LOCAL PARAMETER ESTIMATION AND UNSUPERVISED SEGMENTATION OF SAR IMAGES

H.-Ch. Quelle*, J.-M. Boucher*, W. Pieczynski**

* Département Mathématiques et Systèmes de Communication
Ecole Nationale Supérieure des Télécommunications de Bretagne
** Département Systèmes et Réseaux
Institut National des Télécommunications

1. Abstract

Our work deals with the unsupervised statistical segmentation of SAR images. However the method here developped is a general parameter estimation technique and can be used for most types of images. We adopt a contextual method in which each pixel is classified from the measurements taken in its neighborhood. In this approach the previous statistical problem is the estimation of components of a distribution mixture. We showed in some previous studies that the SEM is well adapted to the problem in this frame, when stationary random fields are considered. In this paper we present a new distribution mixture estimator in which priors can depend on the position of the considered pixel. This makes it valid in the non-stationary case. We describe some situations, based on synthetic images sampled by stationary or non stationary random fields, in which the contextual method based on parameters estimated by our algorithm is more efficient than the same method based on parameters estimated by the SEM algorithm.

2. Introduction

When considering the statistical segmentation of images authors generally suppose the existence of two random fields : the field of "classes" $X = \{X_s : s \in S\}$, and the field of "measurements" $Y = \{Y_s : s \in S\}$. Then at each pixel $s \in S$, the random variable X_s takes its value in a finite set $\Omega = \{\omega_1, \ldots, \omega_K\}$ of classes and Y_s in $\mathbb{R}$. So the segmentation is in fact an estimation of an "ignored" realization of X from an "observed" realization of Y. There are two families of bayesian methods :

2.1. global methods

A global method uses all available spatial information : each X_s is classified from the whole observation $Y = y$. By choosing two different loss functions we get the two global methods : MAP and MPM. Neither the solution of MAP nor the solution of MPM can be computed directly. However their solutions can be approached by the simulated annealing [4] and Marroquin et al's algorithm [8]. Both require the two following hypothesis : the field X is markovian and the random variables Y_s are independent conditionally to every realization of X. When the used model verifies these conditions and when all parameters are known several authors showed the great efficiency of global methods. When the needed parameters are not known (unsupervised case) the problem becomes much more difficult. Methods based on the EM algorithm demand strong hypothesis : the same noise variance for all classes and discrete or Gaussian noise.

2.2. local methods

A local, or contextual, method uses the information contained in the neighborhood of each pixel. Its disadvantage is to loose a great part of information (the size of the neighborhood has to be small, since the methods are rather time consuming). Its advantage is the ability to take the noise correlation into account. This approach does not need the modelling by Markov random fields : if $\nu \in S$ designates the context used we have only to know the distribution of (X_ν, Y_ν), the restriction of (X, Y) to ν. In the non-supervised case we have then to choose a sequence of contexts $\nu_1, \nu_2, \ldots, \nu_n$ in S and estimate the parameters defining this distribution from $Y_{\nu_1}, Y_{\nu_2}, \ldots, Y_{\nu_n}$.

The comparison of the efficiency of these two families is difficult in the general case. Local methods, in appearance more rudimentary, can turn out to be more efficient in some situations [7].

In this paper we present a new unsupervised image segmentation method and show its superiority over the SEM in the case of unsupervised SAR image segmentation.

Its organization is as following.

In the next section we briefly recall the hierarchical image model (HIM) from Kelly and Derin based on Markov random fields, generally agreed to model SAR images.

The third section is devoted to the description of our method and the principle of SEM is carried out.

The forth section contains results of experiments and some comments.

Concluding remarks constitute the fifth and last section.

3. Modeling SAR images

This chapter explains the image model HIM from Kelly and Derin. We use it for the modeling of our test scenes and generally for the Bayesian approach of image segmentation. The main idea of the HIM are : the model consists of two independent random process, and the observed image is the superposition of these two processes. This model can easly be adapted to most image producing systemes by changing the statistics of the texture generating process.

We resume very briefly the chapter "Model for speckled images" from [5]. The hierarchical image model consists of two random fields. One governs the grouping of pixels, called region process. The other consists of K random fields which represent the speckled appearance of the K types of nature, the classes. We have chosen Kelly and Derin's image model for two reasons.

Firstly, this model points clearly out the idea that the "observed speckle" is not only a worthless disturbance, but carries information about the nature of ground.

Secondly, the hierarchical character fits well to Bayesian decision theory and this model allows quite easily the generation of a great number of images with real SAR image statistics.

3.1. Region process

The region process is responsible for the distribution of pixels within the different classes.

The field x is said to be markovian, with :

$$\begin{aligned} &P[\,x_s = \omega_s \mid x_t = \omega_t, (s) \neq (t)] \\ =&P[\,x_s = \omega_s \mid x_t = \omega_t, (s) \in \nu_t] \end{aligned}$$

where ν_s designs a neighborhood of s.

In fact it can be a very simple type of a MF, called multilogistic level (MLL) field [3]. In this model a parameter exists for each clique type : δ_k for the K differently colored singletons, γ for the four types of clique-pairs. For the sake of simplicity all cliques consisting of more than two pixels are ignored. Under a positivity condition the field x has a Gibbs distribution :

$$P[\,x = \omega\,] = \frac{1}{Z} \exp\left\{ -\sum_{c \in C} V_c(\omega) \right\}$$

Where the potential function $V_c(\omega)$ is defined as :
for c a clique of pairs :

$$V_c(\omega) = \begin{cases} \gamma \text{ if all } \omega_s \text{ are equal} & \text{for } (s) \in c \\ -\gamma \text{ otherwise} \end{cases}$$

for c a clique of one singleton :

$$V_c(\omega) = \delta_k \text{ if } x_s = \omega_k \text{ for } (s) \in c$$

3.2. Speckle process

The speckle process generates for each class k a random field $W^{(k)}$ or $V^{(k)}$ which modelize the textures of the different classes. The speckle consists, in analogy with the physical model, of a field of complex, gaussian, zero-mean random variables $Z = \{Z_s : s \in S\}$. At each site s, the intensity w and the phase φ are defined by :

$$w_s = |z_s|^2, \text{ and } \varphi_s = \arctan \frac{\text{Im}\,[z_s]}{\text{Re}\,[z_s]}$$

The joint probability $P[\,\underline{w}\,]$ may be expressed explicitly with :

$$\underline{W} = [\, W_{u,v}, W_{u-1,v}, W_{u+1,v}, W_{u,v-1}, W_{u,v+1} \,] \qquad (u,v) \in S$$

For multi-look images with

$$V = \frac{1}{L} \sum_{1}^{L} W_l$$

the joint density can be expressed by a $(L-1)$ fold convolution of either $P[\,w\,]$ or $P[\,\underline{w}\,]$. Deriving $P[\,\underline{v}\,]$ analytically fails at the evaluation of the convolution integral.

3.3. Hierarchical model

At each pixel s the value of $X_s = \omega_k$ determines the marginal distribution of the speckle process. The color of the pixel s from speckle process $W_s^{(k)}$ gives the final value y_s.

$$Y_s = W_s^{(k)} \text{ or } V_s^{(k)} \quad \text{if} \quad x_s = \omega_k$$

4. Local SEM distribution mixture estimation

The aim of this section is to expose our method which we will note LSEM (local SEM [1],[2]). In order to simplify things we will consider a simple case : the neighborhood is limited to one pixel (ν_s consists of (s) and its neighbor) and a image is binary (two classes, $m = 2$). The generalization is possible, however, the number of parameters to estimate increases quickly. To be more precise, if K is the number of classes and m the cardinal of ν we have a mixture of K^m distributions. So, for $K \geq 3$ it is difficult to consider more than one neighbor.

We first recall the SEM algorithm and expose then the modifications leading to the LSEM. Let ν be of the shape above defined and $\nu_1, \nu_2, \ldots, \nu_n$ a sequence of subsets of this shape in S. This sequence does not necessarily recover S. We will denote by $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ the restrictions of X and Y to $\nu_1, \nu_2, \ldots, \nu_n$ respectively. Thus we have to estimate the parameters defining the distribution of (X_i, Y_i), which is independent from i in the stationary case. In our case this distribution is given by the priors $p_{ij} = P[\,X_\nu = (\omega_i, \omega_j)\,]$ and f_{ij} distribution densities of Y_ν conditional to $X_\nu = (\omega_i, \omega_j)$ $(1 \leq i,j \leq 2)$. According to the hierarchical model the densities f_{ij} are defined by the means $m_1 = E[\,Y_s \mid X_s : x_s = \omega_1\,]$, $m_2 = E[\,Y_s \mid X_s : x_s = \omega_2\,]$ and covariance matrices $C_i = \text{cov}[\,(Y_s, Y_t) \mid (X_s, X_t) = (\omega_i, \omega_i)\,]$ $(1 \leq i \leq 2)$, where s and t are horizontal neighbors. In order to simplify, let us put $q_1 = p_{11}, q_2 = p_{12}, q_3 = p_{21}, q_4 = p_{22}$.

Finally the problem consists of estimating the parameter $\theta = (q_1, q_2, q_3, q_4, m_1, m_2, C_1, C_2)$ from a realization of $X_1, X_2, \ldots, X_n$. As it will appear in the sequence, it is convenient to separate the priors $\alpha = (q_1, q_2, q_3, q_4)$ and the parameter $\beta = (m_1, m_2, C_1, C_2)$ defining the conditional distributions.

4.1. SEM Algorithm

The idea of this algorithm is to use the "artificial" (sampled according to the a posteriori distribution) realizations of $X_1, X_2, \ldots, X_n$. This is an iterativ method which runs as follow.

- initialization : consider a first realization $x^0 = (x_1^0, \ldots, x_n^0)$ of $X_1, \ldots, X_n$ somehow obtained. In the image segmentation problem it can be the restriction to $\nu_1, \nu_2, \ldots, \nu_n$ of a segmentation obtained from a classic method (for instance based on the histogram). In absence of any information each X_i^0 is sampled according to the uniform distribution $L_0 = (0.25, 0.25, 0.25, 0.25)$. Consider $Q_1^0, Q_2^0, Q_3^0, Q_4^0$ a partition of the set of pixels such that for every $j \in Q_i^0$ the realization of X_j^0 corresponds to q_i. Suppose that x^0 is the actual realization of X. The estimation of β is than quite easy : m_1 can be estimated by the empirical mean from the first component of the Y_i such that $i \in Q_1^0 \cup Q_2^0$ and the second component of Y_j such that $j \in Q_3^0$, the same for m_2 by using $Q_3^0 \cup Q_4^0$ and Q_2^0. We estimate C_1 by empirical covariance from Y_i such that $i \in Q_1^0$, the same for C_2 by using Q_4^0. This gives us a first value $\beta^0 = (m_1^0, m_2^0, C_1^0, C_2^0)$.
- For $\tau > 1$ $\theta^{\tau+1}$ is obtained from θ^τ and $(Y_1, \ldots, Y_n) = (y_1, \ldots, y_n)$ as follows :
  - (i) compute, for each $1 \leq i \leq n$, the a posteriori distribution based on $Y_i = y_i$ and θ^τ.
  - (ii) sample, for each $1 \leq i \leq n$, a realization x_i^τ in Ω^m according to $p_{y_i}^\tau$.
  - (iii) estimate $\theta^{\tau+1}$ by using $Q_1^\tau, Q_2^\tau, Q_3^\tau, Q_4^\tau$ as in the initialization.

The SEM was already used as a parameter estimation step in contextual segmentation in the case of optical (SPOT) data [6] and gives satisfying results. This could be due to the gaussian conditional distributions used. In the SAR case they are not Gaussian and the programmation becomes more difficult [9].

4.2. LSEM algorithm

Real images contain at least 256 $\times$256 pixels and can therefore rarly be considered as a realization of a stationary random field. So ordinarly the image is cutted into little stationary windows. This implies the problem of their size : if the size is not importent enough, the estimated parameters are not correct, if the size is to big, the stationarity can not longer be ensured.

The novelty of our approach is to suppose that the non-stationarity of Y is only due to the non-stationarity of X and the K conditional distributions $f_1, f_2, \ldots, f_K$ de Y given $X : X_s = \omega_k \ \forall \ s \in S$ are stationary, with other words that the parameter α depends on the position of s in S but β rests independent of s. As we will see below, this allows us to estimate $(L_s)_{s \in S}$ on a small window and β in the whole scene S. This modelling is much more general than the stationary hierarchical model ; in fact, the "conditional stationarity" of Y seems natural. These distributions model the "natural variability" of the ground and different noises inside the classes which, in principle, do not depend on the position of pixel s in S.

The LSEM runs as follows :

- take the same initialization like for the SEM, with $\alpha = (\alpha_i)_{1 \leq i \leq n}$ depends on s. Let us denote by W_s a window containing the couple of pixels at s.
- For $\tau \geq 1$ $\theta^{\tau+1} = (\theta_i^{\tau+1})_{1 \leq i \leq n}$ with $\alpha^{\tau+1} = (\alpha_i^{\tau+1})_{1 \leq i \leq n}$ and $\beta^{\tau+1}$ is obtained from $\theta^\tau = (\theta_i^\tau)_{1 \leq i \leq n}$ and $(Y_1, \ldots, Y_n) = (y_1, \ldots, y_n)$ by :
  - (i) compute, for each $1 \leq i \leq n$, the a posteriori distribution $\hat{p}_{y_i}^\tau$ on Ω based on $Y_i = y_i$ and θ_i^τ.
  - (ii) sample, for each $1 \leq i \leq n$, a realization x_i^τ in Ω according to $\hat{p}_{y_i}^\tau$.
  - (iii) consider, as in the case of the SEM, the partition $Q^\tau = (Q_1^\tau, Q_2^\tau, Q_3^\tau, Q_4^\tau)$ and put for each $1 \leq i \leq n$, $Q_i^\tau = (Q_1^\tau \cap W_i, Q_2^\tau \cap W_i, Q_4^\tau \cap W_i)$.
- Estimate $\beta^{\tau+1}$ from Q^τ and each $\alpha_i^{\tau+1}$ from Q_i^τ in the same way then in the SEM case. This gives the next value $\theta^{\tau+1}$ of θ.

5. Results

We present some estimation and segmentations results. We processed two images with four different methods. The number of classes is equal in all images, $K = 2$. The theoretical error of blind segmentation is about 25%. Image I.1 contains a special region process which allows us to test the impact of the stationarity hypotheses. Image I.2 was created by a non-stationary version of the hierarchical image model, where the parameters δ_k and γ_k of the region process changes abruptly from a low level (left, upper coin) to a high level (right, downer coin). The notations used in this chapter are : When the neighborhood of s in the speckle process Y is reduced to $\nu_s = s$, we call it a "blind" method. When the neighborhood of s in the speckle process Y consists of $\nu_s = s$ and its neighbors, we call it a "contextual" method. When the a priori probabilities of the classes are estimated in the whole scene S, we call it a "global" estimation. When the a priori probabilities of the classes are estimated locally in W_s, a little (3×3) window centered at s, we call it a local estimation. We segmented the two images I.1 and I.2 with the following methods :

1) blind-global, 2) blind-local, 3) contextual-global, 4) contextual-local.

Let us notice that we use the terms "local" and "global" only to design the parameter estimation but the segmentation is still made by a "local" method.

In Fig. 2 and Fig. 3 image I.1 was segmented by the above defined four methods. In this particular image local and global estimated priors are completly different. The result show clearly the influence of the locally adapted estimation of priors.

The images in Fig. 2/3 are the repartitions Q_1^{20}, Q_2^{20} of the region process after 20 iterations. The final segmentation results are slighty different, because not the statistic picking of the SEM or the LSEM, but the Bayesian rule is used for discriminating the classes.

Fig. 4 shows the region process X and the synthesized image Y of I.2. We tested the influence of the size of the local a priori probabilities estimation window. Fig. 5 shows left the result of a segmentation using a local 3×3 window for estimating the priors and right the result of a segmentation using a local 5×5 window.

We notice here another effect of local parameter estimation which we call the "cleaning effect". Since the probabilities of classes are estimated in a little window, local partition is very determining. The little this window is, the greater is the impact of the local partition.

The results show clearly the superiority of local a prior parameter estimation methods in non-stationary scenes. We notice the influence of contextual segmentation, which is little here due to a very tiny correlation of the noise.

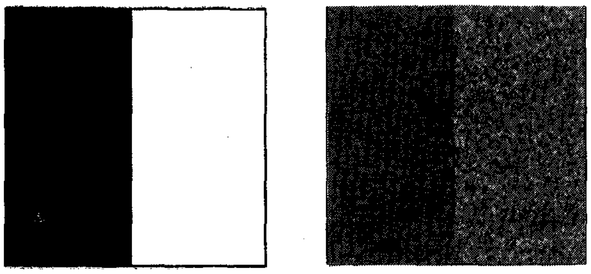

Fig. 1 : *I.1. left : region process X, and right : Y*

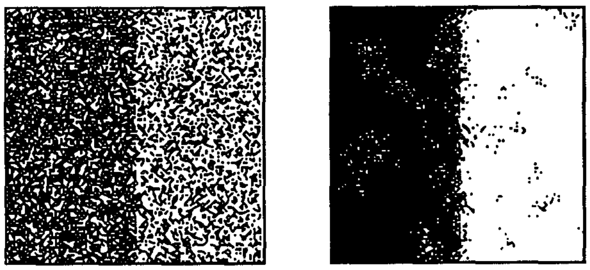

Fig. 2 : *I.1 segmented ; left : blind-global, right : blind-local*

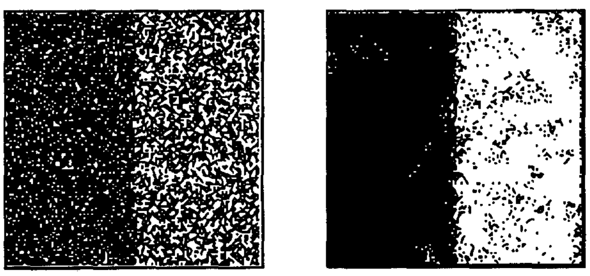

Fig. 3 : *I.1 segmented ; left : contextual-global, right : contextual-local*

Image	blind-global	blind-local	contextual-global	contextual-local
I.1 (stripes)	36,50%	07,07%	31,75%	06,90%

Tab. 1 : Error rates of the classification phase Q_1^{20}, Q_2^{20}

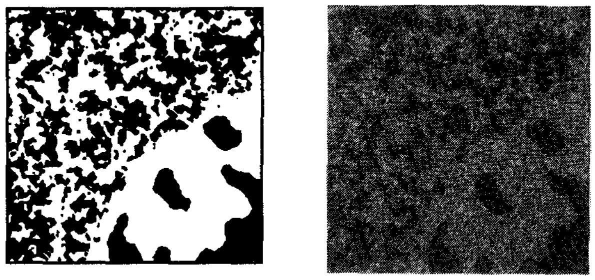

Fig. 4 : *I.2 : left region process X, right Y*

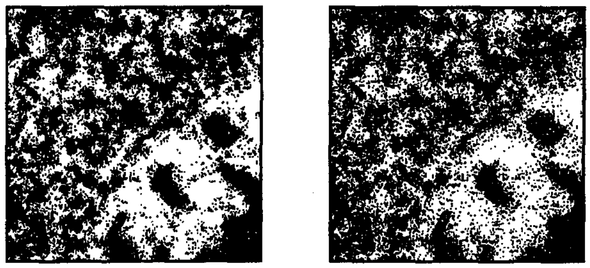

Fig. 5 : *I.2 blind-local, left : 3×3, right : 5×5*

Image	blind-local (3×3)	contextual-local (5×5)	(blind-global, not shown)
I.2 (gibbs)	16,15%	19,80%	34,56%

Tab. 2 : Error rates of the classification phase Q_1^{20}, Q_2^{20}

6. Conclusions

We exposed a new "local" parameter estimating method and showed that in some situations the blind and the contextual segmentation based on locally estimated parameters is more efficient than the segmentation based on globally estimated parameters. Furthermore, our method remaind valid in a more general frame - in fact, the stationarity of the class field X is not required.

7. References

[1] G. Celeux, J. Diebolt, "L'algorithme SEM : Un algorithme d'apprentissage probabiliste pour la reconnaisance de mélange de dénsités, "*Revue de Statisiques Appliquées*, **vol.34**, no.2, 1986.

[2] G. Celeux, J. Diebolt, "Comportement asymptotique d'un algorithme d'apprentissage probabiliste pour les mélanges de lois de probabilités, "*Rapport de la Recherche*, , no.563, 1986.

[3] H. Elliot, H. Derin, R. Christi and D. Geman, "Application to the Gibbs distribution to image segmentation," in *Proceedings of the Int. Conf. Acoust. Speech Signal Processing. ICASSP*, 1984, pp.32.5.1-32.5.4.

[4] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, "*IEEE Transactions on Pattern and Machine Intelligence*, **vol.6**, pp.721-741, 1984.

[5] P.A. Kelly, H. Derin, K.D. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model, "*IEEE Transactions on Acoustics, Speech, and Signal Processing*, **vol.36**, pp.1628-1641, 1988.

[6] P. Masson, W. Pieczynski, "Segmentation of SPOT images by contextual SEM," in *Proceedings of EUSIPCO*, Sept.1990, ,.

[7] Nicole Marhic, W. Pieczynski, "Estimation of mixture and unsupervised segmentation of images," in *Proceedings of IGARSS'91*, June 1991, pp.1083-1086.

[8] J. Marroquin, S. Mitter and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision, "*Journal of the American Statistical Association*, **82**, pp.76-89, 1987.

[9] H.-Ch. Quelle, J.-M. Boucher and W. Pieczynski, "Unsupervised Bayesian classification of SAR-images," in *Proceedings of IGARSS'91*, June 1991, pp.2177-2180.