FAST UNSUPERVISED STATISTICAL IMAGE SEGMENTATION METHOD

M. Emsalem H. Caillol P. Olivie G. Carnat W. Pieczynski

Groupe Image
Département Systèmes et Réseaux
Institut National des Télécommunications
9, rue Charles Fourier 91011 Evry Cedex France

ABSTRACT

This work deals with the statistical unsupervised image segmentation. We propose a new fast algorithm based on hidden Markov chains. The originality of our approach is situated at two levels. First, the pixels are numbered according to a Peano curve and we show that it improves the efficiency of the classical "line by line" numbering. Second, the parameter estimation phasis is performed by the use of a new general method of estimation in the case of hidden data, so called "iterative conditional estimation". The segmentation phasis is performed by the "maximiser of posteriors marginals", where the posterior marginal distributions are computed be the "backward-forward" algorithm. The efficiency of our method is compared with the efficiency of a "classical" one, where the segmentation is performed by the ICM algorithm and the Markov random hidden fields parameters are estimated, using segmentations based on "current values" of parameters, by the estimator of Derin and Elliot.

INTRODUCTION

Our work treats the nonsupervised segmentation of images. When all the useful parameters are known (supervised case) there is a number of efficient methods of statistical segmentation. Some among these, called global (MAP, MPM, ICM) use a Markovian spatial model, while others (local or contextual) satisfy themselves with a stationarity assumption. When the parameters are unknown, what is usually the case in practical applications, the problem becomes more complex. The efficiency of a given method depends as much on the quality of the estimators as on the robustness of the segmentation method used: Dubes and Jain demonstrate in [8] the degradation in performance of the ICM as the parameters diverge from their values.

Several solutions to this problem have been proposed the last few years. Some algorithms (Chalmond ([3]), Devijver ([6], [7]), Masson and Pieczynski ([15]), Qian and Titterington ([18])) use variations of the EM algorithm ([4]), adapted to the models considered. One alternative technique (Besag ([1]), Lakshmanan and Derin ([11])) consists of a re-estimation of the parameters based on the segmentations obtained with the current parameters. Younes ([20]) proposed a new method based on the notion of stochastique gradient. A new method of estimation in the case of incomplete data, called "iterative conditional estimation" (ICE, [17]), can provide a large number of Bayesian unsupervised segmentation algorithms. Braathen et al. ([2]) applied it in the case of global methods and Marhic et al.in the case of local ones ([12], [13]).The algorithms obtained provide quite satisfactory results. The MAP or MPM based unsupervised techniques mentioned above are rather time consuming and in practical applications one often uses the ICM. As ICE can also be applied to the estimation of all parameters of a hidden Markov chain, which is much quicker than the estimation in the hidden Markov fields, the idea of our method is to range pixels in a sequence and to treat the problem by considering Markov random chains. We choose to range the pixels according to a Peano curve. Doing this we make the "past" and the "future" of the Markov chain considered better suit to the spatial context. In fact, results obtained by our method are better than those obtained by the use of the classic "line by line" curve.

The efficiency of the method we propose is compared with the efficiency of the Besag's one ([1]), where the segmentation is performed by the ICM algorithm and the Markov random hidden fields parameters are estimated, using segmentations based on "current values" of parameters, by the estimator of Derin and Elliot ([5]).

ICE PROCEDURE

Let us suppose that we have an estimator $\hat{\theta} = \hat{\theta}(X,Y)$ of θ defined from (X,Y). Conditional expectation, denoted by E [./Y], is the best approximation to $\hat{\theta}$, as far as the mean square error is concerned, by a function of Y : let us put :

$$\theta^* = E[\hat{\theta} / Y] \qquad (1)$$

θ^* is not an estimator of θ: in fact, the conditional expectation, which can be seen as an expectation of the identity function according to the distribution of X conditioned upon Y, depends on θ. So the use of (1) requires a "current" value of θ.This defines an iterative procedure, which is called ICE (iterative conditional estimation). This procedure runs as follows ([17]): using an initial value θ_0 of θ we put:

$$\theta_{n+1} = E_n[\hat{\theta} / Y] \qquad (2)$$

where E_n denotes the conditional expectation based on the current value θ_n of the parameter.

As we will see in next section it can occur that (2) is not workable but simulations of realisations $X_1^n, X_2^n, ..., X_N^n$ of X according to the posterior distribution (conditioned upon Y) based on θ_n are obtainable, then (2) can be replaced by a "stochastic" ICE :

$$\theta_{n+1} = \frac{1}{N}\left[\hat{\theta}(X_1^n, Y) + ... + \hat{\theta}(X_N^n, Y)\right]$$

which converges, according to the law of large numbers, to θ_{n+1} defined by (2).

HIDDEN MARKOV CHAINS

We consider the following modelling: $X = (X_1,...,X_n)$ is a Markov chain and the random variables $Y_1,...,Y_n$ are real, Gaussian and independent conditionally to each realization of X. Every X_i takes its values in a finite space $\Omega = \{\omega_1,..., \omega_k\}$. The distribution of X is given by $q_i = P[X_1 = \omega_i]$ and $q_{ij} = P[X_m = \omega_i, X_{m+1} = \omega_j]$. (q_{ij}) defines the transition matrix (a_{ij}) by :

$$a_{ij} = \frac{q_{ij}}{\sum_j q_{ij}}$$

The likelihood of X at the point $x = (\omega_{i_1}, \omega_{i_2},..., \omega_{i_n})$ is

$$f(x) = q_{i_1} a_{i_1 i_2} a_{i_2 i_3} ... a_{i_{n-1} i_n}$$

and, if we denote by f_i the likelihood of the distribution of each Y_j conditioned by $X_j = \omega_i$, the likelihood of the distribution of (X,Y) at the point (x,y), with $y = (y_1,...,y_n)$, is :

$$h(x,y) = q_{i_1} f_{i_1}(y_1) a_{i_1 i_2} f_{i_2}(y_2) .... a_{i_{n-1} i_n} f_{i_n}(y_n)$$

so, the posterior likelihood of X is

$$h^y(x) = \frac{h(x,y)}{\sum_x h(x,y)}$$

which can be written :

$$h^y(x) = q_{i_1}^{y_1} a_{i_1 i_2}^{y_2} .... a_{i_{n-1} i_n}^{y_n}$$

with

$$a_{i_m i_{m+1}}^{y_{m+1}} = \frac{a_{i_m i_{m+1}} f_{i_{m+1}}(y_{m+1})}{\sum_j a_{i_m j} f_j(y_{m+1})}$$

and

$$q_{i_1}^{y_1} = \frac{q_{i_1} f_{i_1}(y_1)}{\sum_j q_j f_j(y_1)}$$

The result is that the posterior distribution of X is that of a non-stationary Markov chain, which makes its simulation quite easy. Thus we can use the SICE procedure whenever we have at our disposal an estimator $\hat{\theta}=\hat{\theta}(X,Y)$ of the parameter $\theta = (Q, \beta_1,...\beta_k)$, where $Q = (q_{ij})$ defines the distribution of X and $\beta = (\beta_1,...,\beta_k)$ defines the distributions of Y conditional to X (every β_i defines the distribution of each Y_j conditional to $X_j = \omega_i$). For instance, if (Y_j) are Gaussian and take their values in R^m, β_i is given by the mean vector and covariance matrix. We can often find such an estimator because Q can be estimated from X by the frequencies and we dispose, in general, of estimators (β_i) of (β_i) from (X,Y) (which are in the Gaussian case empirical mean vectors and covariance matrices.

FAST UNSUPERVISED ALGORITHM

The fast algorithm we propose runs as follows:

1° Number the pixels according to a Peano curve
2° Suppose the stochastic process obtained is a hidden Markov chain
3° Estimate all parameters by the stochastic ICE
4° Calculate the posterior marginal distributions using the "backward -forward" algorithm ([6])
5° Perform the segmentation using the MPM

Thus our method can be seen as an approximation of the real, based on the modelling by hidden Markov fields, MPM method. The approximation consists on substitution of the spatial context by the "past" and "future" of a stochastic process. The use of a Peano curve makes this approximation less rough. The saving of the computing time is mainly realised in the parameter estimation phasis: when considering Markov chains the simulation of realisations according to the posterior distribution, used in the ICE procedure, is direct and one is not obliged to call on iterative methods, as it is the case when using Markov fields. When the sequence of the estimated parameters becomes steady we stop the parameter estimation step and apply the "backward-forward" algorithm in order to calculate the marginal posterior distributions, and classify each pixel according to the MPM principle. Let us note that the possibility of the computing of the posteriors by the "backward-forward" is another source of time saving, in fact when using the MPM based on Markov random fields modelling these posteriors can not be computed and have to be estimated using simulations of the class field according to the posterior distribution.

SOME NUMERICAL RESULTS

Let us consider two images Im.1 (Fig.1), Im.2 (Fig.2). They are corrupted with the white Gaussian noise of variance 1 and means 0 and 2 respectively, which gives images represented in Fig.3 and Fig.4 respectively. Let us denote by A1 the Besag's unsupervised algorithm, A2 our ICE based fast method using the "line by line" curve and A3 our method using the Peano curve. The results, expressed in the per cent of wrong classified pixels, of the corresponding unsupervised segmentations are given in the Tab.1 below. The segmented images by A2 and A3 are given by Fig.5, Fig.6, Fig.7, Fig.8.

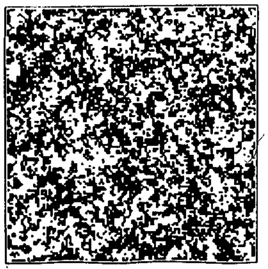

Fig.1 Im.1

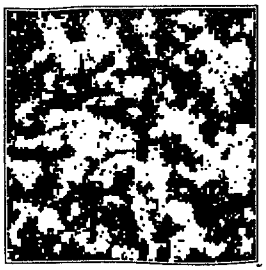

Fig.2 Im.2

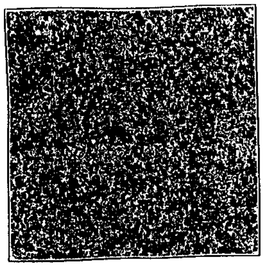

Fig.3 Im.1+noise

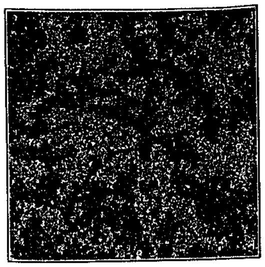

Fig.4 Im.2+noise

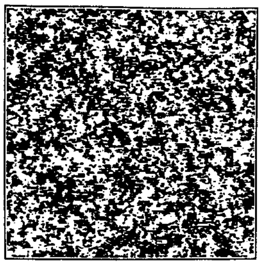

Fig.5 A2(Im.1)

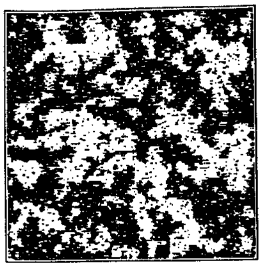

Fig.6 A2(Im.2)

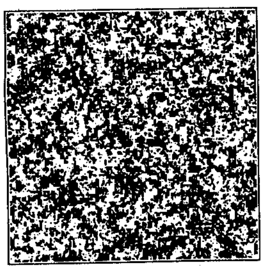

Fig.7 A3(Im.1)

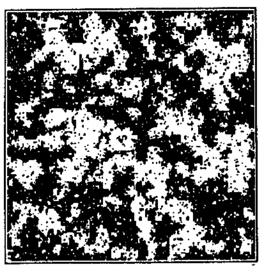

Fig.8 A3(Im.2)

	A1	A2	A3
Im.1	15.67	15.24	11.48
Im.2	7.20	9.10	7.40

Tab.1

CONCLUDING REMARKS

The efficiency of our method seems to be equivalent to the efficiency of the Besag's algorithm in the case of a relatively homogeneous image and is much better in the case of a non homogeneous one. In both cases the use of the Peano curve implies a real improvement with respect to the use of the classic "line by line" one.

REFERENCES :

1° Besag J. - 1986 - On the statistical analysis of dirty pictures - Journal of the Royal Statistical Society - Series B, 48 pp. 259-302.

2° Braathen B.- Marhic N.- Masson P.- Pieczynski W. - 1991 - Sur une nouvelle approche de segmentation bayesienne non supervisée d'images. Proceedings of GRETSI, Juan-Les-Pins, France.

3° Chalmond B. - 1989 - An iterative Gibbsian technique for reconstruction of m-ary images - Pattern Recognition - Vol 22 - n° 6 - pp 747-761.

4° Dempster M.M. - Laird N.M. - Rubin D.B. - 1977 - Maximum likelihood from incomplete data via the EM algorithm - Journal of the Royal Statistical Society - Series B - 39 pp 1-38.

5° Derin H. - Elliot H. - 1987 - Modelling and segmentation of noisy and textured images using Gibbs random fields - IEEE Transactions on PAMI-9 - pp 39-55.

6° Devijver P.A. - Dekesel M. - 1988 - Champs aléatoires de Pickard et modélisation d'images digitales - Traitement du Signal - Vol. 5 - n° 5.

7° Devijver P.A. - Real time modelling of image sequence. Submitted to IEEE Transactions on PAMI.

8° Dubes R.C. - Jain A.K. - 1989 - Random field models in image analysis - Journal of Applied Statistics - Vol. 6 - n° 2.

9° Geman S. - Geman D. - 1984 - Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, IEEE Transactions on PAMI, 6 - pp 721-741.

10° Kelly P.A. - Derin H. - Hartt K.D. - 1988 - Adaptive segmentation of speckled images using a hierarchical random field model - IEEE Transactions on ASSP - Vol. 36 - n° 10 - pp 1628-1641.

11° Lakshmanan S. - Derin H. - 1989 - Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing" - Transactions on PAMI- Vol. 11 - n° 8.

12° Marhic N.- Masson P.- Pieczynski W.- Melange de lois et segmentation non supervisee des donnees SPOT. To appear in Statistique et Analyse des Donnees.

13° Marhic N.- Pieczynski W. -1991- Estimation of mixture and unsupervised segmentation of images. Proceedings of IGARSS, Helsinki, Finland.

14° Marroquin J. - Mitter S. Poggio T. - 1987 - Probabilistic solution of ill-posed problems in computational vision - Journal of the American Statistical Association - 82 - pp. 76-89 - 1987.

15° Masson P. - Pieczynski W. - 1990 - Segmentation of SPOT images by contextual SEM - Proceedings of EUSIPCO - Barcelona.

16° Pieczynski W. - 1989 - Estimation of context in random fields - Journal of Applied Statistics - Vol 6 - n°2.

17° Pieczynski W.-1991- Missing data and iterative conditional estimation. Technical report n° 134, LSTA, Universite PARIS 6, France.

18° Qian W. - Titterington D.M. - 1989 - On the use of Gibbs Markov chain models in the analysis of images based on second-order pairwise interactive distributions. Journal of Applied Statistics - Vol. 16 - n°2.

19° Tilton J. - Vardeman S. - Swain P. - 1982 - Estimation of context for statistical classification of multispectral image data, IEEE Transactions on GRS - GE-20 - pp 445-452.

20° Younes L. - 1989 - Parametric inference for imperfectly observed Gibbsian fields - Probability Theory and Related Fields 82, 625-645.