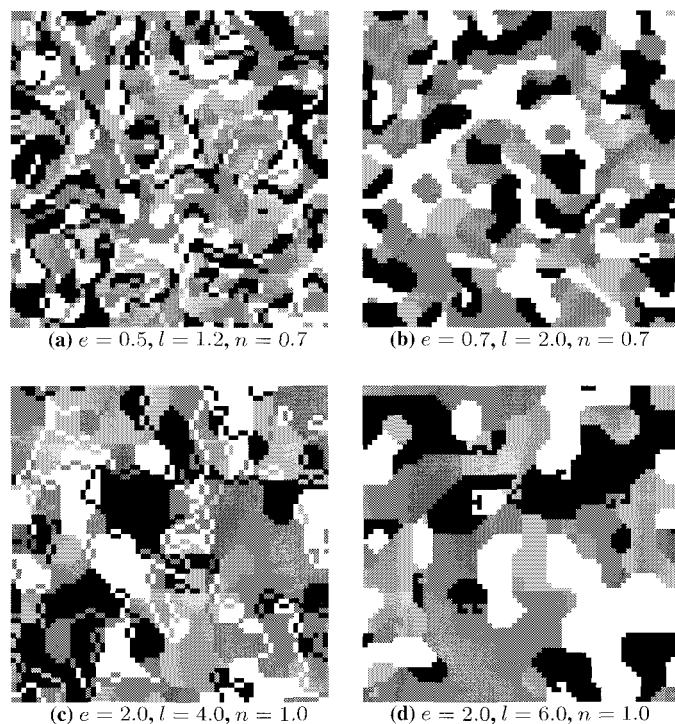


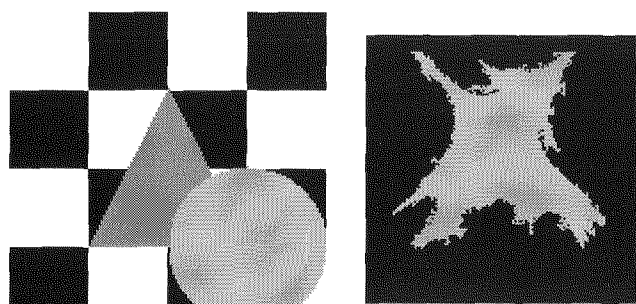
Figure 10. – Réalisations du chien-modèle.

4. résultat obtenus par les deux modèles

Dans ce paragraphe, nous comparons les deux modèles en utilisant une segmentation supervisée. Nous considérons comme connus le nombre de classes et les paramètres qui leur sont associés. En ce qui concerne l'image SPOT, des ensembles d'apprentissage ont été définis manuellement par un expert. Nous reprenons alors les paramètres donnés dans [12]. Pour chacun des deux modèles, nous

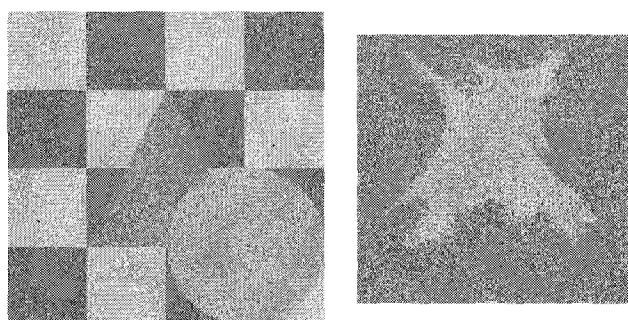
utilisons le critère bayésien qui donnent les meilleurs résultats et que nous expliciterons par la suite.

Les deux premières images sont des images de synthèse pour lesquelles nous connaissons donc « la solution » (voir figure 11). Nous pouvons alors déterminer le nombre de pixels mal classés et les localiser. La première contient des formes aux contours lisses alors que la seconde image contient une forme aux contours très chahutés, ce qui va permettre de tester la finesse des différents modèles de régularisation. Ces images sont bruitées par un bruit additif gaussien (voir figure 12). Pour finir, une troisième image permet de valider ce travail sur des données réelles. Il s'agit d'une image SPOT (canal XS3) de la Hollande (voir figure 13). La scène représente une zone de parcellaire, possédant des structures très fines (routes, chemins, champs étroits,...).



(a) Image 256×256
Figure 11. – Images tests de synthèse.

(b) Image 128×128



(a) Image 1.a bruitée : SNB = 5dB

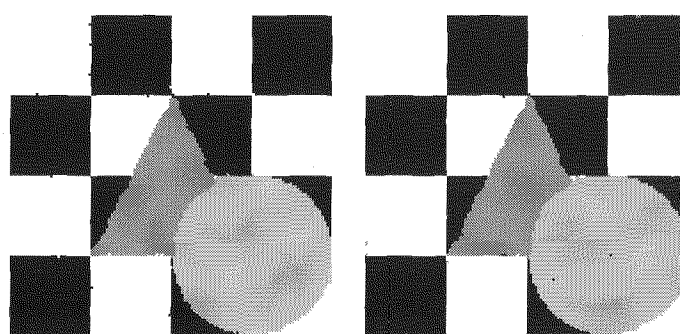
(b) Image 1.b bruitée : SNB = 5dB

Figure 12. – Images tests bruitées.



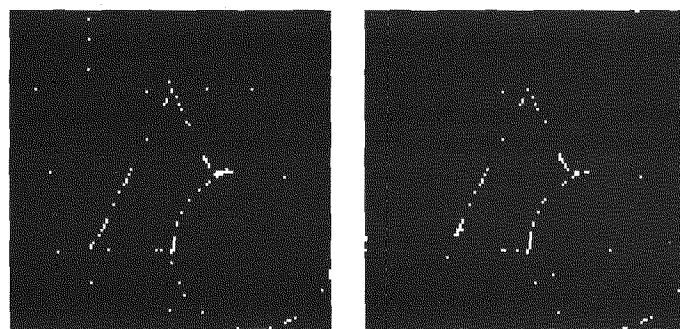
Figure 13. – Image SPOT (canal XS3).

La première image test donne des résultats comparables pour les deux modèles (voir figure 14). En effet, cette image ne possède pas de structures fines et peut donc être traitée correctement par le modèle de Potts. Il est intéressant de noter que le pouvoir régularisateur du chien-modèle est aussi efficace que celui du modèle de Potts. Pour finir, notons que le chien-modèle conduit à 85 pixels mal classés alors que nous en avons 99 pour le modèle de Potts. La différence provient essentiellement des contours de l'échiquier.



a) Modèle de Potts.

b) Chien-modèle.



c) Modèle de Potts :
image des erreurs.

d) Chien-modèle :
image des erreurs.

Figure 14. – Résultats obtenus pour l'image de synthèse # 1.

La seconde image test présente une forme aux contours très accidentés. Le modèle de Potts a tendance à lisser les contours pour minimiser leur longueur, ce qui conduit à 342 pixels mal classés (voir figure 15.a). Le chien-modèle régularise sans trop pénaliser les contours et permet alors de réduire le nombre de pixels mal classés à 299 (voir figure 15.b).

Pour finir, nous présentons les résultats obtenus sur l'image SPOT. Les classes ont été définies à partir d'ensembles d'apprentissage déterminés par un expert. Nous pouvons remarquer sur les figures 16.a et 16.b que le chien-modèle préserve mieux les structures fines comme par exemple les chemins entre les différents champs (structures linéiques). Pour des images plus bruitées, comme les images radar de type RSO par exemple, le modèle *a priori* prend encore plus d'importance. L'avantage du chien-modèle par rapport au modèle de Potts devrait s'avérer encore plus manifeste dans ce cadre.

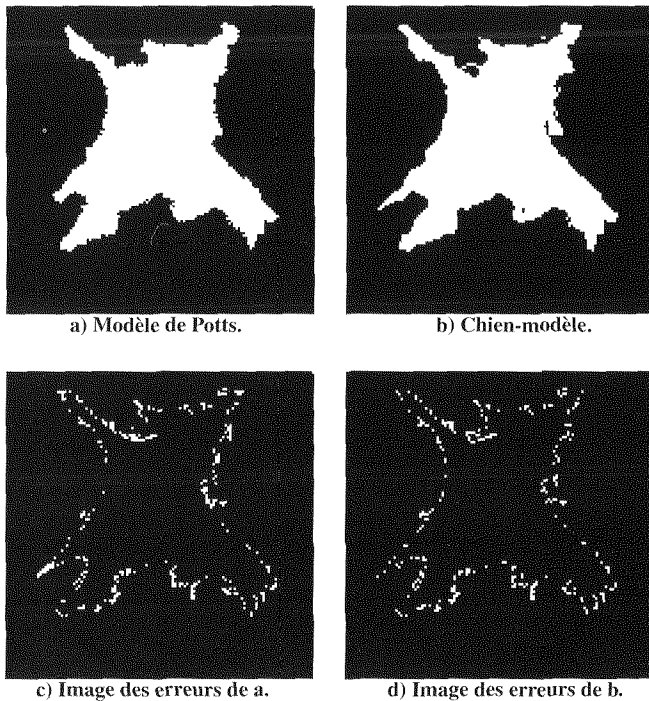


Figure 15. – Résultats obtenus pour l'image de synthèse # 2.

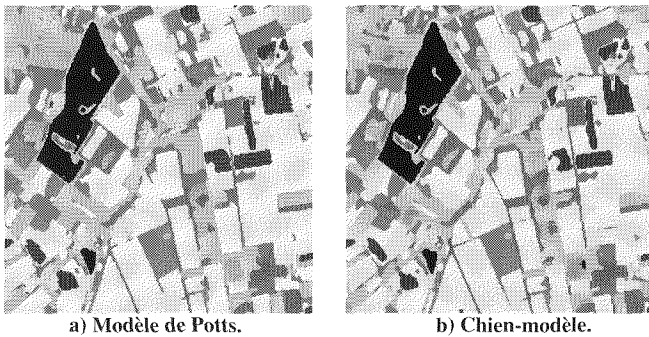


Figure 16. – Résultats obtenus pour l'image de synthèse # 2.

5. à propos du choix du critère d'optimisation

Les différents résultats présentés jusqu'ici ont été obtenus par un critère MPM. Pour converger vers ce critère, il suffit d'échantillonner le processus par un échantillonneur de Gibbs par exemple. Chaque pixel prend alors pour valeur l'étiquette la plus représentée en ce point dans les différents échantillons. Une alternative, couramment employée, consiste à estimer le MAP par un recuit simulé. Optimiser le MAP revient à rechercher une configuration d'énergie globale minimale. Néanmoins, les schémas théoriques de décroissance de la température sont beaucoup trop lents pour être envisageables en pratique. Une décroissance géométrique de la température est le plus souvent utilisée, ce qui conduit à

une approximation du critère MAP. D'autres approximations sont également obtenues avec des algorithmes déterministes comme l'ICM [2] ou le GNC [4]. Le MPM semble plus intuitif puisqu'il permet de minimiser le nombre de pixels mal classés. En revanche, le critère MAP peut être approximé par de nombreux algorithmes déterministes peu coûteux en temps de calcul. Nous comparons ici quelques résultats obtenus avec les deux critères bayésiens : MAP et MPM.

Les résultats obtenus par les deux critères sont comparables lorsque le modèle de régularisation est le modèle de Potts (voir figure 17). En revanche, en utilisant le chien-modèle, nous avons une énergie plus complexe, comportant de forts minimums locaux. Par suite, une décroissance trop rapide de la température lors d'un recuit simulé piège la configuration dans un minimum local. Le critère du MAP n'est pas à remettre en cause car le minimum global de l'énergie correspond bien aux résultats souhaités. Néanmoins, même un algorithme comme le recuit simulé possédant une propriété théorique de convergence vers un minimum global de l'énergie n'est dans la pratique qu'une approximation du critère MAP. Cette approximation s'avère trop grossière dans le cas d'énergies complexes. Le MPM est alors préférable à une approximation du MAP comme le montrent les résultats de la figure 18. Une région de quelques pixels égaux à 1 peut représenter un fort minimum local et apparaître avec une approximation du MAP. En utilisant un échantillonneur de Gibbs cette région apparaît dans certaines réalisations mais pas dans la majorité. Le critère MPM permet alors de la supprimer. Le chien-modèle montre donc encore sa supériorité mais est sensible en présence d'un fort taux de bruit aux approximations d'un algorithme.

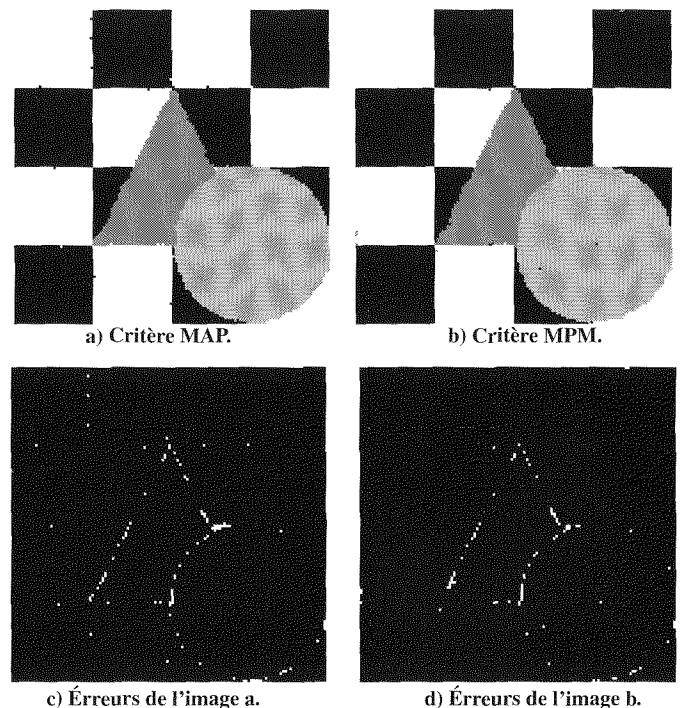


Figure 17. – Résultats obtenus pour l'image de synthèse # 1 avec un modèle de Potts.

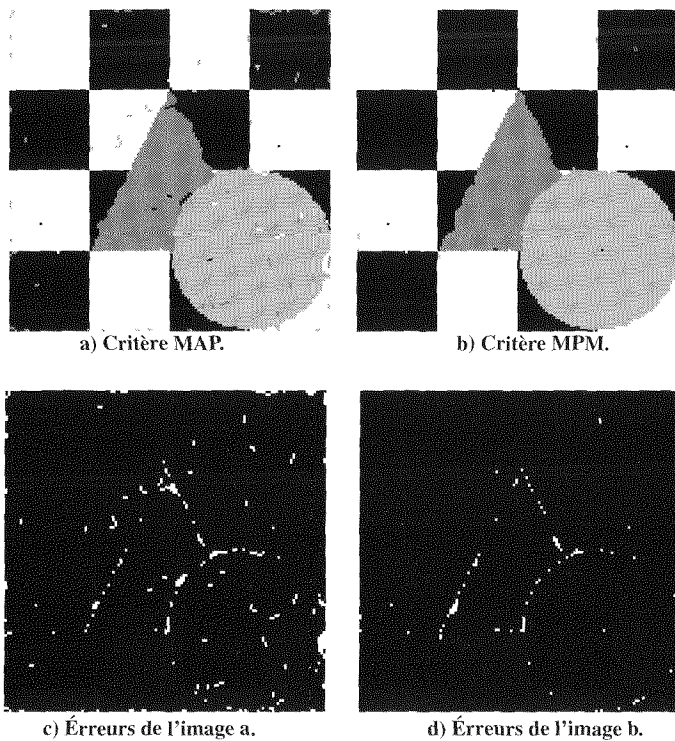


Figure 18. – Résultats obtenus pour l'image de synthèse # 1 avec un chien-modèle.

6. conclusion

Dans cet article, nous avons montré les limites du modèle de Potts utilisé comme loi *a priori* en segmentation d'images. Nous avons défini un nouveau modèle *a priori* palliant ces limites : le chien-modèle. Ce modèle, dont la version binaire est décrite dans [15], permet de contrôler l'énergie associée aux contours et aux lignes d'une réalisation indépendamment de l'énergie liée au bruit. Tout en conservant des propriétés régularisantes, ce modèle permet de préserver les structures fines de la segmentation. L'apport de ce modèle, comparativement au modèle de Potts, est montré à la fois sur des images de synthèse et sur des données réelles (image SPOT).

Cet article consacré aux modèles *a priori* et aux critères bayésiens d'optimisation ne traite pas de l'estimation des paramètres. Le problème de l'estimation du nombre de classes et des paramètres associés à ces classes, permettant de définir le terme d'attache aux données est un problème de classification. Les algorithmes de type K-moyennes ou Isodata permettent donc d'estimer ces paramètres. Une solution utilisant une information de type contextuel est proposée dans [17].

Pour obtenir une méthode de segmentation totalement non-supervisée, il reste à estimer les paramètres du modèle *a priori*. Certaines méthodes sont proposées dans la littérature ([3], [7]) pour estimer les paramètres d'un tel modèle. Ces méthodes permettent d'obtenir des approximations du maximum de vraisemblance tels que les estimateurs au sens du pseudo-maximum de

vraisemblance. De telles approximations donnent des résultats honorables pour le modèle de Potts mais se sont révélées inefficaces pour un modèle plus complexe comme le chien-modèle. Nos recherches s'orientent donc vers des méthodes d'estimation au sens du maximum de vraisemblance.

Pour finir, nous avons comparé les deux critères bayésiens : MPM et MAP. Les résultats obtenus nous font préférer le MPM dont la convergence est mieux contrôlée du point de vue algorithmique que pour le MAP. Cet avantage est d'autant plus marquant pour des modèles dont l'énergie présente de forts minimums locaux, comme c'est le cas pour le chien-modèle.

7. remerciements

Les auteurs tiennent à remercier le GdR TdSI (CNRS) pour l'image de synthèse n° 2 ainsi que le CNES pour l'image SPOT. Les commentaires des relecteurs ont permis d'améliorer le texte, qu'ils en soient remerciés ici.

BIBLIOGRAPHIE

- [1] H. Tjelmeland, J. Besag, Markov Random Fields with higher order interactions. submitted to JASA, Preprint.
- [2] J. Besag, Spatial Interaction and Statistical Analysis of Lattice Systems. *J. Royal Statistical Society, Series B*, 36 : 721-741, 1974.
- [3] J. Besag, On the statistical analysis of dirty pictures. *J. Royal Statistical Society, Series B*, 48 : 259-302, 1986.
- [4] A. Rangarajan, R. Chellappa, Generalised Graduated Non-Convexity algorithm for Maximum A Posteriori image estimation. in *ICPR*, pages 127-133, June 1990.
- [5] J. Zerubia, R. Chellappa, Mean field approximation using Compound Gauss Markov Random Field for edge detection and image restoration. In *Proc. ICASSP, Albuquerque, USA*, pages 2193-2196, 1990.
- [6] X. Descombes, Application of stochastic techniques in Image Processing for automatic tissue classification in MRI and blood vessel restoration in MRA. Technical Report KUL/ESAT/MI2/9603, Laboratory for Medical Imaging Research (ESAT-Radiology), K.U. Leuven, Belgium, 1996.
- [7] H. Derin, H. Elliott, Modelling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 9 (1) : 39-55, January 1987.
- [8] S. Geman, D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 6 (6) : 721-741, 1984.
- [9] H.O. Georgii, *Gibbs Measures and Phase Transitions*. De Gruyter - Studies in Mathematics, vol. 9, 1988.
- [10] P.J. Green, Monte Carlo methods : an overview. In *Proceedings of the IMA Conference on Complex Stochastic Systems and Engineering*, September 1993.
- [11] W.K. Hastings, Monte Carlo sampling methods using Markov Chains and their applications, *Biometrika*, 57/ 97-109, 1970. 97-109,

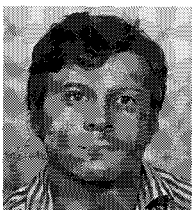
Améliorations à la segmentation d'images : modélisation

- [12] Z. Kato, *Modélisations markoviennes multirésolution en vision par ordinateur. Application à la segmentation d'images SPOT*. PhD thesis, Université de Nice-Sophia Antipolis, December 1994.
- [13] G. Wolberg, T. Pavlidis. Restoration of binary images using stochastic relaxation with annealing, *Pattern Recognition Letters*, 3 : 375-388, 1985.
- [14] A. Maruani, E. Pechersky, M. Sigelle. On Gibbs Fields in Image Processing, *Markov Processes and Related Fields*, 1(3) : 419-442, 1995.
- [15] X. Descombes, J.F. Mangin, E. Pechersky, M. Sigelle. Fine structures preserving model for image processing. In *Proc. 9th SCIA 95 Uppsala, Sweden*, pages 349-356, 1995.
- [16] R.D. Morris, X. Descombes, J. Zerubia. The Ising/Potts model is not well suited to segmentation tasks. In *proc. IEEE Digital Signal Processing Workshop*, 1996. Sept. 1-4 Norway.
- [17] X. Descombes, R. Morris, J. Zerubia. Quelques améliorations à la segmentation d'images bayésienne. *Seconde partie : classification. traitement du signal*, 1997.

Manuscrit reçu le 18 juillet 1996

LES AUTEURS

Xavier DESCOMBES



Xavier Descombes est chercheur invité à l'institut Max Planck pour les Sciences Cognitives à Leipzig (Allemagne) depuis octobre 1996. Auparavant il a travaillé en tant que chercheur post-doctorant successivement à l'ENST (1994), à l'Université Catholique de Louvain (1995) et à l'INRIA au sein du projet PASTIS (1996). Il a obtenu un diplôme d'ingénieur de l'ENST en 89, un DEA de mathématiques pures à l'université Pierre et Marie Curie (Paris VI) en 90 et une thèse de doctorat à l'ENST en 93. Ses domaines d'intérêt sont la modélisation stochastique, l'estimation des paramètres et l'optimisation. Ses travaux, concernant principalement les champs de Markov, ont été appliqués à l'imagerie satellitaire et médicale pour des tâches de segmentation, de restauration et d'extraction d'objets.

Robin MORRIS



Robin Morris a obtenu le "BA degree" au département ingénierie de l'université de Cambridge en 1991. Il obtient son doctorat dans le groupe de traitement du signal du même département en 1995. En Octobre 1994, il obtient une bourse "Junior Research Fellowship" du Trinity College de Cambridge. En 1996, il est à l'INRIA. Depuis Janvier 1997, il travaille dans le groupe d'analyse de données du "NASA Ames research center" grâce à une bourse NRC. Ses domaines d'intérêt sont l'inférence Bayésienne, la modélisation d'images et les méthodes MCMC.

Josiane ZÉRUBIA



Josiane Zerubia est chercheur permanent à l'INRIA depuis 1989. Elle est directeur de recherche depuis juillet 1995 et responsable du laboratoire PASTIS (INRIA Sophia-Antipolis) depuis novembre 1995. Auparavant, elle était à l'institut de traitement des signaux et des images de l'université de Californie du sud à Los Angeles en tant que Post doctorante. Elle a également travaillé comme chercheur au LASSY (Université de Nice et CNRS) entre 1984 et 1988, au laboratoire de recherche de Hewlett Packard en France et à Palo-Alto entre 82 et 84. Elle a obtenu un diplôme d'ingénieur de l'ENSIEG, Grenoble, en 81, un doctorat ingénieur en 86, un doctorat en 88 et une habilitation à diriger des recherches en 94. Elle est membre de IEEE depuis 1988 et de l'académie des sciences de New-York depuis 1996. Son domaine d'intérêt est le traitement des images, en particulier la restauration d'images, la segmentation, classification, le groupement perceptuel, la super-résolution en utilisant des modèles probabilistes ou des réseaux neuronaux. Elle travaille également sur l'estimation de paramètres et les techniques d'optimisation.