

## Fusion d'images multirésolutions par modèles markoviens hiérarchiques.

J.-M. Laferté, F. Heitz, P. Pérez & E. Fabre

IRISA/INRIA - Rennes , Campus de Beaulieu 35042 Rennes Cedex, France  
FAX : (33)99-84-71-71, laferte@irisa.fr

### RÉSUMÉ

Nous développons dans cet article deux algorithmes basés sur des modèles statistiques hiérarchiques non-linéaires pour la fusion d'images multirésolutions. Les processus mis en jeu sont définis sur un graphe hiérarchique simple (quadarbre) et sont supposés markoviens en échelle. Cette structure pyramidale permet d'induire des modèles causaux et ainsi d'en dériver des algorithmes d'optimisation non itératifs. Le premier algorithme de type Viterbi donne la solution exacte au sens du MAP. Le second donne une solution approchée au sens d'un estimateur plus pertinent que le MAP (le SMAP). Ces deux algorithmes sont comparés expérimentalement sur un problème de classification (images synthétiques et réelles).

### 1. INTRODUCTION

La fusion d'observations multisources, dont l'intérêt n'est plus à démontrer, pose de nombreuses difficultés provenant en particulier des caractéristiques variées des capteurs. En imagerie, cela se traduit en particulier par des images provenant de domaines spectraux différents et définies à des résolutions variables (imagerie satellitaire, aérienne ...).

Dans cet article, nous présentons et comparons deux algorithmes destinés à fusionner des images multirésolutions sur un graphe hiérarchique. Les algorithmes développés reposent sur une modélisation statistique hiérarchique des liens de dépendance entre données multirésolutions [6]. Le caractère causal des modèles adoptés permet la dérivation d'algorithmes d'optimisation non-linéaire non itératifs. Ces modèles offrent ainsi une alternative séduisante aux champs markoviens (modèles non causaux), qui conduisent à des algorithmes de relaxation généralement prohibitifs d'un point de vue calculatoire [3, 5]. Un premier algorithme de Viterbi sur l'arbre, permettant de calculer une estimée exacte au sens du Maximum A Posteriori (MAP) d'un champ d'étiquettes est tout d'abord décrit. Un second algorithme basé sur un estimateur plus perti-

### ABSTRACT

This paper presents two non-linear hierarchical algorithms for the fusion of multiresolution image data in low-level vision. The approach combines non-linear causal Markov models defined on hierarchical graph structures, with standard bayesian estimation theory. An optimal algorithm (inspired from the Viterbi algorithm) is developed to compute the MAP estimates on the quadtree. An other multiscale algorithm based on a more relevant estimator (the SMAP estimator) on the same structure is presented. These algorithms are non-iterative and yield a per pixel computational complexity which is independent of image size. They are compared on a classification task on synthetic and real images.

nent (le SMAP) [2] est ensuite présenté et comparé sur un problème de classification d'images.

### 2. MODÈLES STATISTIQUES

#### 2.1. Notations

On cherche à estimer un ensemble de variables  $X$  dites cachées et appelées *étiquettes* à partir d'observations  $Y$ .  $X$  et  $Y$  sont des processus aléatoires indicés par les sommets d'un quadarbre  $\tau$  (figure 1). On note  $s \tilde{\gamma}$  l'unique "père" de  $s$ , l'ensemble des fils de  $t$  est appelé  $\mathcal{D}(t)$ ,  $\mathcal{D}(t) \triangleq \{s : t = s \tilde{\gamma}\}$ . Par ailleurs l'ensemble  $> s$  désigne le "sous-arbre" de  $\tau$  dont la racine  $s$  est exclue ( $\geq s$  pour  $s$  inclus). On définit également l'ensemble des ancêtres de  $s$  par  $< s \triangleq \{s \tilde{\gamma}, s \tilde{\gamma} \tilde{\gamma}, \dots, r\}$ ,  $r$  désignant la racine du quadarbre. On désigne par  $S$  l'ensemble des sommets de  $\tau$ . Les sommets de l'arbre appartenant à la même "génération" forment un ensemble noté  $S^n$  où  $n$  représente l'échelle, l'échelle 0 correspondant à la résolution la plus fine et l'échelle  $L$  (résolution dite la plus grossière) correspondant au singleton  $S^L = \{r\}$  (cf figure 1).

La restriction de  $X$  à  $S^n$  est notée  $X^n \triangleq \{X_s, s \in S^n\}$ , où  $X_s$  représente l'étiquette en un site  $s$ . On définit

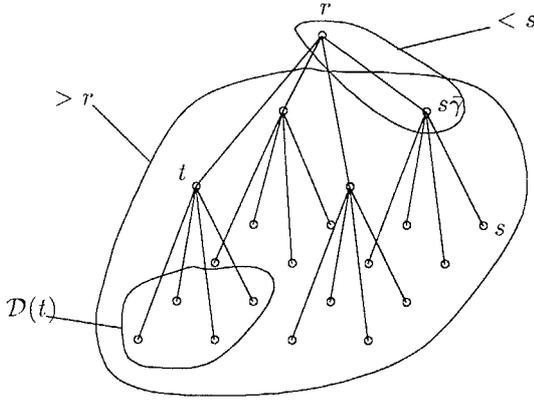


FIG. 1 - Le quadarbre  $\tau$

de même les ensembles d'observations  $Y^n$  (et les variables  $Y_s$ ) à ceci près que  $Y_s$  n'existe pas nécessairement dans le cas de données manquantes dans une image ou de jeu de données inexistant à une résolution donnée. Les observations  $Y^n$ , lorsqu'elles existent, correspondent à une ou plusieurs image(s) à l'échelle  $n$ . Dans le cas général  $Y_s$  est donc un vecteur qui peut représenter par exemple une information multispectrale ( $Y^n$  est alors un  $d$ -uplet d'images définies à la même résolution), de même  $X_s$  peut être vectoriel.

### 2.2. Hypothèses sur le modèle statistique

On considère donc le processus  $X = (X^n)_{n=0\dots L}$  et  $X_s$  à valeurs dans  $\Lambda$ , espace d'états supposé discret. Nous faisons les hypothèses suivantes sur les distributions intervenant dans ce modèle.

- L'hypothèse fondamentale consiste à considérer le processus aléatoire  $X$  markovien en échelle, *i.e.* :

$$P(X^n|X^i, i > n) = P(X^n|X^{n+1})$$

- Les probabilités de transitions inter-échelles sont supposées se factoriser de la façon suivante :

$$P(X^n|X^{n+1}) = \prod_{s \in S^n} g_n(X_s^n|X_{s\bar{\gamma}}^{n+1})$$

- D'autre part nous supposons une dépendance ponctuelle de  $Y$  relativement à  $X$ , c'est-à-dire :

$$P(Y|X) = \prod_n P(Y^n|X^n) = \prod_n \prod_{s \in S^n} f^n(Y_s|X_s)$$

(on dit que  $Y$  est semi-markovien par rapport à  $X$ ). On prendra pour convention  $f^n(y_s^n|\cdot) \equiv 1$  si  $s$  ne porte pas d'observation.

À partir de ces hypothèses on déduit que  $Z = (X, Y)$  est markovien sur le quadarbre, ce qui implique<sup>1</sup> :

$$\forall A, B, C \subset S, \text{ si } B \text{ sépare } A \text{ et } C, \text{ alors} \quad (1)$$

$$P(Z_A, Z_C|Z_B) = P(Z_A|Z_B) P(Z_C|Z_B) \quad (2)$$

### 3. ALGORITHMES BAYESIENS

Nous considérons comme critère d'estimation le critère bayésien de minimisation de l'espérance d'une fonction de coût conditionnellement aux observations :

$$\hat{x} = \arg \min_x \mathbb{E}(C(X, x)|Y = y) \quad (3)$$

où  $X$  est une variable aléatoire représentant la solution exacte et  $C(X, x)$  représente le prix à payer quand on choisit  $x$  alors que la solution exacte est  $X$ . Différents estimateurs bayésiens peuvent être choisis. Nous en proposons deux, correspondant à deux fonctions de coût différentes.

#### 3.1. Estimateur du MAP et algorithme de Viterbi

On cherche tout d'abord à estimer  $X$  au sens du MAP :

$$\hat{x} \triangleq \arg \max_x P(x|y) = \arg \max_x P(x, y)$$

On a

$$\begin{aligned} \max_x P(x, y) &= \max_{x_{\geq r}} P(x_{\geq r}, y_{\geq r}) \quad (4) \\ &= \max_{x_r} \left\{ P(x_r, y_r) \max_{x_{>r}} P(x_{>r}, y_{>r}|x_r) \right\} \\ &= P(\hat{x}_r, y_r) \prod_{s \in \mathcal{D}(r)} \max_{x_{\geq s}} P(x_{\geq s}, y_{\geq s}|\hat{x}_r) \end{aligned}$$

si  $\hat{x}_r$  est l'étiquette réalisant le maximum dans (4).

On voit alors apparaître la formule de récurrence :

$$\begin{aligned} P_s^*(x_{s\bar{\gamma}}) &\triangleq \max_{x_{\geq s}} P(x_{\geq s}, y_{\geq s}|x_{s\bar{\gamma}}) \\ &= \max_{x_s} \left\{ P(x_s, y_s|x_{s\bar{\gamma}}) \prod_{t \in \mathcal{D}(s)} P_t^*(x_s) \right\} \quad (5) \end{aligned}$$

On procède donc ici en deux passes :

- Une première passe montante qui consiste à calculer et "stocker" les  $P_s^*$  en fonction de  $x_{s\bar{\gamma}}$  (on s'arrête juste avant la racine), *i.e.*

- sur  $S^0$  :

$$\begin{aligned} P_s^*(x_{s\bar{\gamma}}) &\triangleq \max_{x_s} P(x_s, y_s|x_{s\bar{\gamma}}) \\ &= \max_{x_s} f^0(y_s|x_s) g^0(x_s|x_{s\bar{\gamma}}) \end{aligned}$$

1. On dit que  $B$  sépare  $A$  et  $C$  ssi pour tout élément de  $A$  et tout élément de  $C$  il n'existe pas de chemin sur le graphe joignant  $A$  et  $C$  ne passant pas par  $B$ .

- sur  $S^n, n < L$ :  $P_s^*$  est calculée à partir des  $P_t^*, t \in S^{n-1}$  et la relation (5).

On mémorise simultanément les  $\hat{x}_s(x_{s\bar{y}})$  réalisant les maxima.

- Une seconde passe descendante qui, partant de la racine, déroule les  $\hat{x}_s$  à l'aide de la première passe, jusqu'aux feuilles :

$$- \hat{x}_r = \arg \max_{x_r} \left\{ f^L(y_r|x_r) \prod_{s \in \mathcal{D}(r)} P_s^*(x_r) \right\} \text{ (on suppose } X_r \text{ uniforme)}$$

- Pour  $s \in S^n, 0 \leq n < L$ ,  $\hat{x}_s$  est l'argument qui réalise  $P_s^*(\hat{x}_{s\bar{y}})$ .

### 3.2. L'estimateur du SMAP

L'estimateur du MAP est malheureusement mal adapté aux modèles hiérarchiques. En effet il pénalise une erreur indépendamment de l'échelle à laquelle elle se produit. Or pour des raisons intuitives on souhaite pénaliser plus sévèrement les erreurs aux échelles grossières. C'est pourquoi on préfère un estimateur du type du SMAP introduit par Bouman [2] et adapté ici à la fusion d'images multirésolutions :

$$C_{SMAP}(X, x) \triangleq \sum_{n=0}^L 2^n \left( 1 - \prod_{i=n}^L \delta(X^i - x^i) \right)$$

On montre alors que le problème (3) revient à chercher :

$$\max_x \sum_{n=0}^L 2^n P(X^i = x^i, i \geq n | Y = y).$$

On définit alors des fonctions de pseudo-vraisemblance :

$$h_s^n(\lambda) \triangleq P(y_{>s} | X_s = \lambda) \quad n \geq 1 \text{ et } \lambda \in \Lambda.$$

les  $h_s^n(\lambda)$  peuvent être calculées par récurrence montante :

$$\begin{aligned} h_s^1(\lambda) &= \prod_{t \in \mathcal{D}(s)} \sum_{\lambda' \in \Lambda} f^0(y_t | \lambda') g^0(\lambda' | \lambda) \\ h_s^n(\lambda) &= \prod_{t \in \mathcal{D}(s)} \sum_{\lambda' \in \Lambda} f^{n-1}(y_t | \lambda') g^{n-1}(\lambda' | \lambda) h_t^{n-1}(\lambda'). \end{aligned} \quad (6)$$

En supposant la distribution  $X_r$  uniforme et en exploitant la markovianité de  $(X, Y)$ , on obtient *in fine* :

$$\begin{aligned} \hat{x}_r &\simeq \arg \max_{k \in \Lambda} h_r^L(k) f^L(y_r | k), \\ \hat{x}_s &\simeq \arg \max_{k \in \Lambda} h_s^n(k) f^n(y_s | k) g^n(k | \hat{x}_{s\bar{y}}), \quad 0 < n < L \quad (7) \\ \hat{x}_s &\simeq \arg \max_{k \in \Lambda} f^0(y_s | k) g^0(k | \hat{x}_{s\bar{y}}), \quad n = 0. \end{aligned}$$

ALGO	Résolutions des images	SNR dB	Taux de bonne classif.
MAP	256 × 256	-10	<b>96.0</b>
SMAP	256 × 256	-10	<b>97.4</b>
MAP	256 × 256	-16	<b>73</b>
SMAP	256 × 256	-16	<b>84</b>
MAP	256 × 256 64 × 64	-16 1	<b>82</b>
SMAP	256 × 256 64 × 64	-16 1	<b>93</b>

TAB. 1 - Résultats sur les images synthétiques.

L'algorithme se déroule donc également en deux passes : la première (montante) consiste à calculer les fonctions  $h_s^n(\cdot)$  pour tous les sommets de l'arbre et la seconde (descendante) consiste à résoudre à chaque site l'équation (7).

## 4. CLASSIFICATION

On cherche à partitionner une image en  $M$  ( $M$  est supposé connu *a priori*) classes, *i.e.*  $\Lambda = \{1, \dots, M\}$ . Les fonctions de vraisemblance  $f$  sont gaussiennes et

$$g^n(m|k) \triangleq \begin{cases} \theta_n & \text{si } m = k \\ \frac{1-\theta_n}{M-1} & \text{sinon} \end{cases} \quad (8)$$

les  $\theta_n$  étant des paramètres estimés par un algorithme de type EM [1].

Les images synthétiques représentent des formes circulaires auxquelles on a ajouté un bruit gaussien. Comme attendu, les meilleurs résultats sont obtenus par le SMAP (cf table 1). La fusion de deux images est dans tous les cas plus robuste au bruit. Malheureusement la structure de quadarbre induit des effets de bloc plus ou moins visibles sur les résultats (figure 2). Afin de réduire ces artéfacts, on peut ajouter des liens inter-échelles (voir par exemple [4]) mais ceci complique considérablement le problème.

On a appliqué les algorithmes précédents à des images aériennes de la région de Saint-Louis prises lors de la crue historique des fleuves Mississippi et Missouri en juillet 1993. Nous considérons ici 4 classes : les deux fleuves, la zone urbaine et la végétation. Les paramètres des classes (moyennes et variances) ont été estimés par apprentissage sur un échantillon représentatif.

Les champs d'étiquettes sont estimés en utilisant l'algorithme SMAP qui, comme on peut le voir figure 3, affine graduellement la classification en intégrant et fusionnant progressivement l'information. Si l'on ne

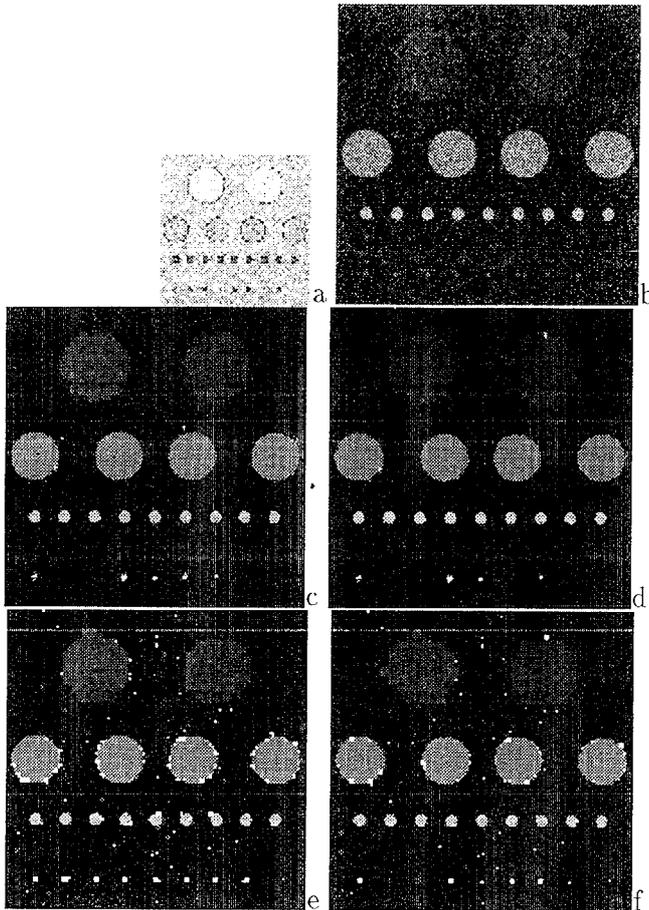


FIG. 2 - Fusion par modèle markovien hiérarchique (données synthétiques) a) Image  $64 \times 64$  (rapport signal sur bruit : 1 dB). b) Image  $256 \times 256$  (rapport signal sur bruit : -10 dB). c) Résultat du critère SMAP avec les deux images. d) Résultat du critère SMAP avec l'image b. e) Résultat du critère MAP avec les deux images. f) Résultat du critère MAP avec l'image b.

prend en compte qu'une seule image, les deux fleuves ne sont pas discriminés.

## 5. CONCLUSION

Nous avons présenté deux algorithmes non-itératifs pour la fusion d'images multirésolutions sur un quadarbre. Si la simplicité de la structure hiérarchique induit des algorithmes simples et efficaces, en contrepartie la présence d'artéfacts importants (effets de bloc) sur les estimées, caractéristiques des quadarbres, demeure un problème majeur. L'étude de graphes plus complexes, bien que posant de nombreuses difficultés théoriques, s'avère donc nécessaire et est actuellement en cours.

## 6. RÉFÉRENCES

- [1] N. M. LAIRD A. P. DEMSTER and D. B. RUBIN. - Mixtures densities, maximum likelihood from incom-

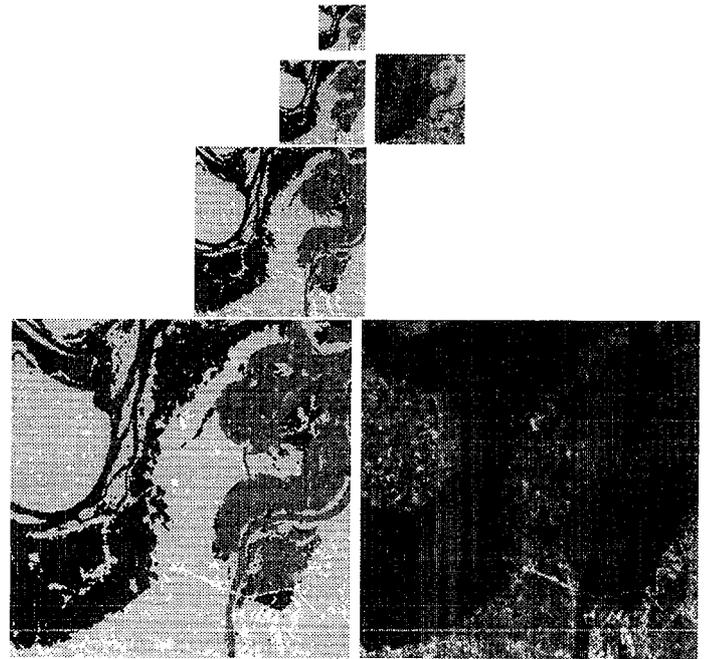


FIG. 3 - Résultat de l'algorithme du SMAP pour les images "Mississippi". Les images sont à droite (image fine  $512 \times 512$ , image grossière  $128 \times 128$ ), les cartes de classification à gauche (à une couleur correspond une classe). Temps CPU 1mn40s sur une station de travail (Sun SPARC 10).

plete data via the EM algorithm. - *Journal of the Royal Stat. Society*, Vol. 39, No 1: pages 1-38, 1977.

- [2] C. BOUMAN and M. SHAPIRO. - A multiscale random field model for bayesian image segmentation. - *IEEE Trans. on Image Processing*, Vol. 3, No. 2: pages 162-177, March 1994.
- [3] Z. KATO, M. BERTHOD, and J. ZERUBIA. - Multiscale Markov random field models for parallel image classification. - In *Proc. Int. Conf. Computer Vision*, pages 253-257, Berlin, May 1993.
- [4] J.-M. LAFERTÉ, F. HEITZ, P. PÉREZ, and E. FABRE. - Hierarchical statistical models for the fusion of multiresolution image data. - In *Proc. Int. Conf. on Computer Vision*, Cambridge, USA, 1995.
- [5] J.-M. LAFERTÉ, P. PÉREZ, and F. HEITZ. - Global non-linear multigrid optimization for image analysis tasks. - In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 533-536, Adelaide, Australia, 1994.
- [6] M. LUETTGEN, W. KARL, A. WILLSKY, and R. TENNEY. - Multiscale representations of Markov Random Fields. - *IEEE Trans. Signal Processing*, Vol. 41, No 12: pages 3377-3395, Dec. 1993.