SPIEs International Symposium on Remote Sensing, September 22-27, Crete, Greece, 2002

SAR Image Segmentation using Generalized Pairwise Markov Chains

Stéphane Derrode[a] and Wojciech Pieczynski[b]

[a]ENSPM, GSM, Institut Fresnel (UMR 6133), DU de Saint Jérôme, F13013 Marseille
[b]INT, Département CITI, 9 rue Charles Fourier, F91011 Evry

ABSTRACT

The efficiency of Markov models in the context of SAR image segmentation mainly relies on their spatial regularity constraint. However, a pixel may have a rather different visual aspect when it is located near a boundary or inside a large set of pixels of the same class. According to the classical hypothesis in Hidden Markov Chain (HMC) models, this fact can not be taken into consideration. This is the very reason of the recent Pairwise Markov Chains (PMC) model which relies on the hypothesis that the pairwise process (X, Y) is Markovian and stationary, but not necessarily X. The main interest of the PMC model in SAR image segmentation is to not assume that the speckle is spatially uncorrelated. Hence, it is possible to take into account the difference between two successive pixels that belong to the same region or that overlap a boundary. Both PMC and HMC parameters are learnt from a variant of the Iterative Conditional Estimation method. This allows to apply the Bayesian Maximum Posterior Marginal criterion for the restoration of X in an unsupervised manner. We will compare the PMC model with respect to the HMC one for the unsupervised segmentation of SAR images, for both Gaussian distributions and Pearson system of distributions.

Keywords: Unsupervised classification, Bayesian restoration, Markov chain, pairwise Markov chain, iterative conditional estimation, generalized mixture estimation, Pearson system, SAR image segmentation.

1. INTRODUCTION

In the only image processing area, the field of applications of Hidden Markov Chains (HMC) is extremely vast. HMC can be used in image segmentation[1], hand-written word recognition[2], vehicle detection[3], or even gesture recognition[4]. Multisensor and multiresolution images can still be segmented using HMC[5,6]. The success of HMC models is mainly due to the fact that when the unobservable process X can be modeled by a finite Markov chain and when the noise is not too complex, then X can be recovered from the observed process Y using different Bayesian classification techniques like Maximum A Posteriori (MAP) or Maximal Posterior Mode (MPM)[7,8].

The efficiency of Markov models in the context of Synthetic Aperture Radar (SAR) image segmentation mainly relies on their spatial regularity constraint. This is justified by the fact that for many natural scenes, neighboring pixels are more likely to belong to the same class than pixels that are farther away from each other. However, a pixel may have a rather different visual aspect when it is located near a boundary or inside a large set of pixels of the same class. According to the classical hypothesis in Markov models, this fact can not be taken into consideration. This is the very reason of the recent Pairwise Markov Field (PMF[9]) and Pairwise Markov Chain (PMC[10,11]) models in which the pairwise process (X, Y) is supposed to be Markovian. In this paper, we deal with PMC. Such model is strictly more general than HMC since the (single) process X is not necessarily a Markov one. In the SAR image segmentation context, the main interest of the PMC model is to not assume that the speckle is spatially uncorrelated. Hence, unlike the HMC model, it is therefore possible to take into account the difference between two consecutive pixels that belong to the same region and two consecutive pixels that overlap a boundary between two regions.

Further author information:
S.D.: E-mail: stephane.derrode@enspm.u-3mrs.fr, Telephone: +33 4 91 28 28 49
W.P.: E-mail: wojciech.pieczynski@int-evry.fr, Telephone: +33 1 60 76 44 25

It has been shown that, similar to HMC, the PMC model allows one to perform Bayesian MAP or MPM restorations[10]. Since we are concerned with unsupervised classification, the statistical properties of the different classes are unknown and must first be estimated. In this work, all of the parameters are learnt from a variant of the general Iterative Conditional Estimation (ICE) method[12]. ICE is an alternative to the Expectation-Minimization (EM) procedure and has been successfully applied in several unsupervised image segmentation applications: sonar[13] and medical[14] images, spatio-temporal video[1], radar[15] and multiresolution[16, 17] images, among others. The only image dependant parameters that must be entered by the user is the number of classes. All other parameters are computed automatically.

The main issue of this work is to test and compare the PMC model with respect to the HMC one for the unsupervised segmentation of SAR images. This article is organized as follows: In the next section, the PMC model and the ICE principle are recalled. Sections 3 and 4 are devoted to the parameters estimation problem. In particular, we specify how the original ICE based method works in the Gaussian case and how it can be extended to the generalized mixture estimation case. Sections 5 and 6 compare the segmentation results obtained from the HMC and PMC based restoration algorithms on two noisy simulated images and a real JERS1 image of rice plantation in Indonesia. Finally, conclusions and further work are drawn in section 7.

2. PAIRWISE MARKOV CHAIN

Since we are concerned with the restoration of images, one have first to transform the bi-dimensional data into a one-dimensional chain. A solution consists in traversing the image line by line or column by column. Nevertheless, only horizontal or vertical neighbourhoods are taken into account. Another alternative is to use a Hilbert-Peano scan which alternates horizontal and vertical transition, as illustrated in figure 1. Generalized Hilbert-Peano scan[18] can be applied to image whose height and width are even.

2.1. Pairwise Markov Chain distribution

Let $X = (X_1, \ldots, X_N)$, and $Y = (Y_1, \ldots, Y_N)$ be two sequences of random variables, corresponding respectively to the unknown class image and the observed image, according to the Hilbert-Peano scan. Each X_n, $1 \leq n \leq N$ takes its values in a finite set $\Omega = \{1, \ldots, K\}$ of classes and each Y_n takes its values in the set of real numbers $\mathbb{R}$. Considering $Z_n = (X_n, Y_n)$, the process $Z = (Z_1, \ldots, Z_N)$ is called a "pairwise" process associated to X and Y. Realizations of such processes will be denoted by lowercase letters. To simplify notations, we will write $p(x_n)$ for $p(X_n = x_n)$ and we will denote different distributions by the same letter p, except for distributions on $\mathbb{R}$ and $\mathbb{R}^2$ where f will be used.

The process Z is then called a "pairwise Markov chain" associated with X and Y if its distribution can be expressed as $p(z) = p(z_1)\, p\,(z_2\,|z_1\,)\, \ldots\, p\,(z_N\,|z_{N-1}\,)$. In this paper, we only consider the "stationary PMC" in which $p(z_n, z_{n+1})$ does not depend on n. The process Z is then a Markov chain whose distribution is defined by

$$p(z_1, z_2) = p(i, j)\; f_{i,j}(y_1, y_2), \tag{1}$$

where $p(i, j)$ is a probability on Ω^2, and $f_{i,j}(y_1, y_2) = p\,(y_1, y_2\,|i, j\,)$ is a distribution density on $\mathbb{R}^2$. Hence, $p(z_1, z_2)$ is a probability distribution on $\Omega^2 \times \mathbb{R}^2$. The distribution of the Markov chain Z will equivalently be determined by the initial probabilities $p(z_1)$, using (1):

$$p(z_1) = \sum_{j=1}^{K} p(i, j) \int_{\mathbb{R}} f_{i,j}(y_1, y_2)\, dy_2 = \sum_{j=1}^{K} p(i, j)\; f_{i,j}(y_1), \tag{2}$$

and the transition matrix $p(z_2|z_1)$, using (1) and (2):

$$p(z_2|z_1) = \frac{p(z_1, z_2)}{p(z_1)} = \frac{p(i, j)\; f_{i,j}(y_1, y_2)}{\displaystyle\sum_{j=1}^{K} p(i, j)\; f_{i,j}(y_1)}. \tag{3}$$

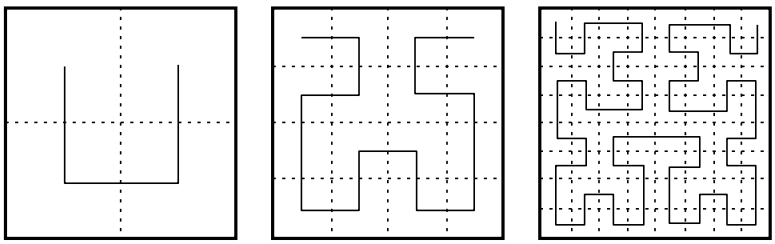


Figure 1. Construction of a Hilbert-Peano scan for a 8×8 image (initialisation, intermediate stage and result).

2.2. Relations to Hidden Markov Chain

It has been shown (see Ref. 10) that (i) $f_{i,j}(y_1) = f_i(y_1)$ implies that X is a Markov chain and, (ii) reciprocally the markovianity of X implies $f_{i,j}(y_1) = f_i(y_1)$. These properties state a necessary and sufficient condition that a PMC be an HMC and show that a PMC is strictly more general than a HMC. Furthermore, the distribution of a classical stationary HMC, in which X is Markovian, can be written as

$$p(z) = p(x, y) = p(x_1)\, f_{x_1}(y_1)\; p(x_2 | x_1)\, f_{x_2}(y_2)\; \ldots\; p(x_N | x_{N-1})\, f_{x_N}(y_N)\,.$$

This is a particular PMC defined by $p(z_1, z_2) = p(i, j)\, f_i(y_1)\; f_j(y_2)$. This illustrates the classical hypotheses according to which the random variables $Y_1, \ldots, Y_N$ are independent conditionally on X, and the distribution of each Y_n conditional to X is equal to its distribution conditional to X_n. The classical HMC cannot take into account situations in which $f_{i,j}(y_1)$ does depend on j. This can be a serious drawback since such dependencies may occur. For example, following Ref. 10, consider the problem of restoring a two classes image with "Forest" and "Water": $\Omega = \{F, W\}$. For $i = F$, the random variable Y_n models the natural variability of the forest and, possibly, other "noise" which is not considered here. Considering (F, F) and (F, W) as two possibilities for (i, j), it seems quite natural that $f_{F,F}(y_1)$ and $f_{F,W}(y_1)$ can be different. In fact, the second case can appear when some trees are near the water, giving them a different visual aspect. More generally, the possible dependence of $f_{i,j}(y_1)$ on j allows one to easily model the fact that the visual aspect of a given class can be different near a boundary than inside a large set of pixels of a same class. Hence, a kind of "non stationary", which models the fact that the noise can be different close to boundaries, is taken into account in the context of a stationary PMC model. Regarding the problem of restoring SAR images, the main interest of PMC is to allow to take into account spatially correlated speckle noise, which can not be done in the HMC case.

2.3. ICE principle and segmentation algorithm overview

As specified above, the distribution of a PMC Z is given by the distribution $p(z_1, z_2)$. From Eq. 1, the set θ of parameters is made up of the K^2 probabilities $p(i, j)$, and the parameters that describe the K^2 densities $f_{i,j}(y_1, y_2)$. In the case of unsupervised classification, these parameters are unknown and must be estimated from the observed data $Y = y$. We use here the ICE procedure whose principle is the following. Let P_Z^θ be the distribution of $Z = (X, Y)$, depending on a parameter $\theta \in \Theta$. The problem is to estimate θ from a sample $y = (y_1, \ldots, y_N)$. The ICE procedure is based on the following assumptions:

1. There exists an estimator of θ from the complete data: $\hat{\theta} = \hat{\theta}(z)$;
2. For each $\theta \in \Theta$, either the conditional expectation $E_\theta\left[\hat{\theta}(Z) | Y = y\right]$ is computable or simulation of X according to its distribution conditional to $Y = y$ is feasible.

The ICE procedure is an iterative method which runs as follows: Initialize $\theta = \theta_0$; For $q > 0$,

- if the conditional expectation is computable, put $\theta_{q+1} = E_{\theta_q}\left[\hat{\theta}(Z) | Y = y\right]$,
- if not, simulate L realizations $x_1, \ldots, x_L$ of X (each x_l is here a sequence) according to its distribution conditional on $Y = y$ and based on θ_q, and put $\theta_{q+1} = \frac{1}{L} \sum_{l=1}^{L} \hat{\theta}(x_l, y)$.

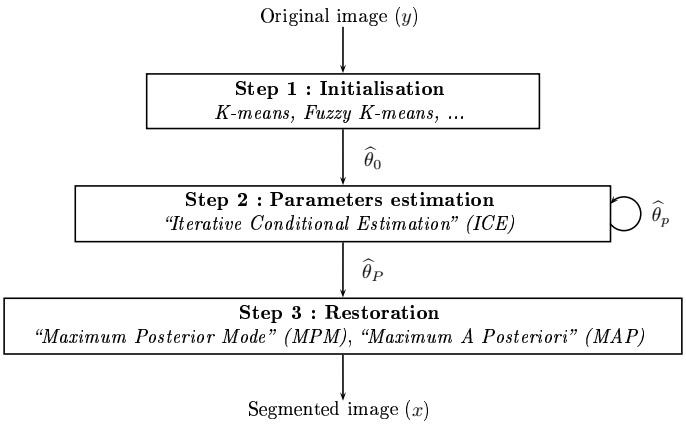


Figure 2. Diagram of the complete ICE based restoration algorithm.

When unsupervised restoration is concerned, the ICE procedure can be incorporated into the algorithm sketched in figure 2. Given the original sequence y (the Hilbert-Peano scan of the image), the first stage allows to obtain an initial estimation $\hat{\theta_0}$ of the set of parameters θ describing the model. Then the ICE procedure is applied iteratively until convergence or until a given number of iterations is reached. The estimated parameters $\hat{\theta}_P$ are then used for Bayesian restoration, and an inverse Hilbert-Peano scan allows to recover the segmented image x.

The next two sections show how the ICE can be used to perform the estimation of the two sets of PMC parameters from the observed data $Y = y$. For later use, let $\mathcal{A}_{i,j}$ denotes the set of indices $1 \leq n \leq N-1$ for which $(x_n, x_{n+1}) = (i, j)$.

3. MODEL PARAMETERS ESTIMATION

Similar to HMC, to estimate the PMC parameters $p(i,j)$, we need the following distributions :

- "Forward" and "backward" probabilities

$$\begin{aligned} \alpha_n(j) = &\quad p(y_1, \ldots, y_n, x_n = i), \quad 1 \leq n \leq N, \\ \beta_n(j) = &\quad p(y_{n+1}, \ldots, y_N | x_n = j), \quad 1 \leq n \leq N. \end{aligned}$$

The "forward" and "backward" probabilities can be computed recursively, using Baum's algorithm[7]. However, the original algorithm is subject to serious numerical problems and we use its "normalized" variant[19], which can be iteratively calculated by:

$$\alpha_1(i) = \frac{p(x_1 = i, y_1)}{\sum_{\omega \in \Omega} p(x_1 = \omega, y_1)}, \quad \alpha_{n+1}(i) = \frac{\sum_{\omega \in \Omega} \alpha_n(\omega)\, p(x_{n+1} = i, y_{n+1} | x_n = \omega, y_n)}{\sum_{(\omega_1, \omega_2) \in \Omega^2} \alpha_n(\omega_1)\, p(x_{n+1} = \omega_2, y_{n+1} | x_n = \omega_1, y_n)}, \tag{4}$$

and

$$\beta_N(j) = 1, \quad \beta_n(j) = \frac{\sum_{\omega \in \Omega} \beta_{n+1}(\omega)\, p(x_{n+1} = \omega, y_{n+1} | x_n = j, y_n)}{\sum_{(\omega_1, \omega_2) \in \Omega^2} \alpha_n(\omega_1)\, p(x_{n+1} = \omega_2, y_{n+1} | x_n = \omega_1, y_n)}. \tag{5}$$

- Joint *a posteriori* probabilities The "joint *a posteriori* probability" of two subsequent classes given all the observations can be expressed as

$$\begin{aligned}\Psi_n(i,j) &= p(x_n = i, x_{n+1} = j \,|y\,), \\ &= \frac{\alpha_n(i) p\,(x_{n+1} = j, y_{n+1} \,|x_n = i,\, y_n)\, \beta_{i+1}(j)}{\sum_{(\omega_1,\omega_2)\in\Omega^2} \alpha_n(\omega_1) p\,(x_{n+1} = \omega_2, y_{n+1} \,|x_n = \omega_1,\, y_n)\, \beta_{n+1}(\omega_2)}, \quad 1 \le n \le N-1. \end{aligned} \tag{6}$$

- Marginal *a posteriori* probabilities The "marginal *a posteriori* probability" of two subsequent classes given all the observations can be expressed as

$$\begin{aligned}\Phi_n(i) &= p(x_n = i \,|y\,), \\ &= \frac{\alpha_n(i)\beta_n(i)}{\sum_{\omega\in\Omega} \alpha_n(\omega)\beta_n(\omega)}, \quad 1 \le n \le N. \end{aligned} \tag{7}$$

Similar to HMC, the last equation allows the Bayesian MPM restoration of the hidden sequence X. Let us also notice that equations (4) to (7) give the classical HMC formulas when the PMC is an HMC.

Otherwise, the joint *a priori* probabilities $p(i,j)$ can be estimated from X by using the following standard estimation $\hat{p}(i,j) = \frac{Card(\mathcal{A}_{i,j})}{N-1}$. Applying ICE, the conditional expectation of $\hat{p}(i,j)$ at iteration $q+1$ is computable and gives

$$\hat{p}^{q+1}(i,j) = E_{\theta_q}\left[\hat{p}(i,j)\,|Y = y\right] = \frac{1}{N-1}\sum_{n=1}^{N-1} p^q\,(x_n = i, x_{n+1} = j\,|y\,), \tag{8}$$

where $p^q\,(x_n = i, x_{n+1} = j\,|y\,)$ are calculated from (6), at iteration q.

Let us notice that, similar to the HMC case, it is possible to simulate X according to its distribution conditional on $Y = y$. Indeed, the *a posteriori* distribution of X, i.e. $p(x\,|y\,)$, is that of a non stationary Markov chain[11] with transition matrix:

$$t_{n+1}(i,j) = p\,(x_{n+1} = i\,|x_n = j, y\,) = \frac{\Psi_n(i,j)}{\Phi_n(i)}, \quad 1 \le n \le N-1. \tag{9}$$

4. MIXTURE PARAMETERS ESTIMATION

The parameters of the K^2 possibly correlated bi-dimensional densities $f_{i,j}\,(y_1, y_2)$ are not known and must be estimated from the observation $Y = y$. In radar image segmentation, the choice for the family of distributions is crucial since it models our *a priori* knowledge on the nature and shape of the noise, induced by the backscattering mechanisms. One simple choice is to consider that all distributions are bi-dimensional Gaussian and we first present the *classic mixture estimation* algorithm. However, on the first hand, it is well known that Gaussian densities are not suited to the segmentation of SAR images. On the other hand, non Gaussian bivariate distributions can sometimes be difficult to obtain explicitly*. Hence, in section 4.2, we propose a method that reduces the problem into the estimation of two 1D independent distributions, each of them belonging to a finite set of possible shapes. Finally, the Pearson system of distributions is introduced and its use in the *generalized mixture estimation* algorithm is presented.

*see Ref. 20 for a synthesis on continuous multivariate distributions.

4.1. ICE in Gaussian PMC

A PMC is called Gaussian if all densities $f_{i,j}(y_1, y_2)$ in (1) are Gaussian. Denoting by $\mu_{i,j}$ the mean vector and by $\mathbf{\Gamma}_{i,j}$ the variance-covariance matrix of $f_{i,j}(y_1, y_2)$, we may use the following estimates from (X, Y):

$$\hat{\mu}_{i,j}(z) = \frac{1}{Card(\mathcal{A}_{i,j})} \sum_{n=1}^{N-1} \mathbf{1}_{\mathcal{A}_{i,j}} \begin{pmatrix} y_n \\ y_{n+1} \end{pmatrix}, \tag{10}$$

and

$$\hat{\mathbf{\Gamma}}_{i,j}(z) = \frac{1}{Card(\mathcal{A}_{i,j})} \sum_{n=1}^{N-1} \mathbf{1}_{\mathcal{A}_{i,j}} \left(\begin{pmatrix} y_n \\ y_{n+1} \end{pmatrix} - \hat{\mu}_{i,j}(z) \right)^t \left(\begin{pmatrix} y_n \\ y_{n+1} \end{pmatrix} - \hat{\mu}_{i,j}(z) \right). \tag{11}$$

The conditional expectations of $\hat{\mu}_{i,j}$ and $\hat{\mathbf{\Gamma}}_{i,j}$ are not computable and thus we must perform simulations by using (9). In practice, just one realization of $X = x^q$ is sampled according to its distribution conditional to $Y = y$ and based on $\hat{\theta}_q$. Thus the next $\hat{\mu}_{i,j}^{q+1}$ and $\hat{\mathbf{\Gamma}}_{i,j}^{q+1}$ are estimated using (10) and (11), with x^q instead of x for the computation of $\mathcal{A}_{i,j}$.

4.2. ICE in generalized PMC

When the nature of the K^2 densities $f_{i,j}(y_1, y_2)$ on $\mathbb{R}^2$ is not known, but each of them belongs to a fixed set $H_{i,j}$ of possible shapes, the problem of finding them is called "generalized mixture estimation". Let us notice that when each set $H_{i,j}$ is reduced to one element, the generalized mixture becomes a classical mixture. Furthermore, when all elements are Gaussian families, we find a Gaussian mixture and the generalized ICE becomes a classical ICE in PMC as specified previously. In some situations, $f_{i,j}(y_1, y_2)$ may be difficult to obtain explicitly. For example, we have an idea about $f_{i,j}(y_1)$ and $f_{i,j}(y_2)$, and we have their correlation. Of course, this does not give us the density $f_{i,j}(y_1, y_2)$ which is needed in generalized ICE iterations. We can then use the following method, inspired from the method successfully applied to bi-sensor image segmentation in Ref. 6.

Let denote σ_1 and σ_2 the standard deviations of Y_1 and Y_2, and ρ their correlation coefficient. Considering

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \underbrace{\frac{|\rho|}{w\, v} \begin{pmatrix} 1 & \frac{w\,\sigma_2 - \sigma_1}{\rho\,\sigma_2} \\ \frac{w\,\sigma_1 - \sigma_2}{\rho\,\sigma_1} & 1 \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, \tag{12}$$

with $w = \sqrt{1-\rho^2}$ and $v = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\,\sigma_1\,\sigma_2\,w}$, we have: (i) $Var(U_1) = Var(U_2) = 1$, (ii) $Cov(U_1, U_2) = 0$, and (iii) $Cov(U_1, Y_1) = Cov(U_2, Y_2)$. The condition (iii) makes the choice for (U_1, U_2) unique. This ensures that the general PMC condition according to which the marginals of $p(z_1, z_2)$ are equal. Such condition was not needed in Ref. 6.

As the random variables (U_1, U_2) with densities g_1 and g_2 are not correlated, we simplify things by assuming that they are independent, and $g(u_1, u_2) = g_1(u_1)\, g_2(u_2)$. At each iteration of the generalized ICE and for each $(i, j) \in \Omega^2$, we seek g_1 and g_2. Since they are densities on $\mathbb{R}$, this is much simpler than to seek a density on $\mathbb{R}^2$. The density f of the distribution of (Y_1, Y_2) is then linked to g_1 and g_2 by the relation

$$f_{i,j}(y_1, y_2) = |\det(\mathbf{A})|\; g_1(u_1)\, g_2(u_2). \tag{13}$$

Hence, the mixture parameters that have to be estimated is first σ_1, σ_2 and ρ and, second, depend on the family of the candidate distributions g_1 and g_2. If both g_1 and g_2 belong to the Gaussian family, we then find again the classical PMC-ICE algorithm. Let us now examine the more general case where the families of g_1 and g_2 belong to the Pearson system.

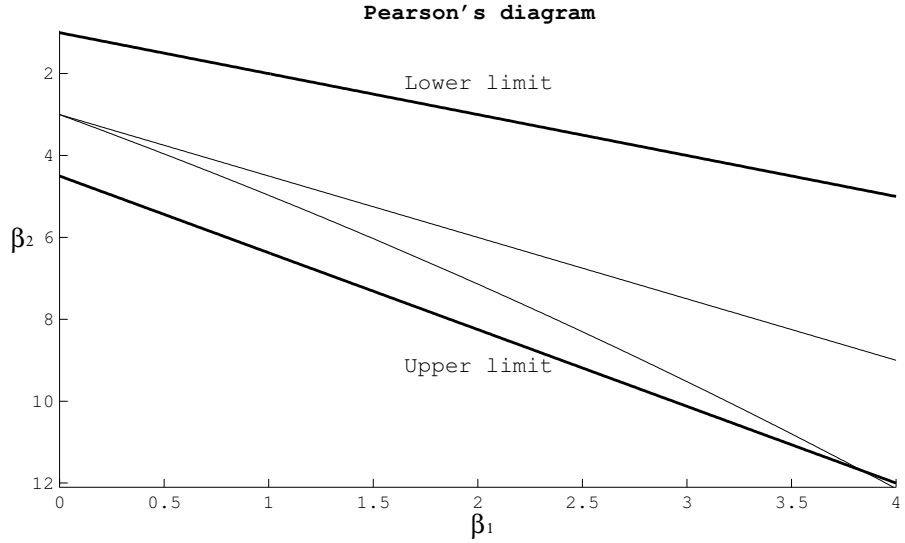


Figure 3. Pearson's diagram in term of skewness β_1 and kurtosis β_2. Note that the β_2 axis is reversed.

4.3. ICE in Pearson PMC

The problem is now reduced to the search for one-dimensional densities. In order to take into account theoretical results in the modelling of backscattering mechanisms, a number of candidate families has been proposed, such as Gamma, K and Beta densities[21, 22]. The interest in these distributions comes essentially from the large variety of possible shapes that can be obtained by modifying a limited number of parameters. In particular, all of them can take into account the dissymmetry of class densities, which is not the case of Gaussian distributions. Another interesting point is that some of them have a finite or semi-finite support, which is of great interest in SAR image processing. In order to enlarge the set of available shapes, one solution is to consider, not only one of the families cited above, but all of them in a unified way with the help of the Pearson system of distributions. Comprehensive introduction and detailed statements on the Pearson system are given in Ref. 23.

This system consists of mainly eight families of distributions of various types with mono-modal and possibly non symmetrical shapes (Gamma, Exponential and Beta distributions among others) and has shown to be efficient in HMC in the context of radar image segmentation[5, 24]. Let μ_2, μ_3 and μ_4 denote the centered moments of order 2, 3 and 4. All the families can be expressed in terms of the mean (μ_1), variance (μ_2), skewness ($\beta_1 = \mu_3^2/\mu_2^3$) and kurtosis ($\beta_2 = \mu_4/\mu_2$) and located in the so-called Pearsons diagram, cf. Fig. 3. Gaussian distributions are located in ($\beta_1 = 0$, $\beta_2 = 3$). Type II and Student't distributions are respectively situated according to ($\beta_1 = 0, 1 < \beta_2 < 3$) and ($\beta_1 = 0, 3 < \beta_2 < 4.5$). Gamma distributions are located on the straight line, and inverse Gamma distributions are located according to the second curve[†]. Beta distributions of the first kind are situated between the lower limit and the Gamma line, Beta distributions of the second kind are located between the Gamma and the inverse Gamma distributions, and Type IV distributions are located between the inverse Gamma distribution and the upper limit.

We thus seek both the family of the distribution and the parameters that best describe its samples. From a realization x of X by using (9), one can estimate the empirical moments using classical estimators (or more sophisticated estimators as the ones proposed in Ref. 25) and compute (β_1,β_2). Given the Pearson's diagram, it becomes possible to select the corresponding member family and recover the parameters, which precisely identify the distribution, from the estimated moments.

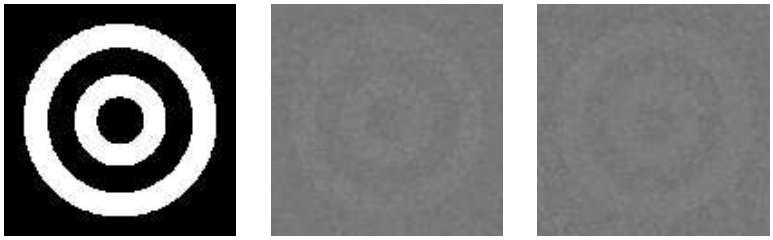
[†]Gamma line : $\beta_2 = 1.5\,\beta_1 + 3$, inverse Gamma curve: $\frac{3}{32-\beta_1}\left(13\beta_1 + 16 + 2\,(\beta_1+4)^{1.5}\right)$ for $0 \leq \beta_1 \leq \frac{96}{25}$.

5. EXPERIMENTS ON NOISY SIMULATED IMAGES

In this section, we present a set of experiments concerning the segmentation of a noisy simulated image corrupted with some correlated noise and then restored by HMC and PMC based Bayesian segmentation methods. The question is whether PMC have to be used instead of HMC or not, in such situation. Fig. 4 shows the original image and the two noisy image with correlated noise generated according to : (1) the two classes have been corrupted with two independent noises (parameters are reported in Tab. 1); (2) the correlation is obtained by performing the following filtering on the entire image :

$$\frac{1}{3.8}\begin{pmatrix} 0 & 0.7 & 0 \\ 0.7 & 1 & 0.7 \\ 0 & 0.7 & 0 \end{pmatrix}.$$

In this way, the test images are corrupted with two correlated noises. Note that such images neither represent a HMC nor a PMC process. However, they are segmented in an unsupervised manner by HMC and PMC based MPM restoration methods, with all parameters estimated by ICE algorithm, as specified in Sections 3 and 4. We test HMC and PMC classical mixture estimation (Gaussian densities) on the first noisy image and HMC and PMC generalized mixture estimation (Pearson system) on the second noisy image. The number of ICE iterations is set to 100, and we compute only one *a posteriori* realization per iteration for the estimation of the mixture parameters.



(a) (b) (c)

Figure 4. Original image and noisy simulated images (128×128). $\Omega = \{Black, White\}$. The "noise parameters" are reported in Table 1.

Table 1. Noise parameters used for the two experiments.

Image	Noise parameter values							
	Law		μ_1	μ_2	μ_3	μ_4	β_1	β_2
(b)	f_{Black}	Gaussian	120.0	49.0	0.0	7203.0	0.0	3.0
	f_{White}	Gaussian	125.0	75.0	0.0	16875.0	0.0	3.0
(c)	f_{Black}	Gamma	120.0	49.0	343.0	10804.5	1.0	4.5
	f_{White}	Inv. Gamma	125.0	75.0	918.6	40159.1	2.0	7.1

The segmentation results are reported in Fig. 5. Both the error rates and a visual inspection show that ICE-PMC based segmentations using classical and generalized mixture estimation [images (b) and (d)] give better results than the corresponding ICE-HMC ones [image (a) and (c)]. We can also note that the error rates are a little bit bigger in the case of generalized mixture. One reason is that the mixture is much more difficult to estimate since the "degree of freedom" is greater than in the Gaussian case. Table 2 gives the estimated parameters found with PMC-ICE and generalized mixture estimation. Most of the 1D densities selected among the Pearson system are Beta of the first kind. In order to illustrate the approximation presented in section 4.2, we also reported in Fig. 6 the bi-dimensional density obtained from equation (13), for $(i, j) = (0, 0)$.

(a) G-HMC - 13.98% (b) G-PMC - 6.66% (c) P-HMC - 15.34% (d) P-PMC - 8.36%

Figure 5. Segmentation results obtained with ICE estimation and MPM classification for both HMC et PMC models. 'G' and 'P' stand respectively for Gaussian and Pearson. The percentages give the error rates of misclassification according to the original image in Fig. 4.

Table 2. Parameter values estimated using Pearson PMC based ICE algorithm.

$$\hat{p}\,(i,j) = \begin{pmatrix} 0.3399 & 0.0086 \\ 0.0086 & 0.6428 \end{pmatrix},$$

$f_{i,j}$	Estimated parameters for the 1D densities							
	ρ	1D Law	μ_1	μ_2	μ_3	μ_4	β_1	β_2
$f_{0,0}$	0.53	$g_1^{0,0}$ Type IV	24.69	1.00	0.44	3.40	0.20	3.40
		$g_2^{0,0}$ Beta 1	26.52	1.00	0.37	3.18	0.14	3.18
$f_{0,1}$	0.60	$g_1^{0,1}$ Beta 1	37.21	0.99	0.85	2.93	0.74	2.97
		$g_2^{0,1}$ Beta 1	-22.50	0.99	0.72	3.13	0.52	3.18
$f_{1,0}$	0.35	$g_1^{1,0}$ Beta 1	24.89	0.99	0.72	2.71	0.53	2.75
		$g_2^{1,0}$ Beta 1	3.15	0.99	0.15	2.49	0.02	2.53
$f_{1,1}$	0.48	$g_1^{1,1}$ Beta 1	31.56	1.00	0.32	3.08	0.10	3.08
		$g_2^{1,1}$ Beta 1	31.60	1.00	0.34	3.12	0.12	3.12

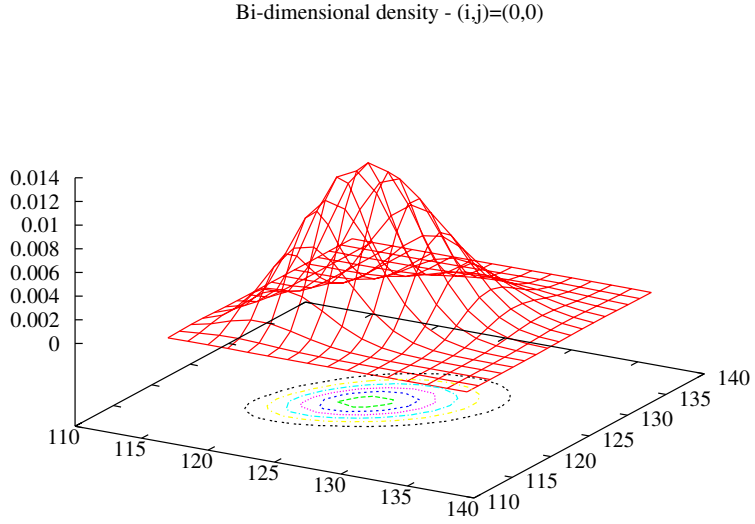


Figure 6. Plot of the bi-dimensional density $f_{0,0}$ obtained from the estimation of $g_1^{0,0}$ and $g_2^{0,0}$.

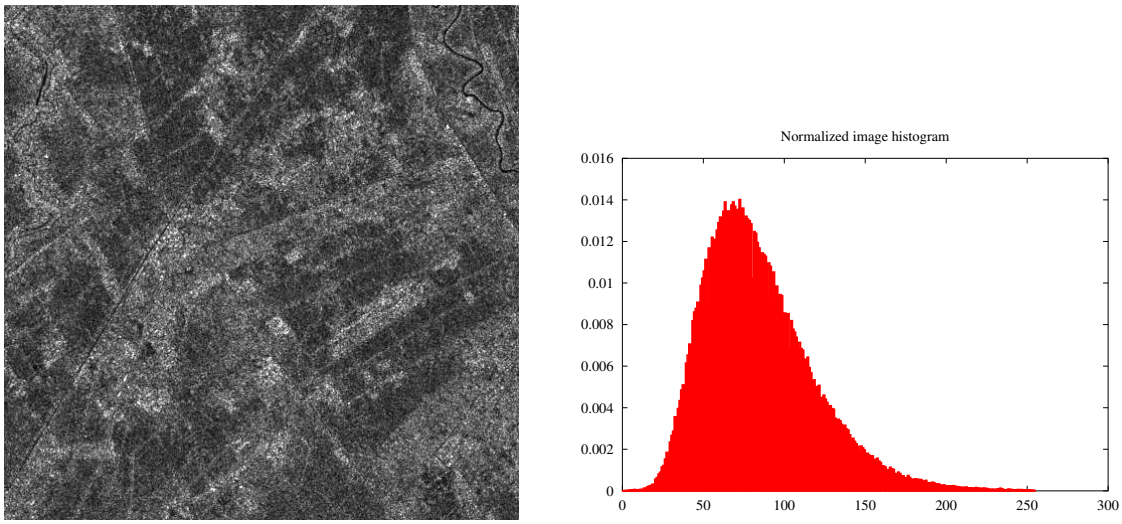


Figure 7. ERS-1 SAR image of a rice plantation in Java Island (Indonesia) and histogram. Image size : 512×512. Date : 1994, 10 february.

6. EXPERIMENTS ON A REAL SAR IMAGE

Figure 7 shows an extract of a JERS image of rice growing in Semarang (Java island) with mainly early rice, late rice and other cultivations[26]. It was decided to classify the image into four classes with again 100 ICE iterations and one realization per iteration. We only test the generalized mixture case for both HMC and PMC models. The MPM classification results are reported in figure 8, together with the normalized histogram of each class. It is clear that the PMC based segmentation is much more regular than the HMC one and rice plots seem to be better extracted. Note also that only three classes have been detected with the HMC model (four attempted), and the configurations of class histograms are quite different for the two models.

The computation time between the two algorithms is quite different. The program based on HMC spent about 35 minutes on a PC with Pentium IV 1.3 GHz processor running Linux, whereas the program based on PMC needs 2 hours and 40 minutes. The complexity of the PMC-ICE algorithm is more important since not only $K = 4$ densities have to be estimated (HMC case), but $2\,K^2 = 32$ one-dimensional densities (PMC case). Another time consuming point of the algorithm comes from the estimation of $f_{i,j}(y_1) = \int f_{i,j}(y_1, y_2)\, dy_2$ which happens in the computation of $p(z_n)$ and $p(z_{n+1}\,|z_n\,)$. In the generalized mixture case, no analytic calculation can be made and numerical integration must be used.

7. CONCLUSIONS AND FURTHER WORK

This article describes unsupervised classification of SAR images in the framework of the recent model called "Pairwise Markov Chain". The method is based on a variant of the general Iterative Conditional Estimation (ICE) and Bayesian Maximum a Posteriori Mode (MPM). As the PMC model is more general – and more complex – than the HMC one, the aim of our work was to examine whether the use of PMC instead of HMC is justified in this context. Numerous experiments, the results of some of them are presented in the paper, show that the use of PMC is of interest and can significantly improve the segmentation obtained using HMC. Methods based on PMC constitute an interesting alternative mainly because the spatial correlated speckle noise can be directly taken into account in the model. This generally gives more regular class in the segmented images. However, compared to ICE-HMC, the parameters estimation is time consuming since the complexity of the ICE-PMC algorithm is more important.

As perspectives for further work, we may mention the extension of PMC to multisensor image processing. Both spatial and inter-sensor correlation should be taken into account at the same time. The extension of PMC

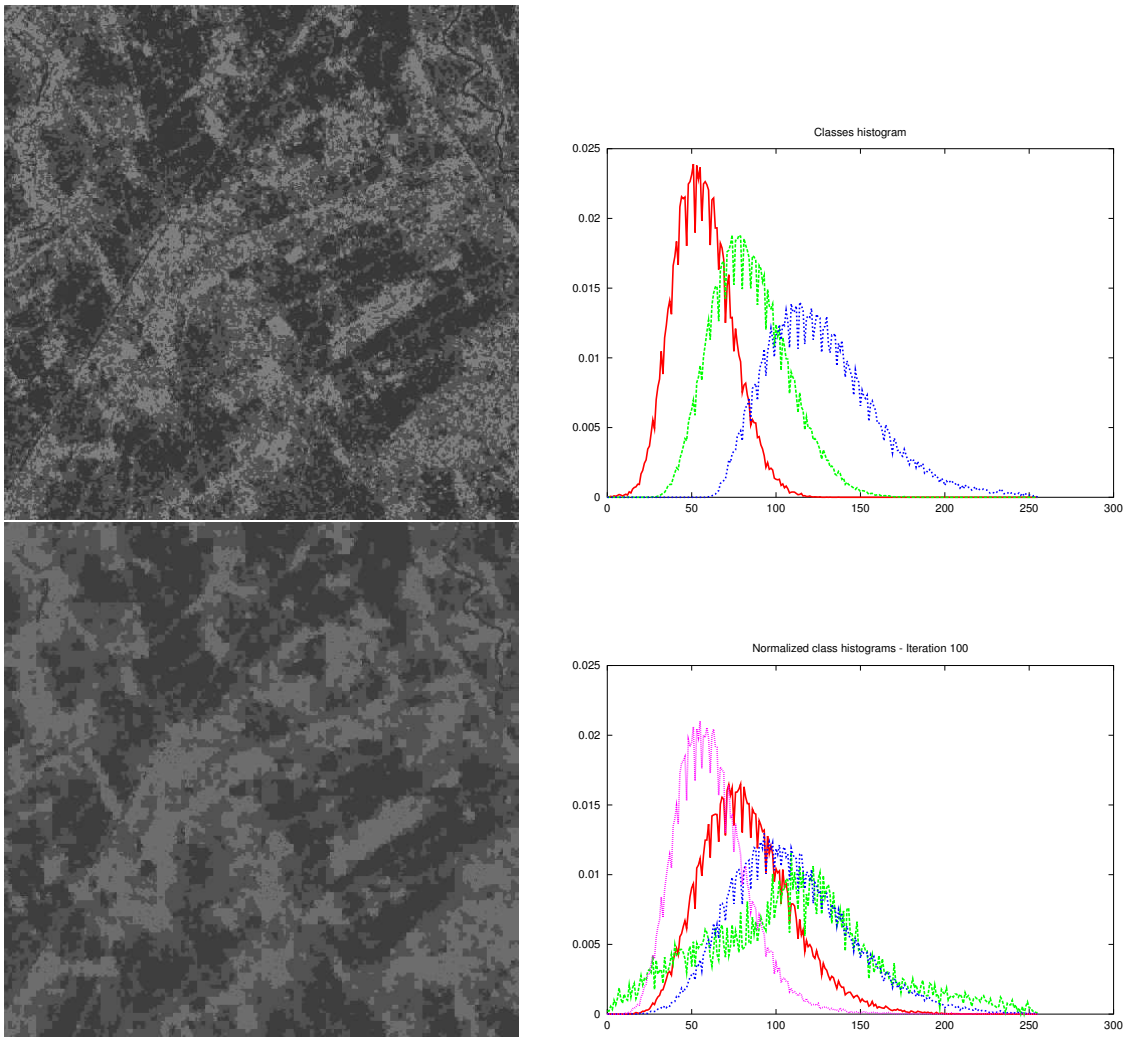


Figure 8. Segmentation results and classes histograms for both ICE and HMC classification (up) and ICE and PMC classification (down). The class values $(1, ..., 4)$ has been changed to the mean values of the classes to facilitate visual comparison with the original image.

to the family of models called "Hidden Markov Tree", described in Ref. 27, would also possibly be of interest for multiresolution image processing. Another direction concerns the new model called "Triplet Markov Chain"[28] and its application in unsupervised image segmentation.

REFERENCES

1. B. Benmiloud and W. Pieczynski, "Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images," *Traitement du Signal* **12**(5), pp. 433–454, 1995. in french.
2. A. El-Jacoubi, M. Gilloux, R. Sabourin, and C. Suen, "An HMM-based approach for off-line unconstrained handwritten word modeling and recognition," *IEEE Trans. on PAMI* **21**(8), pp. 752–760, 1999.
3. K. Aas, L. Eikvil, and R. Huseby, "Applications of hidden Markov chains in image analysis," *Pat. Recog.* **32**(4), pp. 703–713, 1999.
4. A. Wilson and A. Bobick, "Parametric hidden Markov models for gesture recognition," *IEEE Trans. on IP* **8**(9), pp. 884–900, 1999.

5. N. Giordana and W. Pieczynski, "Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation," *IEEE Trans. on PAMI* **19**(5), pp. 465–475, 1997.
6. W. Pieczynski, J. Bouvrais, and C. Michel, "Estimation of generalized mixture in the case of correlated sensors," *IEEE Trans. on IP* **9**(2), pp. 308–311, 2000.
7. L. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statistic.* **41**, pp. 164–171, 1970.
8. G. Fornay, "The Viterbi algorithm," *Proc. of the IEEE* **61**(3), pp. 268–277, 1973.
9. W. Pieczynski and A. Tebbache, "Pairwise Markov random fields and segmentation of textured images," *Machine Graphics and Vision* **9**(3), pp. 705–718, 2000.
10. W. Pieczynski, "Pairwise Markov chains," *IEEE Trans. on PAMI* . To appear.
11. W. Pieczynski, "Pairwise Markov chains and bayesian unsupervised fusion," in *Proc. of the Int. Conf. FUSION'00*, **1**, pp. 24–31, (Paris, France), 10-13 July 2000.
12. W. Pieczynski, "Statistical image segmentation," *Machine Graphics and Vision* **1**(2), pp. 261–268, 1992.
13. M. Mignotte, C. Collet, P. Pérez, and P. Bouthémy, "Sonar image segmentation using an unsupervised hierarchical MRF model," *IEEE Trans. on IP* **9**(7), pp. 1216–1231, 2000.
14. M. Mignotte, J. Meunier, J. Soucy, and C. Janicki, "Comparison of deconvolution techniques using a distribution mixture parameter estimation: Application in SPECT imagery," *J. of Electr. Imaging* **1**(1), pp. 11–25, 2002.
15. R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin, "Unsupervised segmentation of radar images using hidden Markov chains and hidden Markov fields." Submitted to IEEE Trans. on GRS, 2002.
16. Z. Kato, J. Zeroubia, and M. Berthod, "Unsupervised parallel image classification using Markovian models," *Pat. Recog.* **32**(4), pp. 591–604, 1999.
17. L. Fouque, A. Appriou, and W. Pieczynski, "Multiresolution hidden Markov chain model and unsupervised image segmentation," in *Proc. of the IEEE SSIAI'2000*, pp. 121–125, (Austin, Texas), 2-4 April 2000.
18. W. Skarbeck, "Generalized Hilbert scan in image printing," in *Theoretical Foundations of Computer Vision*, R. Klette and W. G. Kropetsh, eds., Akademie Verlag, Berlin, 1992.
19. P. Devijver, "Baum's Forward-Backward algorithm revisited," *Pat. Recog. Let.* **3**, pp. 369–373, 1985.
20. S. Kotz, N. Balakrishnan, and N. Johnson, *Continuous multivariate distributions, Volume 1 : models and applications*, John Wiley and Sons, Inc, second ed., 2000.
21. J. Jao, "Amplitude distribution of composite terrain radar clutter and the K-distribution," *IEEE Trans. on Antennas and Propagation* **32**(10), pp. 1049–1062, 1984.
22. A. Maffett and C. Wackerman, "The modified Beta density function as a model for SAR clutter statistics," *IEEE Trans. on IP* **29**(2), pp. 277–283, 1991.
23. N. Johnson and S. Kotz, *Continuous univariate distributions, Volume 1*, John Wiley and Sons, Inc, second ed., 1994.
24. S. Derrode, G. Mercier, J. Le Caillec, and R. Garello, "Estimation of sea-ice SAR clutter statistics from Pearson's system of distributions," in *Proc. of the IEEE IGARSS'01*, (Sydney, Australia), 9-13 July 2001.
25. R. Parrish, "On an integrated approach to member selection and parameter estimation for Pearson distributions," *Computational Statistics & Data Analysis* **1**, pp. 239–255, 1983.
26. T. Le Toan and al, "Assessment of rice using ERS-1 SAR data : the SARI project (Satellite Assessment of Rice in Indonesia)," in *Proc. of Remote Sensing and GIS Environmental Resources Management*, (Jakarta), 1995.
27. W. Pieczynski, "Pairwise Markov tree," *C. R. Acad. Sci. Paris* **Ser. I 335**, pp. 79–82, 2002. in french.
28. W. Pieczynski, C. Hulard, and T. Veit, "Triplet Markov chains in hidden signal restoration," in *Proc. of the SPIE'02*, (Crete, Greece), 22-27 September 2002.