

**SEGMENTATION BAYESIENNE PAR ECHANTILLONNAGE  
BOOTSTRAP  
APPLICATION A UNE SEQUENCE D'IMAGES  
DU FOND DE L'OEIL**

L. MOREAU, P.M. ALLIOUX, C. BANGA et F. GHORBEL

Groupe Image de l'Institut National des Télécommunications et de l'Ecole Nouvelle d'Ingénieurs en Communication,  
Cité Scientifique Rue Guglielmo MARCONI, 59650 Villeneuve d'Ascq, France

RÉSUMÉ

ABSTRACT

Dans le développement de cet article, nous proposons l'utilisation des méthodes EM pour la segmentation d'une séquence d'images du fond de l'oeil. Un échantillonnage bootstrap, consistant en une sélection aléatoire des pixels de l'image originale, est présenté. Cette méthode amène d'une part à une meilleure estimation des paramètres lors de l'identification du mélange grâce à la diminution de la dépendance au sein de l'échantillon et, d'autre part, à une réduction considérable du temps de calcul (au moins d'un facteur 260) puisqu'une optimisation de la taille de l'échantillon, au sens d'un critère de représentativité, devient possible. Ces gains permettent également l'exploration de segmentations contextuelle et spatio-temporelle réputées pour leur exigences en termes de volume de calculs.

In this paper, we propose using the EM method for the segmentation of an eye's fundus image sequence. A bootstrap sampling is presented, consisting on random pixel choice from the original image. This method allows a better estimation of the parameters during the identification of the mixture by reducing the pixel dependency. Further more, a great decrease of the computing time (at least a factor of 260) is allowed by an optimization of the sampling size in the representative criterion meaning. We show in our application the Bootstrap method interest versus the classical method. We propose a Bootstrap sampling scheme for the EM algorithm in the blind and contextual approach.

Le domaine médical est un gros consommateur d'images. Radiologie, échographie, scanner, RMN, etc... sont autant de sources qui alimentent quotidiennement les archives des hôpitaux et occupent un volume considérable[2],[4]. Certaines de ces images sont pauvres en terme de définition, d'autres, très riches: quelques milliers de pixels en endoscopie à plusieurs millions pour une radiologie conventionnelle. Toutes ces informations ne sont pas également pertinentes pour le médecin et il apparaît qu'une technique permettant de réduire leur nombre à celui réellement utile à l'interprétation et à la prise de décision serait bénéfique tant pour le stockage de ces images que pour l'aide au diagnostic. Une méthode telle que JPEG peut réduire de manière considérable le volume brut des informations contenues dans l'image (dans un rapport 10 au moins), à condition d'accepter une certaine dégradation à la limite du perceptible; mais il s'agit d'une approche qui ne tient pas compte de la nature même de l'image. Une autre procédure consiste dès lors à effectuer un classement des pixels de l'image en fonction des grandes structures qu'elle contient et à ne retenir qu'un nombre limité de classes. Certes, le taux de compression obtenu est plus faible qu'avec JPEG mais ce pré-traitement met en relief les zones d'intérêt du médecin et laisse évidemment la possibilité d'appliquer ultérieurement des algorithmes de compression sans dégradation.

tant d'illuminer le réseau vasculaire rétinien. Une segmentation de cette séquence doit permettre de tracer une cartographie du réseau vasculaire, puis, dans une phase ultérieure, de tenter d'évaluer ponctuellement les débits sanguins.

### 1) SEGMENTATION BAYESIENNE

L'opération de segmentation consiste à subdiviser une image en zones homogènes. L'approche statistique consiste à affecter à chaque pixel un label correspondant à une classe. Dans notre cas, nous utilisons la Règle de Classification de Bayes (idéale puisqu'elle minimise la probabilité d'erreur de classification) au prix d'une hypothèse paramétrique sur le modèle de l'image. Elle peut être mise en oeuvre grâce aux algorithmes de type EM (Estimation-Maximisation).

#### a) Les algorithmes EM, SEM, SAEM

Supposons qu'une image soit une réalisation d'un mélange de loi normales. Ceci implique que sa densité de probabilité peut être approchée par :

$$f(x) = \sum_{j=1}^k \pi_j g_j(x), \quad 0 \leq \pi_j \leq 1, \quad \sum_{j=1}^k \pi_j = 1$$

où  $\pi_j$  est la probabilité a priori de la classe  $j$ ,  $g_j$  est la densité de probabilité conditionnelle de la classe  $j$ ,  $k$  est le nombre de classes et  $f$  représente la loi du mélange. La classification bayésienne paramétrique non supervisée nécessite d'abord l'identification du mélange par l'estimation de ses paramètres. Pour ce

Les images que nous traitons proviennent d'un scanning laser ophtalmoscope et sont obtenues lors d'une angiographie d'un fond d'oeil, l'injection d'un produit de contraste permet-



faire, une hypothèse paramétrique est nécessaire. Dans ce type d'image, nous supposons que la probabilité conditionnelle à la classe  $j$  suit une loi normale avec  $\theta_j = (\mu_j, \sigma_j^2)^T$ :

$$g_j(x) = g(x/\theta_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp - \frac{(x - \mu_j)^2}{2\sigma_j^2}$$

L'identification du mélange revient à l'estimation des paramètres  $\{(\pi_j, \theta_j), j = 1, \dots, k\}$ , le maximum de la vraisemblance n'admettant pas de solutions analytiques, un algorithme itératif, appelé estimation-maximisation et noté EM, est utilisé pour la résolution de la vraisemblance. Après la constitution d'un échantillon représentatif de l'image et l'initialisation des probabilités a priori  $\pi_j$ , l'algorithme EM comporte 4 étapes :

- L'étape **estimation**: elle comporte le calcul de la probabilité a posteriori pour que le pixel  $x_i^*$  appartenant à la classe  $j$  à l'itération  $m$  soit donnée par :

$$t_j^m(x_i^*) = \frac{\pi_j^m g(x_i^*/\theta_j^m)}{\sum_{p=1}^k \pi_p^m g(x_i^*/\theta_p^m)}$$

- L'étape **stochastique**: au cours de cette étape, nous construisons une variable aléatoire de Bernoulli de paramètre  $t_j^m(x_i^*)$  que l'on note par  $z_j^m(x_i^*)$ .
- L'étape **recuit**: elle permet de tempérer l'aspect stochastique. A partir de  $z_j^m$  et  $t_j^m$ , une autre variable aléatoire peut être construite de la façon suivante:

$$r_j^m(x_i^*) = t_j^m(x_i^*) + h_m [z_j^m(x_i^*) - t_j^m(x_i^*)]$$

- L'étape **maximisation**: dans cette étape, la variable aléatoire  $r_j^m(x_i^*)$  est artificiellement considérée comme la probabilité a posteriori des  $x_i^*$ . Ainsi, à l'itération numéro  $m$  nous avons:

$$\pi_j^{m+1} = \frac{1}{n} \sum_{i=1}^n r_j^m(x_i^*); \quad \mu_j^{m+1} = \frac{\sum_{i=1}^n x_i^* r_j^m(x_i^*)}{\sum_{i=1}^n r_j^m(x_i^*)}$$

$$(\sigma_j^{m+1})^2 = \frac{\sum_{i=1}^n (x_i^* - \mu_j^{m+1})(x_i^* - \mu_j^{m+1})^T r_j^m(x_i^*)}{\sum_{i=1}^n r_j^m(x_i^*)}$$

Pour ces algorithmes, il n'existe pas de critère d'arrêt. En pratique, nous l'arrêtons au bout d'un certain nombre d'itérations quand les probabilités a priori ne varient presque plus. Les classes ayant des valeurs de  $\pi_j$  faibles ne seront pas considérées.

L'algorithme EM est celui qui ne contient pas d'étape stochastique, c'est dire  $h_m$  nul. Il a été montré que le EM est un bon estimateur du mélange. Cependant, les résultats obtenus dépendent parfois de l'initialisation. Dans ce cas, nous obtenons une mauvaise solution qui peut correspondre à un minimum local ou un col de la vraisemblance.

L'algorithme SEM a été introduit afin de pallier cet inconvénient du EM. Il est obtenu en considérant  $h_m = 1$ . Dans ce cas, la troisième étape n'est pas nécessaire. Le SEM introduit une randomisation au niveau du EM. Il converge plus rapidement à la bonne solution.

L'algorithme SAEM est une version plus récente. Il a été proposé par Soubiran et al [5]. Il est intermédiaire dans la mesure où il comporte l'étape stochastique mais d'une manière tempérée. Toutes les quatre étapes sont nécessaires.

Une fois le mélange identifié, la règle de décision de Bayes peut être appliquée. La numéro de la classe  $k_0(x)$  correspondant à un pixel de l'image de niveau de gris  $x$  est donnée par :

$$k_0(x) = \text{Arg} [ \max \{ \pi_j g(x/\theta_j), 1 \leq j \leq k \} ]$$

Cependant, l'application directe de cette classification se heurte à deux difficultés: l'une concerne la forte dépendance entre pixels voisins, l'autre réside dans l'importance des temps de calculs mis en jeu. Pour pallier ces deux inconvénients, une méthode efficace a été mise en oeuvre: l'échantillonnage Bootstrap [1].

## 2) L'ECHANTILLONNAGE BOOTSTRAP

Le modèle d'échantillonnage aléatoire proposé dans ce paragraphe repose sur la théorie classique d'échantillonnage et notamment sur la technique Bootstrap. Cette technique regroupe l'ensemble des méthodes de rééchantillonnage des données qui recherchent uniquement dans les données observées les informations qui serviront à estimer les paramètres d'une distribution à partir d'une de ses réalisations. Le modèle proposé est alors le suivant :

### a) Modèle

L'expérience qui consiste à tirer aléatoirement suivant une loi uniforme et avec remise  $n$  pixels d'une image de taille  $N$  peut être modélisée par la donnée d'un échantillon  $X = (X_1, X_2, \dots, X_n)$  de  $n$  variables aléatoires réelles. Soit  $x = (x_1, x_2, \dots, x_n)$  une réalisation de cet échantillon, sachant qu'à chaque pixel  $x_i$  correspond un niveau de gris  $j$ , et soit le nombre  $N_j$  de pixels tirés de niveau  $j$  et de probabilité a priori  $\pi_j$  (estimée par les proportions). Les variables aléatoires  $N_j$  suivent alors individuellement une loi binomiale, tandis que le vecteur aléatoire  $(N_1, N_2, \dots, N_k)$  suit une distribution multinomiale. Etant donné que l'on tire aléatoirement un petit échantillon de taille  $n$  (exemple  $n=1000$ ) parmi tous les  $N$  pixels de l'image, il est indispensable de définir un critère de représentativité qui permettra par ailleurs de déterminer la taille optimale  $n_0$  à partir de laquelle l'échantillon pourrait être considéré comme représentatif.

### b) Critère de représentativité

L'échantillon sera représentatif de toute l'image si et seulement si chaque niveau de gris de l'image apparaît au moins une

fois au cours des  $n$  tirages. La probabilité qu'un tel évènement se réalise, avec  $(1 \leq j \leq k)$ , est donnée par:

$$P_n = \text{Prob} [ N_1 > 0, \dots, N_k > 0 ] = \prod_{j=1}^k (1 - e^{-n p_j})$$

Tant que les  $N_j$  peuvent être considérées comme des variables aléatoires de Poisson indépendantes de paramètre  $n p_j$  soumises à la condition  $\sum n_j = n$ ; où  $n_j$  est une réalisation de  $N_j$ . Le critère de représentativité exige donc que  $P_n$  tende vers 1. En pratique, certains niveaux de gris seront probablement rares dans l'image ( $p_j$  faible) et auront ainsi une influence dominante sur la probabilité  $P_n$  qu'ils feront tendre rapidement vers 0; ce qui est contraire à l'effet recherché. Considérons alors la fonction:

$$p(x) = \prod_{j=1}^k (1 - e^{-x p_j}) ; x \in \mathbb{R} ; 0 < p_j < 1$$

$p(x)$  peut alors être vu comme une fonction de vraisemblance dont l'évolution du maximum sera décrite par la fonction

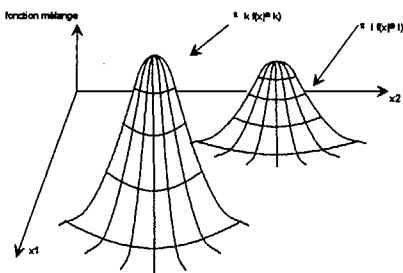
$$B(x) = \sum_{j=1}^k \frac{p_j e^{-x p_j}}{1 - e^{-x p_j}} \text{ et } B(n) = \sum_{j=1}^k \frac{p_j e^{-n p_j}}{1 - e^{-n p_j}}$$

avec  $n p_j \rightarrow \infty$  quand  $B(n) \rightarrow 0$

En pratique, la taille optimale  $n_0$  de l'échantillon est choisie pour  $B(n_0) = 0.05$ . L'utilisation d'un modèle d'échantillonnage aléatoire en segmentation des images présente deux grands avantages liés d'une part au choix d'un échantillon de pixels indépendants qui permettraient d'estimer les paramètres statistiques de l'image dans les meilleures conditions d'indépendance. L'aspect sondage de l'échantillonnage aléatoire permet d'autre part de réduire la redondance d'information liée à la corrélation de voisinage des pixels grâce au choix d'un petit échantillon représentatif, ce qui permet de gagner un facteur  $N/n$  en temps de calcul.

Cependant, la corrélation spatiale entre les pixels est ignorée dans le modèle présenté ci-dessus, qualifié couramment d'approche aveugle. Par opposition, la classification d'un pixel, conditionnellement à son voisinage, définit la segmentation dite contextuelle [3].

### 3) SEGMENTATION CONTEXTUELLE PAR BOOTSTRAP



La segmentation contextuelle se traduit par l'application de la règle de Bayes dans le cas multivarié. Associée à la méthode de Bootstrap, cette approche consiste à tirer aléatoirement, suivant une loi uniforme et

avec remise,  $n$  pixels accompagnés de leurs  $p-1$  voisins. Ceci peut être modélisé par la donnée d'un échantillon  $X = (X_1, X_2, \dots, X_n)$  de  $n$  vecteurs aléatoires de dimension  $p$  sup-

posé suivre un mélange gaussien. Après identification du mélange par la détermination de ses  $K$  composantes, l'application de la règle de Bayes consiste à classer un pixel conditionnellement à son voisinage, ceci en réalisant:

$$K_0 [ x = (x_1, \dots, x_p)^T ] = \text{Arg} [ \max \{ \pi_j g ( x / \theta_j ) \} ]$$

avec  $1 \leq j \leq K$ . Le nombre de classes  $K$  dans le cas contextuel ne peut qu'être supérieur au nombre  $k$  de classes dans le cas aveugle sans qu'il y ait une relation analytique entre eux.

### 4) RESULTATS ET PERSPECTIVES

Nous avons appliqué l'algorithme SAEM à l'image originale de 256 niveaux de gris (fig. 1) tirée d'une séquence enregistrée sur le SLO du service d'ophtalmologie du Pr J.C. Hache du CHR de Lille. Les résultats des différentes segmentations sont présentés dans les fig. 2,3,4,5,6. Par rapport à l'image originale, les structures sont conservées. Le critère de représentativité nous permet de choisir une taille d'échantillon  $n_0 = 1000$ . Dans ces conditions, l'approche Bootstrap présente une nette supériorité tant au niveau du temps de calcul (facteur gain de plus de 260) qu'au niveau de la qualité subjective de la segmentation, ceci grâce à une meilleure justification de l'hypothèse d'indépendance des échantillons qui est à la base des algorithmes de type EM.

Il nous semble important d'explorer la piste contextuelle. Cependant, dans la limite des résultats actuels, l'apport de ce type de segmentation, au niveau de la qualité du rendu, semble négligeable: ceci est dû probablement d'une part au type même des images du fond de l'oeil et à l'information recherchée (vaisseaux sanguins), d'autre part à la taille réduite (1 seul pixel) du contexte. Enfin, la réduction du temps de calculs autorise la segmentation temporelle (contexte temps) de la séquence d'images permettant ainsi son traitement (calcul du débit sanguin) et, éventuellement, son archivage ou sa transmission à faible débit.

### REFERENCES :

[1] BANGA C. and GHORBEL F.  
Optimal Bootstrap sampling for fast image segmentation: Application to retina image. *IEEE-ICASP 1993, Minneapolis USA, Vol. V, pp 638-641*

[2] LEI T. and al.  
Statistical approach to Xrays CT imaging and its application in image analysis. Part II : A new stochastic model-based... *IEEE Trans. on Medical Imaging, vol 11, no1, mars 1992, pp. 62-69.*

[3] MASSON P.  
Etude d'algorithmes de classification contextuelle et application à la segmentation d'images satellites *Thèse de Doctorat, Université de Bretagne Occidentale, no 168, 1991, pp 39-57*

[4] PREUBNER P.R., RICHARD G., DARRELMAN O., WEBER J., KREISSIG :  
Quantitative measurement of retinal blood flow in humans beings by application of digital; image processing methods to television fluorescein angiograms. *GRAEFE'S ARCH. CLIN. EXP. OPHTALMOL., vol 221, 1983, pp. 110-112.*

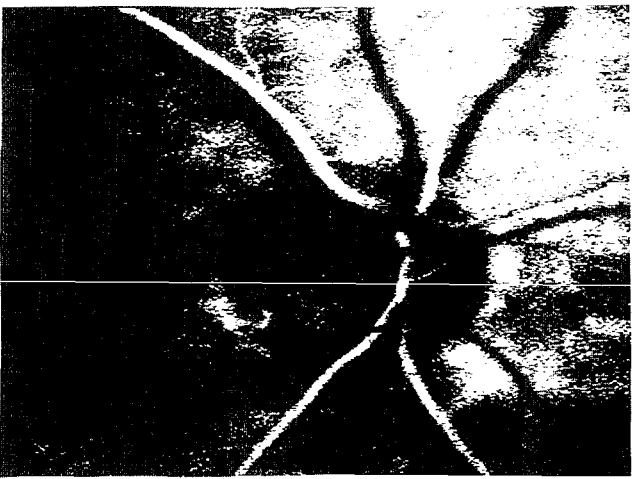
[5] SOUBIRAN C. & al.  
Analyse de mélanges gaussiens pour de petits échantillons: Application à la cinématique stellaire *Revue de Statistique Appliquée, 1991, vol 39, n° 3, pp 17-35*



*Image originale (256 niveaux de gris)*



*Segmentation aveugle classique*



*Segmentation contextuelle classique*



*Segmentation aveugle avec Bootstrap (4 classes)*



*Segmentation contextuelle avec Bootstrap (4 classes)*



*Segmentation contextuelle dans le temps (image suivante)*