

## Modèles de Potts et relaxation d'images de labels par champs de Markov

### *Potts Models and Image Labelling Relaxation by Random Markov Fields*



**Marc SIGELLE**

E.N.S.T.

46, rue Barrault,  
75634 Paris Cedex 13

Marc Sigelle est né à Paris le 18 mars 1954. Diplômé de l'École Polytechnique en 1975 et de l'École Nationale Supérieure des Télécommunications en 1977, il a travaillé au Centre National d'Études des Télécommunications où il a mené des travaux de recherche en Physique et en Informatique. Il est à Télécom Paris depuis 1989, où il mène des travaux sur les Champs de Markov en Traitement d'Images.



**Rémi RONFARD**

DASSAULT

Systèmes Modeleur géométrique  
38 avenue Charles de Gaulle,  
92150 Suresnes

Rémi Ronfard est né à Cambrai le 25 juin 1963. Diplômé de l'École des Mines de Paris en 1986, il a travaillé en tant qu'attaché au Centre de Télédétection de l'École des Mines à Sophia Antipolis. Il a soutenu une thèse de Doctorat en février 1991 sur l'extraction de contours dans les images multi-spectrales et en couleur. Il travaille actuellement au département modeleur Géométrique de Dassault Systèmes.

### RÉSUMÉ

Nous montrons dans cet article la relation profonde entre certains modèles d'énergie provenant de la Physique Statistique utilisés et les modèles utilisés en champ de Markov pour l'étiquetage d'images. Nous présentons comme application une méthode markovienne de relaxation et d'amélioration d'images préclassifiées. On définit pour cela une fonction énergie ne dépendant que des labels et de leur valeur initiale, la connaissance *a priori* sur l'image provenant de la matrice de confusion déduite des échantillons de référence utilisés pour la classification initiale. La fonction à minimiser inclut divers termes assurant la régularité spatiale des labels, la croissance ou la disparition de certaines classes. Cette méthode permet en particulier

de reclasser contextuellement les pixels d'une classe de rejet. Enfin, nous présentons des résultats obtenus sur des images multispectrales en télédétection et en géologie, où nous comparons les résultats des modes conditionnels itérés et du recuit simulé. La méthode n'opérant que sur un processus label s'avère être très performante.

### MOTS CLÉS

Images, classification, relaxation, champs de Markov, physique statistique.

### ABSTRACT

We show in this paper the deep relationship between classic models from Statistical Physics and Markovian Random Fields models used in image labelling. We present as an application a markovian relaxation method for enhancement and relaxation of previously classified images. An energy function is defined, which depends only on the labels and on their initial value. The main *a priori* pixel knowledge results from the confusion matrix of the reference samples used for initial classification. The energy to be minimized includes also terms ensuring simultaneous spatial label regularity, growth of some classes and disparition of some others. The method allows for example to reclassify previous rejection class pixels in

their spatial environment. Last we present some results on Remote Sensing multispectral and geological ore images, comparing the performances of Iterated Conditional Modes (ICM) and Simulated Annealing (SA). Very low CPU time was obtained due to the principle of the method, working on labels instead of gray levels.

### KEY WORDS

Images, classification, relaxation Markov Random fields, statistical physics.

## 1. Introduction

Nous présentons dans cet article une méthode à base de Champs de Markov pour le traitement général d'une classe de problèmes de relaxation de labels dans les images [1], [2], [3], [4]. Nous décrivons dans ce but un cadre théorique issu de la Physique Statistique permettant de généraliser le modèle d'Ising (restauration d'images binaires) à la restauration d'images multi-classes connu sous le nom de modèle de Potts en Physique [5]. L'approche usuelle utilisée en segmentation markovienne fait intervenir la référence aux données originales [6], [7], que ce soit une image mono- ou multispectrale. La méthode exposée ici s'applique dans le cas où l'on dispose de peu de connaissances *a priori* spécifiques, et où les seuls éléments d'information proviennent d'une classification initiale des données, qui sera prise comme référence. Dans cette approche, l'algorithme d'étiquetage initial reste inchangé, et la méthode de relaxation par Champs de Markov intervient comme seconde étape permettant de « retoucher » la classification initiale en fonction des voisins requis. Le déroulement de chaque étape du processus global peut ainsi être contrôlé et modifié si nécessaire. De plus, l'algorithme de relaxation n'opérant que sur un processus label entraîne un faible temps calcul. Nous espérons que le cadre ainsi décrit fournisse une approche cohérente à ce genre de problèmes.

Le paragraphe 2 décrit le problème de classification contextuelle étudié et sa décomposition en deux étapes. Les paragraphes 3 et 4 présentent le cadre théorique employé ainsi que ses relations avec la théorie des réseaux de spins en Physique Statistique. Dans le paragraphe 5 nous discutons les paramètres du modèle et donnons un aperçu de leur estimation à partir de données issues de la classification initiale et d'autres connaissances spécifiques *a priori* lorsqu'elles sont disponibles. Nous présentons enfin dans le paragraphe 6 un ensemble de résultats obtenus sur des images de télédétection en classification multispectrale et sur des images géologiques de fond de mines, puis décrivons les performances de la méthode.

## 2. Relaxation stochastique de données classifiées de façon optimale

Notons  $X$  le processus *pixel* (par exemple une image Landsat multispectrale) et  $L$  le processus *label* (par exemple une classification thématique du terrain). Nous proposons ici le schéma suivant de classification en deux étapes :

$$X \rightarrow L_0 \rightarrow L$$

première classification (A)    relaxation markovienne (B)

$$\text{Max } P(X/L_0) P(L_0) \quad \text{Max } P(L_0/L) P(L).$$

Lorsque la première classification (A) est bayésienne (ce qui sera le cas envisagé dans la suite), l'étiquetage initial  $L_0$  est obtenu en maximisant  $P(X/L_0) P(L_0)$ , les probabilités conditionnelles  $P(X/L_0)$  étant connues à partir des

histogrammes multispectraux d'échantillons de référence convenablement choisis. Le processus de relaxation stochastique (B) permet de reclasser les pixels situés aux frontières entre régions ainsi que les pixels non classés dans l'étape (A). Les pixels frontières sont reclassés de façon contextuelle grâce au Champ de Markov  $P(L)$ . En ce qui concerne le traitement des pixels de la classe de rejet, deux possibilités se présentent :

— l'échantillonnage initial n'est pas complet en nombre de classes : ces pixels constituent donc une nouvelle classe, dont la régularité spatiale sera assurée au même titre que les autres régions par le processus (B) ;

— l'échantillonnage initial est complet : ces pixels doivent alors être reclassés dans les autres régions, et ceci de façon contextuelle.

Il est clair que dans ce schéma, aucune référence aux probabilités initiales de classification n'est effectuée pendant la relaxation, de sorte que l'attache aux données doit être précisée de façon fiable à partir de l'étiquetage initial. Nous suggérons d'utiliser pour cela les résultats de la première classification sur les échantillons de référence eux-mêmes, par une utilisation non conventionnelle de la matrice de confusion qui en résulte (voir § 5).

## 3. Modèle d'Ising et régularisation d'images binaires

Le modèle d'Ising a été développé en 1925 dans le cadre de la théorie du ferromagnétisme pour rendre compte des propriétés de transition de phase ordre-désordre dans les solides. Dans le modèle du réseau plan isotrope et 4-connexe, un spin en un site  $s$  du réseau interagit avec ses 4-voisins (notés  $r$ ) par une constante de couplage  $J$  et avec un champ magnétique  $B$  de sorte que l'énergie du site  $s$  conditionnellement à ses voisins s'écrit

$$(1) \quad W_s = -J \sum_r \sigma_r \sigma_s - B \sigma_s \quad (\sigma_s = \pm 1).$$

L'énergie totale  $E$  du réseau est la somme des termes d'interaction site-site (cliques d'ordre 2) et des termes d'interaction site-champ magnétique (cliques d'ordre 1). Une forme équivalente de la fonction énergie est la forme de Heisenberg

$$(2) \quad E = \frac{J}{2} \sum_{(r,s)} \|\sigma_r - \sigma_s\|^2 + \frac{B}{2} \sum_r \|\sigma_r - b\|^2$$

ou encore

$$(3) \quad E = \frac{J}{2} \sum_{(r,s)} \|\hat{u}_r - \hat{u}_s\|^2 + \frac{B}{2} \sum_r \|\hat{u}_r - \hat{b}\|^2$$

la première somme portant sur les cliques d'ordre 2, la deuxième sur les cliques d'ordre 1, où les  $\hat{u}_r$  sont les vecteurs unitaires associés aux directions des spins en chaque site et  $\hat{b}$  le vecteur unitaire associé au sens du champ magnétique.

Rappelons maintenant la démonstration de Carnevali, Coletti et Patarnello [8] établissant le lien entre régularisation d'une image binaire et modèle d'Ising. Elle est la source des extensions que nous avons effectuées pour les images multi-classes. Considérons une image à régulariser  $f^0(x, y)$  comme continue dérivable dans un domaine de  $\mathbf{R}^2$ , on cherche  $f(x, y)$  solution du problème

$$\min E = E_1 + E_2$$

où

$$(4) \quad E_1 = \lambda \|f - f^0\|^2 = \lambda \iint_{\text{Image}} (f(x, y) - f^0(x, y))^2 dx dy$$

est le terme d'écart à l'image originale

$$(5) \quad E_2 = \mu \|\nabla f\|^2 = \mu \iint_{\text{Image}} \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 dx dy$$

est le terme de régularisation d'image adopté (norme  $L^2$  du gradient de l'intensité).  $E$  est l'énergie totale de l'image. En discrétisant sur l'image, il vient

$$(6) \quad E_1 = \lambda \sum_{i,j} \|f_{i,j} - f_{i,j}^0\|^2$$

$$(7) \quad E_2 = \mu \left( \sum_{i,j} \|f_{i,j} - f_{i+1,j}\|^2 + \|f_{i,j} - f_{i,j+1}\|^2 \right).$$

(Les effets de bord peuvent être pris en compte en raccordant l'image de façon torique ou en l'encadrant d'un bord fixe.)

Lorsque l'image est binaire :  $f_{i,j} = 0$  ou  $1 \forall i, j$  on associe à chaque pixel sa « variable de spin »  $\sigma_{i,j}$

$$(8) \quad f_{i,j} = \frac{1 + \sigma_{i,j}}{2} \quad (\sigma_{i,j} = \pm 1).$$

Les équations (4-5) deviennent

$$(9) \quad E = E_1 + E_2$$

$$(10) \quad E_1 = \frac{\lambda}{4} \sum_{i,j} \|\sigma_{i,j} - \sigma_{i,j}^0\|^2$$

$$(11) \quad E_2 = \frac{\mu}{4} \left( \sum_{i,j} \|\sigma_{i,j} - \sigma_{i+1,j}\|^2 + \|\sigma_{i,j} - \sigma_{i,j+1}\|^2 \right).$$

Comparant à (3), on trouve que l'énergie totale  $E$  est celle d'un modèle d'Ising 2D isotrope de constante de couplage entre sites  $J = \frac{\mu}{2}$  en présence d'un champ magnétique de module constant  $B = \frac{\lambda}{2}$  mais inhomogène et dont le sens

en chaque pixel dépend de sa valeur originale binaire (c'est-à-dire un champ magnétique presque « homogène par morceaux » si l'image de départ n'est pas trop bruitée).

*N.B.* : Une série de tests menés sur des images binaires donne un choix des paramètres  $J = 2$ ,  $B = 3$ , température

initiale  $T_0 = 2$ , en accord avec les valeurs trouvées par Carnevali *et al.* [8]. Nous verrons dans la suite comment généraliser ce choix aux images multi-classes.

## 4. Restauration d'images multi-classes et modèle de Potts

Le potentiel de cliques d'ordre 2 le plus couramment utilisé dans les méthodes de régularisation d'images de classes est [4], [6], [7], [9], [11], [12]

$$(12) \quad V(r, s) = \begin{cases} -\beta & \text{si } \ell_r = \ell_s \\ +\beta & \text{si } \ell_r \neq \ell_s \end{cases}$$

(on supposera la 4(connexité dans la suite de l'exposé).

Il s'agit là à une constante près du modèle dit (standard) de Potts en Physique Statistique

$$(13) \quad V(r, s) = -K\delta(\ell_r, \ell_s) \quad (K = 2\beta)$$

$\delta$  est le symbole de Kronecker défini par

$$\delta(x_1, x_2) = 1 \quad \text{si } x_1 = x_2, \quad 0 \quad \text{sinon}$$

Ce modèle a été abondamment étudié pour ses propriétés de Transitions de Phase [5], qui sont maintenant bien connues et qui dépendent fortement du nombre de classes  $q$ . Le modèle de Potts coïncide avec le modèle d'Ising lorsque  $q = 2$ , en vertu de la formule

$$\delta(\ell_r, \ell_s) = \frac{1 + \sigma_r \sigma_s}{2}$$

(d'où il résulte  $K_{\text{Potts}} = 2J_{\text{Ising}}$ ), les  $\sigma_i$  étant les variables de spin  $\pm 1$  associées aux valeurs (binaires) des pixels. Il en est donc une extension naturelle lorsque  $q$  est quelconque. Nous allons maintenant montrer comment étendre les résultats du paragraphe à partir de ce formalisme. Partons de la formule suivante [5]

$$(14) \quad \delta(m, n) = \frac{1}{q} (1 + (q-1) \hat{u}_m \cdot \hat{u}_n)$$

— où  $m$  et  $n$  sont deux entiers appartenant à l'intervalle  $[0 \dots q-1]$ ,

— où les  $\hat{u}_i$  sont  $q$  vecteurs unitaires « régulièrement répartis » au sommet d'un hypertétraèdre de  $\mathbf{R}^{q-1}$  (voir fig. 1), c'est-à-dire tels que

$$(15) \quad \sum_{k=0}^{q-1} \hat{u}_k = \vec{0} \quad \text{et} \quad \hat{u}_m \cdot \hat{u}_n = -\frac{1}{q-1} \quad \text{si } m \neq n.$$

Le cas  $q = 2$  est bien celui du modèle d'Ising : on obtient deux valeurs de spin de direction opposée. Les vecteurs  $\hat{u}_i$  étant unitaires, la fonctionnelle énergie obtenue en (9-10-11) s'étend de façon naturelle sous la forme suivante

$$(16) \quad E = E_1 + E_2$$

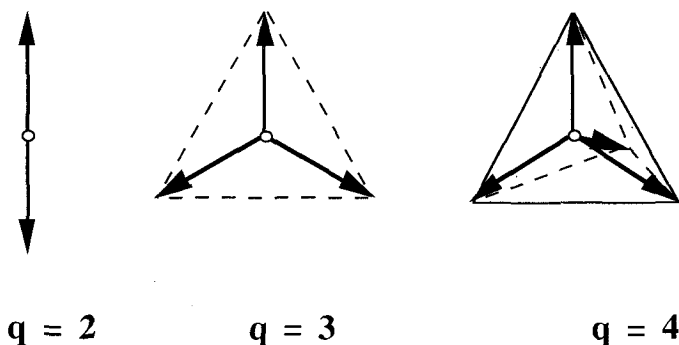


Fig. 1. — Le modèle de Potts pour plusieurs valeurs du nombre de classes (d'après [5]).

$$(17) \quad E_1 = \lambda \sum_{i,j} \|\hat{u}_{i,j} - \hat{u}_{i,j}^0\|^2$$

$$(18) \quad E_2 = \mu \left( \sum_{i,j} \|\hat{u}_{i,j} - \hat{u}_{i+1,j}\|^2 + \|\hat{u}_{i,j} - \hat{u}_{i,j+1}\|^2 \right)$$

Le contenu scalaire de l'image  $f(x, y)$  a été remplacé par une expression vectorielle dans  $\mathbf{R}^{q-1}$ , et la forme d'énergie obtenue est une généralisation à  $q-1$  dimensions du modèle de Heisenberg. L'énergie conditionnelle en un site  $s$  prend la forme

$$(19) \quad W_s = -K \sum_r \delta(\ell_r, \ell_s) - B\delta(\ell_s, \ell_s^0)$$

où

$$(20) \quad \lambda = \frac{q-1}{q} \frac{B}{2} \quad \text{et} \quad \mu = \frac{q-1}{q} \frac{K}{2}$$

les notations étant celles des formules (4) et (5).

La forme (19), qui est probablement le modèle le plus simple associé à la restauration d'images de classes, est indépendante de toute « relation d'ordre » entre labels distincts ou de toutes connaissances *a priori* sur les régions (proximité, etc...). Nous développons au paragraphe suivant une extension de cette formulation afin de permettre leur prise en compte.

NB. : De la même façon que  $K_{\text{Potts}} = 2 J_{\text{Ising}}$ , on a  $B_{\text{Potts}} = 2 B_{\text{Ising}}$ .

Cela nous a incité à adopter dans nos applications  $K = 4$ ,  $B = 6$ , température initiale  $T_0 = 4$ .

*Interprétation physique du potentiel de Derin*

Besag, Derin, Kelly et Karssemeijer [4], [9], [10], [11], [12] incluent dans leur modèle d'énergie le potentiel suivant (d'ordre 1)

$$(21) \quad V_c(n) = \alpha_n \quad \text{pour} \quad n \in [0 \dots q-1]$$

qui permet de favoriser ou défavoriser l'apparition de certaines classes, ce que nous allons retrouver à partir du formalisme précédent. En effet, on peut écrire

$$(22) \quad V_c(n) = \sum_{k=0}^{q-1} \alpha_k \delta(k, n)$$

c'est-à-dire, en employant les expressions vectorielles développées plus haut

$$(23) \quad V_c(n) = \frac{q-1}{q} \sum_{k=0}^{q-1} \alpha_k \hat{u}_k \cdot \hat{u}_n + W = -\vec{C} \cdot \hat{u}_n + W$$

où  $W$  est une constante scalaire et  $\vec{C}$  un vecteur de  $\mathbf{R}^{q-1}$ , entièrement analogue à un champ magnétique constant

$$(24) \quad \vec{C} = -\frac{q-1}{q} \sum_{k=0}^{q-1} \alpha_k \hat{u}_k \nu = \hat{c}$$

où  $\nu = \|\vec{C}\|$ . L'énergie totale correspond alors au modèle de Potts en présence d'un champ magnétique constant. Les pixels (« spins ») vont « s'orienter » selon la direction de ce champ, c'est-à-dire selon les labels  $k$  associés aux composantes  $\alpha_k$  les plus faibles. Cette modélisation est cohérente : ainsi lorsque les  $\alpha_k$  sont égaux,  $\vec{C}$  est le vecteur nul  $\left( \sum_{k=0}^{q-1} \hat{u}_k = \vec{0} \right)$ . Effectivement aucune classe n'a été

particularisée. Notons que le terme supplémentaire associé au potentiel de Derin apparaissant dans la fonctionnelle énergie est

$$(25) \quad E_3 = \frac{\nu}{2} \sum_{i,j} \|\hat{u}_{i,j} - \hat{c}\|^2$$

Dans un de nos applications, nous avons pris en compte un terme de champ  $\vec{C}$  dirigé selon un vecteur  $\hat{u}_L$  particulier ( $L \in [0 \dots q-1]$ ), afin d'augmenter ou de diminuer la probabilité de présence de la classe  $L$  dans l'image à restaurer. La forme de l'énergie conditionnelle au pixel  $s$  devient

$$(26) \quad U = -B \sum_s \delta(\ell_s, \ell_s^0) + \alpha_L \sum_s \delta(\ell_s, L) - K \sum_{(r,s)} \delta(\ell_r, \ell_s)$$

Tous ces développements montrent donc le lien profond avec la « Physique des Transitions de Phase » pour les modèles envisagés dans ce travail.

## 5. Extension à la restauration d'images de classes

On peut toujours mettre la probabilité *a posteriori* à maximiser  $P(L_0/L) P(L)$  sous forme d'une distribution de Gibbs « composée » :

$$(27) \quad P(L_0/L) P(L) = \frac{1}{Z} e^{-U(L_0/L) - U(L)}$$

En effet le terme contextuel  $P(L)$  peut déjà être supposé champ de Markov, c'est-à-dire vérifiant la propriété de Gibbs. Quant à la conditionnelle  $P(L_0/L)$ , on peut également la mettre sous forme de distribution de Gibbs sous l'hypothèse de stationnarité et d'indépendance des pixels

conditionnellement à leurs valeurs initiales. Il est alors facile de généraliser le modèle d'énergie développé au paragraphe précédent en définissant les « éléments de matrice » suivants :

— le terme d'énergie d'ordre 1 est

$$(28) \quad U(\ell_s, \ell_s^0) = - \sum_{ij} B_{ij} \delta(i, \ell_s^0) \delta(j, \ell_s)$$

— le terme d'énergie d'ordre 2 est

$$(29) \quad U(\ell_r, \ell_s) = - \sum_{ij} K_{ij} \delta(i, \ell_r) \delta(j, \ell_s)$$

— [K] est une matrice générale décrivant l'interaction entre sites, dont les paramètres sont en général difficiles à évaluer et au prix de méthodes complexes (ne serait-ce que pour le modèle de Potts lui-même). Nous choisissons par la suite une représentation diagonale « scalaire »

$$[K] = - K[1]$$

où [1] désigne la matrice identité. La seule connaissance *a priori* supplémentaire sur la configuration spatiale des régions que nous avons incluse dans [K] concerne le cas de deux régions *i* et *j* non adjacentes. Le coefficient  $K_{ij}$  associé prend alors une valeur négative de grande amplitude ;

— [B] est une matrice dont chacun des éléments  $B_{ij}$  contrôle la probabilité de transition de l'étiquetage initial  $L_0 = i$  vers l'étiquetage relaxé  $L = j$ . Ses coefficients peuvent être déduits de la classification initiale. En effet, soient  $S_j$  des échantillons de référence parfaitement classifiés, de taille  $n_j$ . On peut en déduire la probabilité conditionnelle d'un étiquetage initial *i* étant donné l'étiquetage « vrai » *j* :

$$(30) \quad P_{ij} = P(L_0 = i / L = j) = \frac{n_{ij}}{n_j}$$

où  $n_{ij}$  est l'occurrence de la classe *i* dans l'échantillon *j*. La matrice [P] n'est autre que la transposée de la matrice de confusion calculée sur les échantillons de référence eux-mêmes. En effet, la matrice de confusion associée à ces échantillons donne, ligne par ligne, la probabilité de classification initiale obtenue pour chacun des échantillons supposés parfaits (voir une application au § 6.4). De plus, le processus de relaxation (B) maximisant la probabilité *a posteriori*  $P(L_0/L)P(L)$ , et en supposant l'hypothèse d'ergodicité ainsi que l'indépendance des pixels vis-à-vis de  $P(L_0/L)$ , on peut déduire les coefficients  $B_{ij}$  :

$$(31) \quad P(\ell_s^0 = i / \ell_s = j) = \frac{\exp(B_{ij})}{Z_j} \approx P_{ij}$$

où les coefficients  $Z_j$  assurent la normalisation des probabilités conditionnelles  $P_{ij}$

$$(32) \quad \sum_i P_{ij} = 1.$$

Les  $B_{ij}$  sont définis à une constante additive près. Dans ce

contexte nous pouvons maintenant formuler la contribution des divers termes d'énergie  $-\alpha_L \delta(1, L)$  comme suit :

$$(33) \quad [B] = - \sum_{L=0}^{q-1} \alpha_L \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdot & 0 \\ 0 & 1 & 0 & \cdot & 0 \\ 0 & 1 & 0 & \cdot & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

En effet, pour chaque valeur de L, on a  $U(\ell_s, \ell_s^0) = -\alpha_L$  lorsque  $\ell_s = L$  (0 sinon) indépendamment du label initial  $\ell_s^0$ .

## 6. Implémentation et résultats

### 1. Conditions générales d'expérimentation

— L'image est raccordée de façon « torique » de façon à obtenir la même cardinalité du voisinage pour tout point de l'image. De façon à accélérer le déroulement de l'algorithme, nous avons mis en place les procédures suivantes.

— Les matrices [K] et [B] sont à coefficients entiers, aux coefficients de pondération respectifs K et B près, de sorte que la probabilité de transition en tout pixel peut se mettre sous la forme :

$$(34) \quad P = \min \left[ \exp \left( \frac{KM + BN}{T} \right), 1.0 \right]$$

où M et N sont des entiers finis. On peut donc connaître à température fixée l'ensemble des probabilités de transition possibles, sans avoir à les recalculer en chaque pixel.

— De la même façon, nous avons implanté une liste circulaire (2 048 éléments par exemple) prédéfinie d'« objets aléatoires », contenant à la fois des labels et des nombres générés de façon aléatoire. A l'examen de chaque pixel, l'élément courant de la liste indique une transition possible, ainsi qu'un nombre aléatoire qui est comparé à la probabilité P définie par (34). Lorsque le pixel suivant est examiné on passe à l'élément de liste suivant, et ainsi de suite pendant l'ensemble des itérations.

Les paramètres de l'algorithme de recuit simulé que nous avons employés sont les suivants :

— La température initiale est  $T_0 = 4$ . Elle décroît ensuite d'un facteur  $\rho = 5\%$  à chaque itération, c'est-à-dire à chaque balayage d'image.

— La convergence est supposée atteinte quand le nombre de pixels ayant changé de valeur entre deux itérations est inférieur à  $\frac{1}{160}$  (taille de l'image).

— En général, de 30 à 50 itérations suffisent pour assurer la convergence pour des images comprenant un faible nombre de classes (de 4 à 7). Une image de taille  $256 \times 256$  nécessite donc un temps CPU moyen de 1,6 s par itération sur VAX 8550 et de 6 s sur PC-AT 286 (équipé d'une carte de visualisation en temps réel), d'où un temps CPU total compris entre 45 s et 1 mn 20 s sur VAX et entre 3 et 5 mn sur PC respectivement (i.e. environ 4 fois plus de temps pour des images  $512 \times 512$ ).

Nous avons comparé les résultats issus de la méthode du recuit simulé et des Modes Conditionnels Itérés (ICM). Le recuit simulé s'est révélé plus efficace et mieux adapté à notre problème, car la classification initiale peut être proche d'un minimum local de la fonction énergie, surtout lorsque la région à reclasser est importante (voir fig. 2 : images LANDSAT). En effet, dans ce cas la contribution des termes d'attache aux données et d'interaction pixel-pixel est prépondérante de sorte que cette région ne peut être facilement reclassifiée par l'ICM.

## 2. Reclassification d'images LANDSAT

Dans notre premier exemple, une image multispectrale LANDSAT de la région de Toulon avait été classifiée précédemment en sept régions : sols nus — culture — sites urbains — feuillus — conifères — eau — classe de rejet (fig. 2.1). La classe de rejet correspond au label  $L = 0$  (en noir sur la fig. 2). On note qu'elle forme presque une région en elle-même. la matrice d'interaction de cliques d'ordre 2 est  $[K] = -4.0 [1]$  dans tous les cas. Deux matrices  $[B]$  ont été utilisées dans nos expériences (coefficient de pondération dans la fonction énergie  $B = 1.0$ )

$$\begin{bmatrix} 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 6 & 0 & \cdot & \cdot & 0 \\ 0 & \cdot & 0 & 6 & 0 & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & 6 & 0 & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

$[B]_1$

sans reclassification

$$\begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ -8 & 6 & 0 & \cdot & \cdot & \cdot & 0 \\ -8 & 0 & 6 & 0 & \cdot & \cdot & 0 \\ -8 & 0 & 0 & 6 & 0 & \cdot & 0 \\ -8 & 0 & \cdot & 0 & 6 & 0 & 0 \\ -8 & 0 & \cdot & \cdot & 0 & 6 & 0 \\ -8 & 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$$

$[B]_2$

reclassification de la classe 0

Les matrices  $[B]$  utilisées

dans l'étude d'images LANDSAT .

Les résultats correspondants sont indiqués en figure 2.

— la figure 2.2 a été obtenue pour la fonctionnelle d'énergie décrite en (26) avec la matrice  $[B]_1$  par recuit simulé. Les régions correspondant aux diverses classes ont été « homogénéisées » de façon « uniforme », indépendamment de leur classe. L'image obtenue peut déjà être utilisée de façon exploitable ;

— la figure 2.3 a été obtenue par ICM pour la même fonctionnelle d'énergie avec la matrice  $[B]_2$  (correspondant au label de rejet  $L = 0$ ). La convergence a été obtenue en

10 itérations. Comme précisé auparavant, on n'obtient pas une reclassification totale des pixels de rejet. C'est ici l'un des cas où le recuit simulé s'avère indispensable pour obtenir l'optimum désiré.

— la figure 2.4 a été obtenue avec la matrice  $[B]_2$  par recuit simulé. L'ensemble des pixels noirs a été reclassifié de façon contextuelle dans les autres classes, ce qui est en accord avec la connaissance *a priori* que l'on avait de l'image (6 classes seulement). On remarque que les autres classes ont été peu modifiées par rapport à la figure 2.2, ce qui indique la stabilité de ce type de méthode. La validité du résultat reste à être étayée par comparaison avec une carte thématique réelle.

## 3. Reclassification d'images de roches

La série de figures 3 et 4 présente des résultats d'expérimentations effectuées sur des images de roches prises en imagerie vidéo couleur.

L'image étudiée comprend 6 classes

- 0  $\Rightarrow$  Fond (En noir sur l'image)
- 1  $\Rightarrow$  Quartz (En sombre sur l'image)
- 2  $\Rightarrow$  Schistes (En blanc sur l'image)
- 3  $\Rightarrow$  Blende (En gris foncé sur l'image)
- 4  $\Rightarrow$  Oxydes (En gris clair sur l'image)
- 5  $\Rightarrow$  Classe de rejet (En noir sur l'image).

Un des buts de cette étude était d'estimer la teneur réelle en blende dans l'image initiale (classe 3). Il s'agissait comme dans l'exemple précédent de reclasser les pixels initialement rejetés, dont le taux considérable dans cette image (37,18 %) est dû à deux raisons :

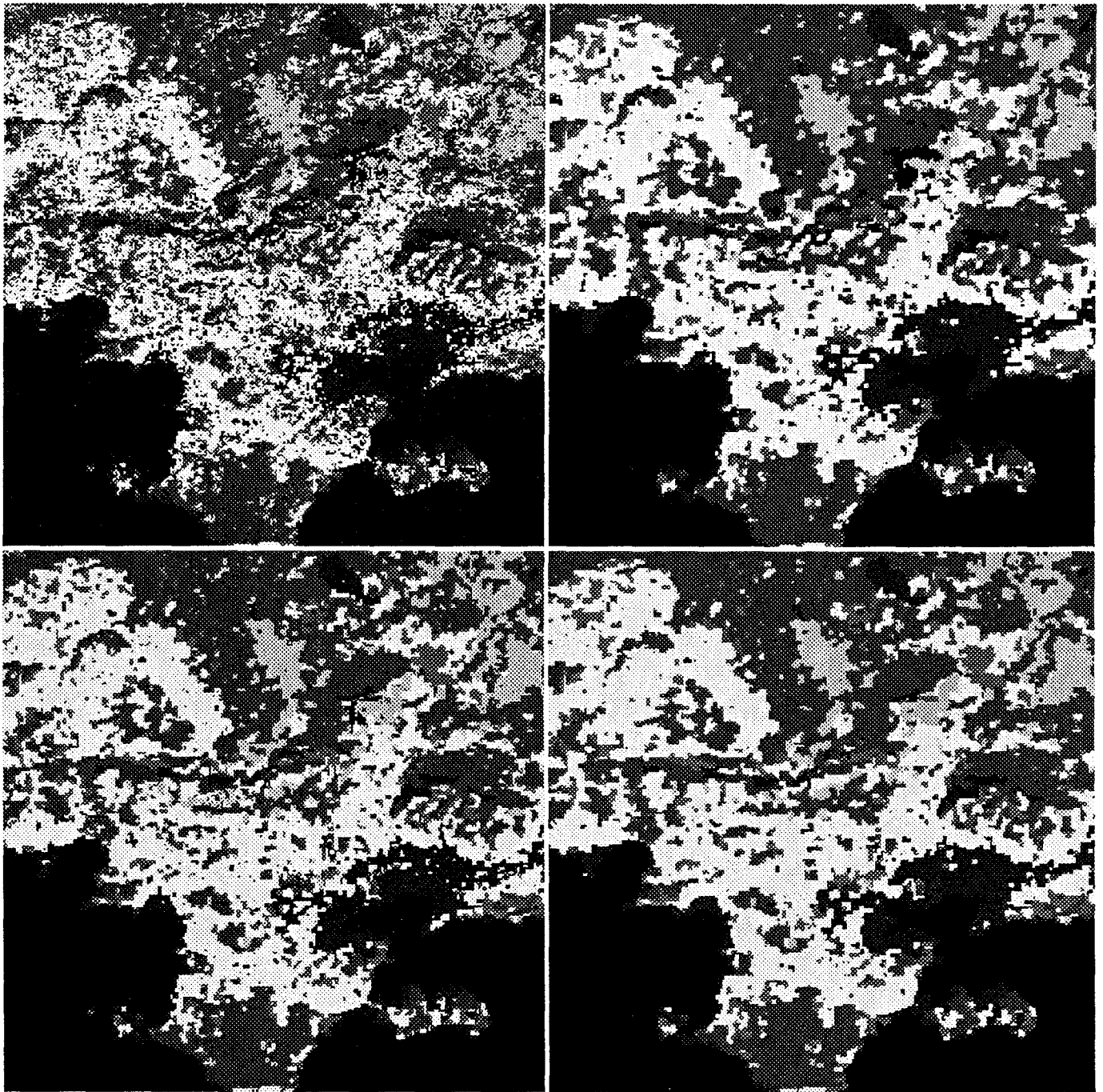
— Les distributions spectrales des minerais présents sur les trois canaux vidéo couleur se recouvrent considérablement, d'où un seuil de rejet élevé dans la classification initiale.

— L'image de teinte initiale est très texturée. Les points de fort gradient ont alors été inclus dans la classe de rejet pour permettre à l'algorithme d'homogénéiser la distribution spatiale des classes. Dans cette étude, elle est la suivante :

échantillon	classe				
	1	2	3	4	5
1	80,05	2,08	0	0,35	17,51
2	5,14	93,76	0	0	1,09
3	21,35	10,82	64,24	2,64	0,94
4	25,20	10,87	3,25	52,24	7,53
5	0	0	0	0	0

La matrice de confusion sur les données de la classification initiale (les probabilités sont exprimées en %).

On note une « corrélation » considérable entre les diverses classes, que l'on peut plus ou moins prendre en compte dans l'algorithme de reclassification. Le lien avec la matrice  $[B]$  s'effectue, comme exposé précédemment, grâce à (31). Nous avons testé notre méthode avec les matrices  $[B]$  suivantes (le coefficient de pondération dans



1	2
3	4

Fig. 2.

1) Classification bayésienne multispectrale d'une image LANDSAT de la région de Toulon.  
2) Reclassification de cette image par recuit simulé avec la matrice  $[B]_1$  (voir texte).

3) Reclassification de cette image par Conditionnels Modes Itérés avec la matrice  $[B]_2$ .  
4) Reclassification de cette image par recuit simulé avec la matrice  $[B]_2$  (voir texte). Les pixels précédemment non classifiés (pixels noirs) ont disparu.



Fig. 3. — Etude d'images géologiques. Classification initiale d'une image de roche (voir texte).

par la « sous-matrice » de confusion associée aux vraies classes. Le traitement avec  $[B]_5$  effectue au contraire une approximation « diagonale » de la matrice de confusion et assure ainsi une reclassification selon les proportions des vraies classes dans la classification initiale. Les résultats sont indiqués respectivement sur les figures 3.2 et 3.3.

Les taux de présence des minerais dans les diverses images sont les suivants :

	Initiale	$[B]_3$	$[B]_4$	$[B]_5$	$[B]_6$
quartz	15.64	37.74	21.71	26.05	39.70
schistes	33.64	41.85	53.25	54.05	41.30
blende	5.38	5.93	12.27	9.12	5.90
oxydes	8.16	12.72	12.77	10.78	13.10
rejet	37.18	1.76	0.0	0.0	0.0

La variation de ces taux de présence d'un cas à l'autre est cohérente avec le choix correspondant de la matrice  $[B]$ . Cette étude reste toutefois une approche d'école demandant à être étayée auprès de spécialistes du domaine.

la fonction énergie est  $B = 0,6$

$$\begin{array}{cc}
 \begin{bmatrix} 10 & 4 & 7 & 7 & 0 \\ 2 & 10 & 5 & 5 & 0 \\ 0 & 0 & 9 & 3 & 0 \\ 0 & 0 & 2 & 9 & 0 \\ 6 & 0 & 0 & 4 & 0 \end{bmatrix} & \begin{bmatrix} 10 & 4 & 7 & 7 & 0 \\ 2 & 10 & 5 & 5 & 0 \\ 0 & 0 & 9 & 3 & 0 \\ 0 & 0 & 2 & 9 & 0 \\ 7 & 7 & 7 & 7 & 0 \end{bmatrix} \\
 [B]_3 & [B]_4 \\
 \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 9 & 0 \\ 7 & 7 & 7 & 7 & 0 \end{bmatrix} & \begin{bmatrix} 10 & 4 & 7 & 7 & -5 \\ 2 & 0 & 5 & 5 & -5 \\ 0 & 0 & 9 & 3 & -5 \\ 0 & 0 & 2 & 9 & -5 \\ 6 & 0 & 0 & 4 & -5 \end{bmatrix} \\
 [B]_5 & [B]_6
 \end{array}$$

Les différentes matrices  $[B]$  utilisées dans l'étude d'images de roches (La classe « fond » n'est pas incluse).

— Les matrices  $[B]_3$  et  $[B]_6$  prennent effectivement en compte la confusion entre classes. Les pixels de rejet seront reclassés par préférence vers les classes 1 (Schistes) et 4 (Oxydes) (voir dernière ligne de la matrice). La matrice  $[B]_6$  ne diffère de  $[B]_3$  que par un terme de type  $\alpha_L$ ,  $L = 5$ , assurant la reclassification totale de la classe de rejet. Les résultats sont indiqués respectivement sur les figures 3.1 et 3.4.

— Les matrices  $[B]_4$  et  $[B]_5$  assurent une répartition « uniforme » de la classe de rejet vers les autres classes. Le traitement avec  $[B]_4$  permet une reclassification gouvernée

## 7. Conclusion

En conclusion, nous avons mis en place une méthode efficace et fiable de relaxation d'images de classes qui pourrait en particulier être en classification multispectrale. Le modèle énergétique décrit dans cet article pourrait être amélioré de façon profitable en y incluant la connaissance des probabilités ponctuelles mises en jeu dans la première classification. Il en résulterait un processus de Markov non stationnaire [12]. Parmi les problèmes théoriques liés à l'algorithme du recuit simulé qui restent à aborder, on peut citer les suivants :

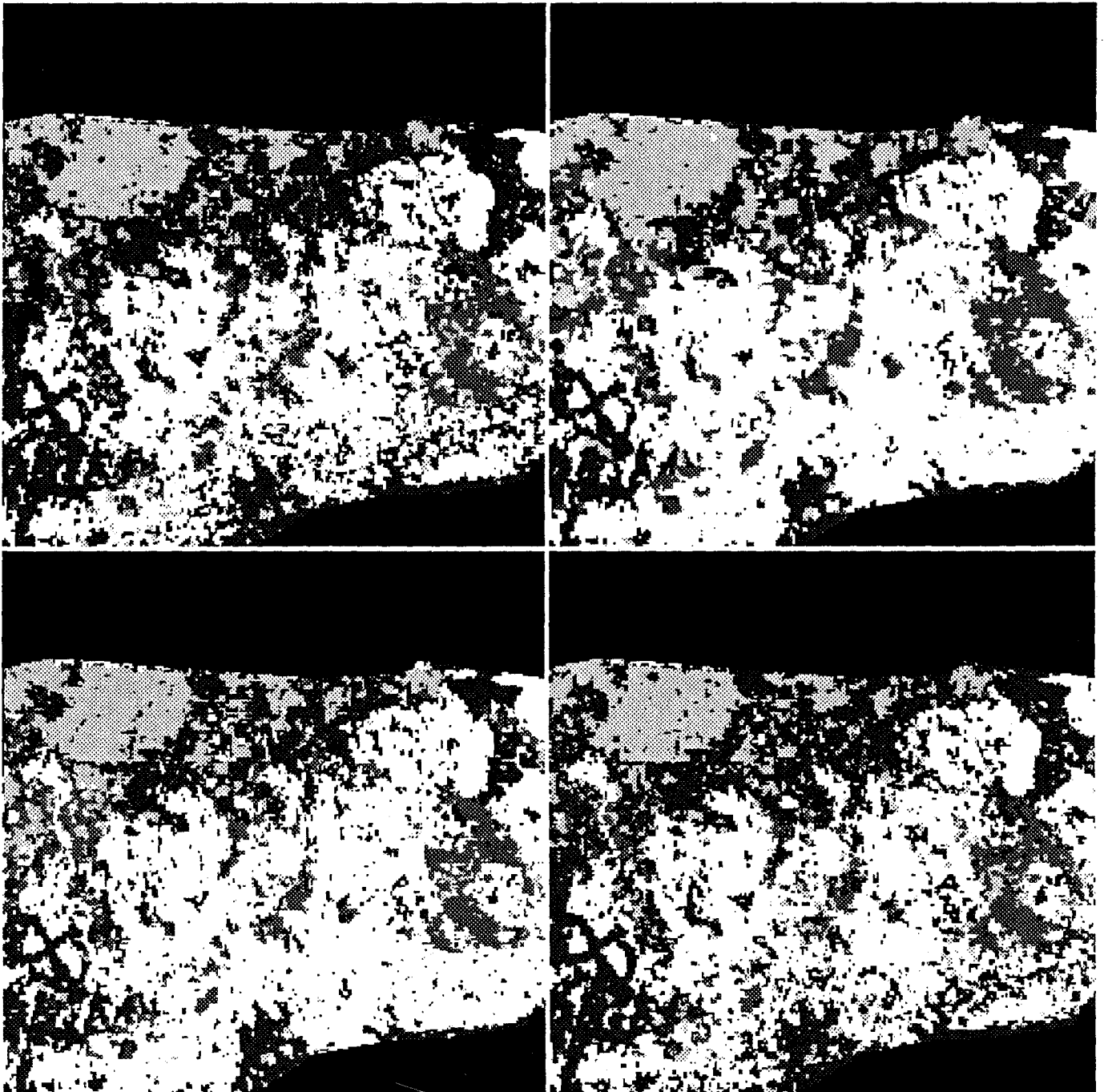
— a) Estimer la température de départ du recuit simulé en fonction de la « qualité » de l'image originale (basse si l'image est peu bruitée, forte sinon), ainsi que la température finale à atteindre (en fonction d'un taux final de contours par exemple).

— b) Examiner le comportement de l'image au voisinage d'une température critique. On sait en particulier qu'en 4 connexité et en champ magnétique nul, l'énergie moyenne par site possède une discontinuité à la température critique quand le nombre de classe  $q$  est supérieur à 4 ! Les propriétés critiques sont connues pour toute valeur de  $q$ , excepté  $q = 3$  [5]. Ainsi la température critique vérifie

$$(35) \quad \exp \frac{K}{T_c} = 1 + \sqrt{q}.$$

L'énergie par site au point critique, ainsi que son saut lorsque  $q \geq 4$  peuvent être calculés analytiquement [5].





1	2
3	4

Fig. 4.

Etude d'une image de roche.

- 1) Reclassification de cette image par recuit simulé avec la matrice  $[B]_3$ .
- 2) Reclassification de cette image par recuit simulé avec la matrice  $[B]_4$  (voir texte).

- 3) Reclassification de cette image par recuit simulé avec la matrice  $[B]_5$  (voir texte).

- 4) Reclassification de cette image par recuit simulé avec la matrice  $[B]_6$  (voir texte).

Ces caractéristiques critiques peuvent d'ailleurs facilement être retrouvées par l'étude d'une série d'échantillons de Gibbs effectués à diverses températures. Il serait nécessaire de généraliser ces résultats en présence d'un « champ magnétique » uniforme (ou uniforme par morceaux dans le cas de la restauration), ainsi que de les étendre au cas de la 8-connexité. l'ensemble des recherches menées dans ce domaine fait apparaître l'importance du nombre de classes *a priori* pour l'analyse d'images par des méthodes markoviennes [13] !

Manuscrit reçu le 18 novembre 1991.

## Remerciements

Les auteurs tiennent à remercier l'ensemble du Département Image de Télécom Paris pour les fructueuses discussions et les nombreux conseils qui ont mené à la réalisation de ce travail.

## BIBLIOGRAPHIE

- [1] A. ROSENFELD, R. A. HUMMEL and S. W. ZUCKER, « Scene Labelling by Relaxation Operation », *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 6, p. 420-433, 1976.
- [2] J. KITTNER and J. FÖGLEIN, « Contextual classification of multispectral data », *Image and Vision Computing*, Vol. 2, No. 1, p. 13-29, 1984.
- [3] S. GEMAN and D. GEMAN, « Stochastic relaxation, Gibbs Distribution, and the Bayesian restoration of Images », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, p. 721-741, 1984.
- [4] N. KARSSMEIJER, « A Relaxation method for Image Segmentation using a Spatially Dependant Stochastic Model », *Pattern recognition Letters*, Vol. 11, No. 1, p. 13-23, 1990.
- [5] F. Y. WU, « The Potts Model », *Review of Modern Physics*, Vol. 54, No. 1, 1982.
- [6] B. CHALMOND, « An iterative Gibbsian technique for reconstruction of m-ary images », *Pattern recognition*, Vol. 22, No. 6, p. 747-762, 1989.
- [7] D. GEMAN, S. GEMAN, C. GRAFFIGNE and P. DONG, « Boundary Detection by Constrained Optimization », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 7, p. 609-628, 1990.
- [8] P. CARNEVALI, L. COLETTI and S. PATARNELLO, « Image processing by Simulated Annealing », *IBM Journal of Research and Development*, Vol. 29, No. 6, p. 569-579, 1985.
- [9] J. BESAG, « On the Statistical Analysis of Dirty Pictures », *Journal of the Royal Statistical Society*, Vol. B-48, p. 259-302, 1986.
- [10] DERIN and H. ELLIOTT, « Modeling and Segmentation of Noisy and Textured Images using Gibbs Random Fields », *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol. 9, No. 1, p. 39-55, 1987.
- [11] P. A. KELLY, H. DERIN and K. D. HARTT, « Adaptive Segmentation of Speckled Images Using a Hierarchical random Field Model », *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 26, No. 10, p. 1628-1641, 1988.
- [12] N. KARSSMEIJER, « A Statistical Method for Automatic Labelling of Tissues in Medical images », *Machine Vision and Applications*, Vol. 3, No. 2, p. 75-86, 1990.
- [13] T. N. PAPPAS and N. S. JAYANT, « An Adaptive Clustering Algorithm for Image Segmentation », *ICASSP Conference Glasgow*, p. 1667-1669, 1989.