

# Estimation hiérarchique semi-itérative du Mode des Marginales a Posteriori en imagerie

Annabelle CHARDIN, Patrick PÉREZ

IRISA/INRIA-Rennes

Campus de Beaulieu, F-35042 Rennes cedex, France

Annabelle.Chardin@irisa.fr, Patrick.Perez@irisa.fr

**Résumé** – Dans le contexte des problèmes inverses en analyse d’images, cet article aborde les modèles stochastiques hiérarchiques qui permettent d’améliorer les approches basées sur les champs de Markov. Nous introduisons une nouvelle structure hiérarchique hybride qui permet de combiner des *a priori* spatial et hiérarchique : il s’agit d’un quadarbre tronqué dont les nœuds du niveau le plus grossier supportent un champ de Markov spatial sur une grille. Sur cette structure, l’estimation des inconnues au sens du Mode des Marginales a Posteriori (MPM) est calculée de manière semi-itérative. Un gain en qualité et en encombrement mémoire est ainsi obtenu pour des temps de calcul proches de ceux des algorithmes non-itératifs sur un arbre complet.

**Abstract** – Within the framework of inverse problems from image analysis, this paper meets with the hierarchical stochastic models which allow the improvement of Markov-Random-Field-based approaches. We introduce a new hybrid hierarchical structure which allows us to combine spatial and hierarchical *a priori*: it is a truncated quadtree whose coarsest level nodes support a grid-based MRF. On this structure the Mode of Posterior Marginals (MPM) inference is conducted in a semi-iterative way. Gains in terms of both quality of results and memory requirements are thus provided for a computational load near the one of non-iterative algorithms on a complete tree.

## 1 Contexte : Les modèles hiérarchiques

Il est possible de réduire la plupart des problèmes inverses d’analyse d’images à l’étude d’une fonction d’énergie  $U(x, y)$  qui traduise les interactions entre des variables mesurées  $y = (y_j)_j$  et de nombreuses variables  $x = (x_i)_i$ , inconnues, à déterminer. Dans une approche markovienne  $x$  et  $y$  sont définis comme des vecteurs aléatoires et nous avons alors la relation suivante entre leur distribution jointe et la fonction d’énergie :

$$\mathbf{P}(x, y) \propto \exp\{-U(x, y)\}.$$

La théorie de l’estimation Bayésienne propose deux estimateurs principaux pour déterminer  $x$  : l’estimateur du Maximum A Posteriori (MAP) qui correspond au minimum global de la fonction d’énergie ( $\hat{x} = \arg \max_x \mathbf{P}(x|y) = \arg \min_x U(x, y)$ ) et l’estimateur du Mode des Marginales a Posteriori (MPM) ( $\forall i, \hat{x}_i = \arg \max_{x_i} \mathbf{P}(x_i|y)$ ) qui nécessite le calcul des probabilités marginales en sommant sur un très grand nombre de configurations.

Bien que la fonction d’énergie puisse généralement se décomposer en une somme de termes locaux ne mettant en jeu qu’un petit nombre de variables, il est généralement impossible d’obtenir une estimation exacte de  $x$  selon le critère du MAP ou du MPM. Des algorithmes itératifs d’approximation déterministes ou stochastiques ont été imaginés, mais ils exhibent souvent un coût calculatoire très important. Ce constat a motivé la recherche de modèles plus spécifiques induisant des algorithmes d’estimation plus efficaces.

C’est pour cette raison, entre autres, que les modèles hiérarchiques ont été introduits. Certains d’entre eux sont basés sur les nœuds d’un quadarbre dont les feuilles sont associées aux pixels des images (de résolution maximale)[2, 5, 6, 7]. Dans ce cas, ils permettent d’obtenir des estimations exactes à l’aide de procédures non-itératives en deux passes. Par conséquent, de tels modèles hiérarchiques apportent un gain en temps de calcul. Ils peuvent aussi améliorer la qualité des résultats. Toutefois, les contraintes imposées par leur structure peuvent paraître artificielles pour certains types de problèmes ou de données ; en particulier comment interpréter les estimées au niveau des résolutions les plus grossières ? De plus, la manipulation d’une structure d’arbre complet peut devenir très lourde dans le cas d’images de très grande taille.

Pour éviter ces inconvénients, nous proposons ici un modèle hiérarchique basé sur une structure hybride qui est constituée d’une grille spatiale de taille réduite au niveau le plus grossier, à laquelle des sous-arbres sont rattachés (voir Fig.1). Les feuilles de ces derniers correspondent au niveau le plus fin. L’étude de l’estimateur du MAP sur cette structure a déjà été menée dans [4]. Cependant, l’estimateur du MPM est connu pour être plus pertinent que le MAP et fournit les probabilités marginales *a posteriori* qui sont une brique indispensable pour les techniques d’estimation de paramètres de type EM. Pour cette raison, nous étudions maintenant l’estimateur du MPM sur cette structure hybride.

TAB. 1: Estimation semi-itérative du MPM sur l'arbre tronqué

<p>▲ <u>pas</u>se montante</p> <p>Feuilles (<math>i \in S^N</math>):</p> $\mathbb{F}_i(x_{\bar{\tau}}, y_i) = \sum_{x_i} h_i(y_i, x_i) f_i(x_i, x_{\bar{\tau}})$ <p>Récurrance (pour <math>n = N - 1 \dots 1</math>, <math>i \in S^n</math>):</p> $\mathbb{F}_i(x_{\bar{\tau}}, y_{\underline{i}}) = \sum_{x_i} h_i(y_i, x_i) f_i(x_i, x_{\bar{\tau}}) \prod_{j \in \underline{i}} \mathbb{F}_j(x_i, y_{\underline{j}})$ <p>◀ calcul des marginales <i>a posteriori</i> au niveau le plus grossier:</p> <p>génération d'échantillons <math>x_{S^0}(1), \dots, x_{S^0}(m)</math> à partir de:</p> $P(x_{S^0}   y) \propto \prod_{i \in S^0} h_i(x_i, y_i) \times \prod_{\langle i, j \rangle \in S^0} g_{ij}(x_i, x_j) \times \prod_{i \in S^0} \prod_{j \in \underline{i}} \mathbb{F}_j(x_i, y_{\underline{j}})$ <p>approximation de <math>P(x_i   y) \approx \frac{1}{m-k} \sum_{j=k+1}^m \delta(x_i(j) - x_i)</math></p> <p>▼ <u>pas</u>se descendante</p> <p>Récurrance (pour <math>n = 1 \dots N</math>, <math>i \in S^n</math>):</p> $P(x_i   y) = \sum_{x_{\bar{\tau}}} P(x_{\bar{\tau}}   y) \frac{h_i(y_i, x_i) f_i(x_i, x_{\bar{\tau}})}{\mathbb{F}_i(x_{\bar{\tau}}, y_{\underline{i}})} \prod_{j \in \underline{i}} \mathbb{F}_j(x_i, y_{\underline{j}})$ <p>MPM au niveau des feuilles:</p> $\hat{x}_i \triangleq \arg \min_{x_i} P(x_i   y)$
---

## 2 Modèle hiérarchique hybride et MPM semi-itératif

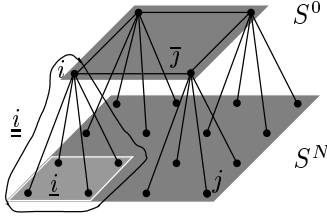


FIG. 1: Exemple de structure hiérarchique hybride

Le modèle hiérarchique que nous utilisons est basé sur la structure hybride décrite dans [4]. Un exemple est présenté sur la figure 1 pour un seul niveau au-dessous de la grille grossière. Brièvement, la structure hybride peut se décrire à travers la définition d'une grille rectangulaire  $S^0$  dont chaque site est la racine d'un quadarbre, de telle sorte que les nœuds du niveau  $n$  ( $0 < n \leq N$ ) forment une grille  $S^n$ ,  $2^n \times 2^n$  fois plus grande que  $S^0$ . Nous notons  $\bar{\tau}$  le parent du site  $i$  (si  $i$  n'est pas au niveau le plus grossier),  $\underline{i}$  l'ensemble des enfants de  $i$  (si  $i$  n'est pas au niveau le plus fin), et enfin  $\underline{j}$  l'ensemble des sites formant l'arbre de racine  $i$ . Les vecteurs  $x$  et  $y$  sont maintenant indexés par les nœuds de  $S \triangleq \bigcup_{n=0}^N S^n$ . Enfin la distribution *a posteriori* est supposée de la forme suivante :

$$P(x, y) \propto \underbrace{\prod_{\langle i, j \rangle \in S^0} g_{ij}(x_i, x_j) \prod_{i \notin S^0} f_i(x_i, x_{\bar{\tau}}) \prod_{i \in S} h_i(x_i, y_i)}_{=\exp\{-U(x, y)\}}$$

où  $\langle i, j \rangle$  désigne les paires de voisins dans  $S^0$ ,  $g_{ij}$  et  $f_i$  sont des fonctions locales qui expriment l'*a priori* respectivement spatial et hiérarchique (ordinairement ils privilégient des estimées similaires entre voisins et entre parent et enfants, resp.), et  $h_i$  représente la relation point-à-point entre les variables observées  $y_i$  et les inconnues  $x_i$ .

Dans le cas d'un arbre complet où  $S^0$  est réduite à un seul site, l'estimation exacte du MPM peut être calculée pour chaque nœud [5] en utilisant une extension de l'algorithme de Baum-Welch sur une chaîne [1]. Dans le cas

général, la récurrence descendante sera maintenant basée sur la relation suivante :

$$\forall i \notin S^0, P(x_i | y) = \sum_{x_{\bar{\tau}}} P(x_i | x_{\bar{\tau}}, y) P(x_{\bar{\tau}} | y)$$

où  $P(x_i | x_{\bar{\tau}}, y) = P(x_i | x_{\bar{\tau}}, y_{\underline{i}})$  d'après la propriété markovienne de séparation.

$P(x_i | x_{\bar{\tau}}, y_{\underline{i}})$ ,  $\forall i \in S \setminus S^0$ , et  $P(x_i | y)$ ,  $\forall i \in S^0$ , doivent être calculés lors d'une passe montante préalable, par sommations successives sur les  $x_i$  tels que  $i \notin S^0$ . A l'issue de cette première étape, nous obtenons à une constante multiplicative près la loi *a posteriori* associée au niveau le plus grossier  $P(x_{S^0} | y) = \sum_{x_{S \setminus S^0}} P(x | y)$ . La structure non-causale de  $S^0$  impose d'approximer les  $P(x_i | y)$  en échantillonnant la distribution  $P(x_{S^0} | y)$  à l'aide d'un échantillonneur de Gibbs. La procédure complète est décrite dans le tableau 1.

## 3 Comparaisons de classifications supervisées

Nous illustrons la pertinence de notre approche au travers du traitement d'un problème de classification supervisée. Pour cette étude de cas, nous avons choisi un *a priori* de type Potts:  $g_{ij}(x_i, x_j) \triangleq \exp\{-2^N \alpha [1 - \delta(x_i, x_j)]\}$  et  $f_i(x_i, x_{\bar{\tau}}) \triangleq \exp\{-\beta [1 - \delta(x_i, x_{\bar{\tau}})]\}$ , avec des vraisemblances gaussiennes pour des données localisées uniquement au niveau le plus fin :

$$h_i(x_i = k, y_i) \triangleq \begin{cases} (2\pi\sigma_k^2)^{-\frac{1}{2}} \exp\left(-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}\right) & \text{si } i \in S^N, \\ 0 & \text{sinon,} \end{cases}$$

où  $\delta$  est la fonction de Kroneker.

Des comparaisons ont été menées pour  $N \in \{0, 2, 3, 4, p\}$  (si la taille de l'image est  $2^p \times 2^p$ ) avec des images synthétiques et des images réelles satellitaires. Nous pouvons remarquer que pour  $N = 0$  il s'agit du modèle standard non-hiérarchique utilisant un échantillonneur de Gibbs, tandis que  $N = p$  correspond à l'arbre complet ( $|S^0| = 1$ ) et permet une estimation exacte non-itérative des marginales *a posteriori* [5].

### 3.1 Images synthétiques

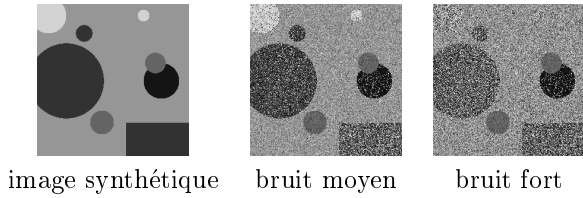


FIG. 2: Données synthétiques

Nous désirons partitionner une image synthétique  $256 \times 256$  en 5 classes (Fig.2). L'image a été bruitée avec un bruit blanc gaussien additif d'écart-type différent pour chaque classe. Ainsi  $(\mu_k, \sigma_k^2)_{k=1}^5$ , les moyennes et les variances en luminance, sont connues. Nous avons choisi d'utiliser deux images bruitées. La deuxième image présentant des variances supérieures, il nous sera ainsi possible d'étudier la robustesse de notre approche par rapport au bruit.

Les résultats sont présentés sur les figures 3 et 4 avec leur taux de bonne classification respectifs et leur temps de calcul (station Sun Ultra Sparc I 170 MHz). De plus, il faut souligner le fait que les algorithmes du MPM donnent accès en chaque site à une mesure de confiance vis-à-vis de l'étiquette estimée. Cette mesure (Fig.5), basée sur un calcul d'entropie

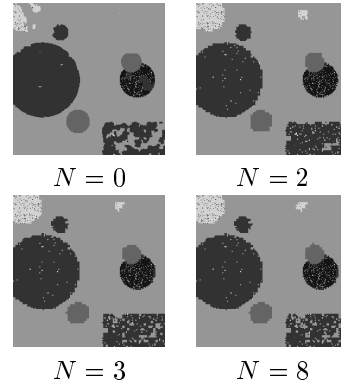
$$c_i \triangleq - \sum_{x_i} P(x_i|y) \log P(x_i|y) ,$$

permet de mieux apprécier la qualité des estimations et d'adapter leur utilisation en conséquence.

Tout d'abord, quant à l'aspect visuel des classifications hiérarchiques, il est à noter la présence d'effets de blocs dus à la structure en quadarbre, mais cet effet est de moins en moins prononcé à mesure que le nombre de niveau dans la structure hybride diminue.

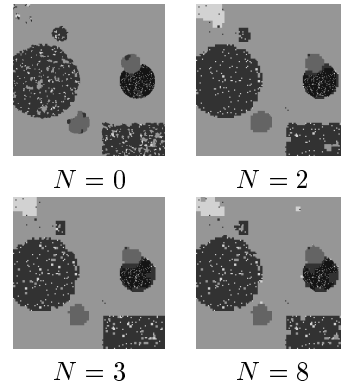
Ensuite, il est remarquable que les méthodes hiérarchiques fournissent de meilleurs résultats que le MPM spatial et demandent beaucoup moins de temps, environ quarante fois moins. Pour l'estimation des marginales *a posteriori* avec la procédure d'échantillonnage, nous avons ajusté le nombre d'échantillons retenus ( $m-k$ , voir synopsis du tab.1) à la taille de la grille concernée (200 pour la grille complète et 20 pour la grille réduite de la structure hybride). Nous pouvons aussi constater que les méthodes hiérarchiques résistent mieux à l'augmentation du bruit dans l'image. En effet, la classe où la variance du bruit a été le plus augmentée (les cercles les plus clairs) n'est quasiment pas présente dans la classification fournie par le MPM spatial et sa carte de confiance présente un aspect beaucoup plus bruité.

Enfin, l'estimation semi-itérative apporte une légère amélioration par rapport au non-itératif, pour des temps comparables et un encombrement mémoire bien moindre, le nombre de sites ayant été réduit. Donc l'utilisation d'une procédure itérative d'échantillonnage au niveau le plus grossier ne semble pas impliquer une surcharge calculatoire significative tout en améliorant les résultats.



Modèle	bonne classif.	temps cpu
itératif $N = 0$	93%	7min
semi-itératif $N = 2$	96.4%	10.5s
semi-itératif $N = 3$	95.5%	9.3s
non-itératif $N = 8$	95.3%	9s

FIG. 3: Résultats comparatifs de classification avec l'image moyennement bruitée



Modèle	bonne classif.	temps cpu
itératif $N = 0$	90%	7min
semi-itératif $N = 2$	93.8%	10.5s
semi-itératif $N = 3$	92.1%	10s
non-itératif $N = 8$	92%	9s

FIG. 4: Résultats comparatifs de classification avec l'image fortement bruitée

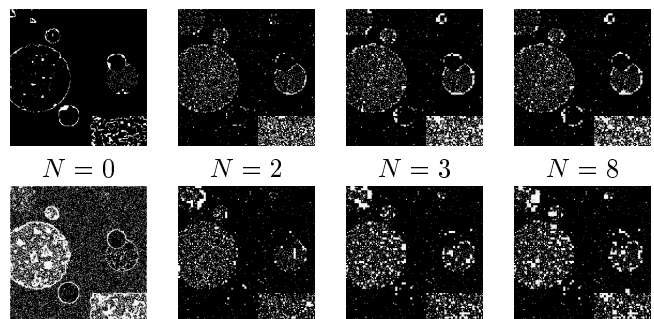


FIG. 5: Cartes de confiance associées aux classifications de l'image moyennement bruitée (au-dessus) et de l'image fortement bruitée (au-dessous).

### 3.2 Images réelles satellitaires

Nous avons poursuivi nos comparaisons sur un exemple réel d'images satellitaires SPOT qui nous ont été fournies par Costel (Université de Rennes2) dans le cadre d'études en télédétection au sein du Groupement Scientifique en Télédétection Bretagne (GSTB). La scène (Fig.6) est composée de trois images  $512 \times 512$  acquises dans trois longueurs d'onde différentes du visible et représente la baie de Lannion (France) en décembre 1996. L'objectif de ces études était de déterminer l'occupation du sol de cette région. Pour cela, les géographes de Costel ont établi une liste de huit classes : Mer et eau, Sable et sols nus, Urbain, Forêts et landes, Prairies artificielles, Prairies naturelles, Colza, Légumes.

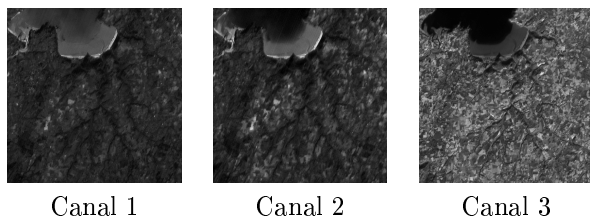


FIG. 6: Images satellitaires SPOT  $512 \times 512$  (avec la gracieuse permission de Costel, Université de Rennes 2, et du GSTB).

Grâce à des prélèvements sur le terrain et à des photo-interprétations, des petites portions d'images « échantillons » représentatives de chacune des classes ont pu être extraites. Nous avons partagé cet ensemble d'échantillons en deux parties : une partie a été utilisée pour l'apprentissage des paramètres des classes et la seconde a été réservée pour évaluer la qualité des classifications (les taux de classification donnés plus loin seront donc calculés sur ces échantillons).

Pour prendre en compte l'aspect multispectral des données, nous considérons les trois canaux indépendants et la fonction d'attache aux données devient le produit des vraisemblances gaussiennes associées à chaque canal.

Les quatre algorithmes présentent des résultats similaires de bonne qualité (Fig.7), puisque environ 92% des pixels des échantillons sont bien classés. Cependant, nous pouvons constater le substantiel avantage qu'apportent les méthodes hiérarchiques en terme de temps de calculs. Ainsi plus d'une heure est nécessaire au MPM itératif, là où les estimations hiérarchiques ne demandent que quelques minutes (moins de 3min). Ceci nous permet d'être optimiste vis-à-vis de l'estimation de paramètres avec des méthodes itératives de type EM pour lesquelles chaque itération nécessite, comme le MPM, le calcul de marginales locales *a posteriori*.

## 4 Extension

Avec l'estimateur du MAP [4], nous avons conclu que la structure hybride était un bon compromis pour l'analyse hiérarchique d'images. Cette étude le confirme avec l'estimateur du MPM. Quant à la suite de ce travail, nous envisageons de traiter le problème critique de l'estimation

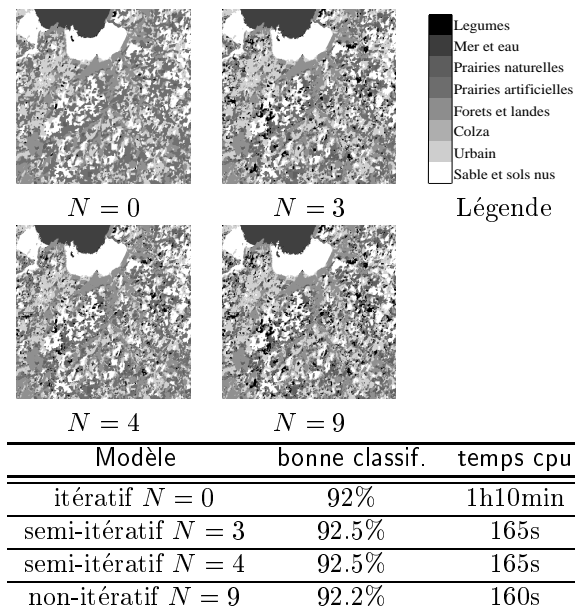


FIG. 7: Résultats comparatifs de classification sur les images SPOT multispectrales.

de paramètres: en utilisant le calcul des marginales introduit ici, nous devrions être capables de construire des algorithmes de type EM spécifiques à la structure hybride, comme cela a déjà été proposé pour les arbres complets [2, 5] et pour les grilles bi-dimensionnelles [3]. Nous avons aussi en projet d'aborder la question de l'estimation automatique du nombre optimal de niveaux dans la structure.

## Références

- [1] L.E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.*, Vol 41: pp. 164–171, 1970.
- [2] C. Bouman and M. Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. Image Processing*, 3, No. 2:162–177, March 1994.
- [3] B. Chalmond. An iterative Gibbsian technique for reconstruction of m-ary images. *Pattern Recognition*, Vol. 22, No 6: pages 747–761, 1989.
- [4] A. Chardin and P. Pérez. Semi-iterative inference with hierarchical models. In *Int. Conf. on Image Processing*, pages 630–634, Chicago, USA, october 1998.
- [5] J-M Laferté, P. Pérez, and F. Heitz. Discrete markov image modeling and inference on the quad-tree. *IEEE Trans. Image Processing*, Accepted for publication, 1999.
- [6] M. Luetzgen, W. Karl, and A. Willsky. Efficient multiscale regularization with applications to the computation of optical flow. *IEEE Trans. Image Processing*, 3(1):41–64, 1994.
- [7] P. Pérez, A. Chardin, and J-M. Laferté. Noniterative manipulation of discrete energy-based models for image analysis. *Pattern Recognition*, Accepted for publication, 1999.