

Unsupervised Statistical Segmentation of Nonstationary Images Using Triplet Markov Fields

Dalila Benboudjema and Wojciech Pieczynski

Abstract—Recent developments in statistical theory and associated computational techniques have opened new avenues for image modeling as well as for image segmentation techniques. Thus, a host of models have been proposed and the ones which have probably received considerable attention are the hidden Markov fields (HMF) models. This is due to their simplicity of handling and their potential for providing improved image quality. Although these models provide satisfying results in the stationary case, they can fail in the nonstationary one. In this paper, we tackle the problem of modeling a nonstationary hidden random field and its effect on the unsupervised statistical image segmentation. We propose an original approach, based on the recent triplet Markov field (TMF) model, which enables one to deal with nonstationary class fields. Moreover, the noise can be correlated and possibly non-Gaussian. An original parameter estimation method which uses the Pearson system to find the natures of the noise margins, which can vary with the class, is also proposed and used to perform unsupervised segmentation of such images. Experiments indicate that the new model and related processing algorithm can improve the results obtained with the classical ones.

Index Terms—Triplet Markov fields, statistical image segmentation, parameter estimation, Pearson system, iterative conditional estimation, nonstationary images, textures classification.

1 INTRODUCTION

THIS paper addresses the problem of statistical unsupervised image segmentation and its main purpose is to deal with nonstationary images. The nonstationarity considered in this paper is understood in the following probabilistic sense. In the classical Markov field context, the distribution of the hidden field is defined by some functions specified on cliques; a field will be considered nonstationary when these functions depend on the position of the cliques in the set of pixels. Therefore, nonstationary fields will produce, on average, nonstationary images in which the visual aspect of the spatial organization of different labels varies with pixels. Of course, a nonstationary field can produce a stationary image and a stationary field can produce a nonstationary image. However, real images are often nonstationary and, thus, on average, using nonstationary fields in different statistical processing should give better results. While doing so, an important problem arises: How do we estimate the parameters when they vary with pixels? Answering this question via the recent “triplet Markov field” model is the main purpose of the paper. Starting from the classical hidden Markov fields (HMF), we propose different extensions allowing one to deal with nonstationary images and possibly non-Gaussian correlated noise. The classical HMF and

Bayesian segmentation based on them can be of outstanding efficiency when dealing with the difficult—and important—problem of unsupervised image segmentation. Hundreds of papers have been written on the subject since the seminal articles [14], [22] and a rich bibliography can be seen in [11], [16], [30], [32], [37], among others. In such models, we have the hidden Markov field, $X = (X_s)_{s \in S}$, and the observed one, $Y = (Y_s)_{s \in S}$, and the problem is to estimate $X = x$ from $Y = y$. The first models, in which X is a Markov field and the random variables (Y_s) are independent conditionally on X , can give good results in many situations; however, they turn out to be too simple when considering very complex images (nonstationary, textured, strongly noisy, ... [25]). A pairwise Markov field (PMF) model has then been proposed, which consists of directly considering that the pair $Z = (X, Y)$ is a Markov field [34]. This implies that both conditional distributions $p(y|x)$ and $p(x|y)$ are Markovian: The former fact allows one to better model complex noises and the latter one still enables one to apply Bayesian segmentation. Afterward, triplet Markov fields (TMF) were proposed in which one introduces a third random field, $U = (U_s)_{s \in S}$, and assumes the Markovianity of the triplet $T = (X, U, Y)$ [1], [35]. This third field can have some physical interpretation or not; however, when the set of its values is not too large, analogous Bayesian processing can still be used to estimate $X = x$ from $Y = y$. Different ways of defining such TMF are described in [1], along with a parameter estimation method making possible unsupervised Bayesian segmentation.

In particular, one possible meaning for $U = (U_s)_{s \in S}$ is to assume that $U = u$ defines different homogeneities of (X, Y) . This means that the Markov field distribution $p(x, y|u)$ is a nonstationary one and, thus, such models enable one to deal with nonstationary (X, Y) [2]. Let us mention that a similar

• The authors are with INT/GET, Département CITI, CNRS UMR 5157, 9, rue Charles Fourier, 91000 Evry, France.
E-mail: Dalila.Benboudjema@ens.fr, Wojciech.Pieczynski@int-evry.fr.

Manuscript received 31 Jan. 2006; revised 29 June 2006; accepted 6 Sept. 2006; published online 18 Jan. 2007.

Recommended for acceptance by J. Goutsias.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0069-0106. Digital Object Identifier no. 10.1109/TPAMI.2007.1059.

study related to the use of triplet Markov chains to manage nonstationary images can be seen in [19].

Otherwise, an important problem is to manage non-Gaussian and correlated noise. In fact, such noises occur in many situations, like those related to radar images [5], [6], [13], [20], [21], [36] or sonar ones [17], [26], among others. This has not been solved, to our knowledge, in the hidden Markov fields context and we thus propose here a new model and a new related parameter estimation method.

Finally, the paper contains the following contributions:

1. The first ideas related to the use of triplet Markov fields in nonstationary images presented in [2] are extended and some new experiments are provided;
2. A new pairwise Markov field model, enabling one to deal with non-Gaussian and correlated noises, is proposed and validated by some experiments;
3. Points 1 and 2 are considered simultaneously, resulting in a new nonstationary model with non-Gaussian and correlated noise;
4. A new parameter estimation method which is based on the Pearson system and extends to the model in item 3 above, the methods proposed in [1] and [5] (some first results concerning this new model are presented in [3]).

2 TRIPLET MARKOV FIELDS AND NONSTATIONARY IMAGES

2.1 The m -Markov Nonstationary Fields

Let S be the set of pixels. The Hidden Markov field (HMF) model contains two stochastic processes, $X = (X_s)_{s \in S}$ and $Y = (Y_s)_{s \in S}$, in which $X = x$ is unobservable—or hidden—and has to be recovered from the observed $Y = y$. Therefore, $Y = y$ can be seen as a noisy version of $X = x$. In this paper, we consider the problem of digital image segmentation: In the whole paper, each X_s takes its values from a finite set of classes $\Omega = \{\omega_1, \dots, \omega_k\}$, whereas each Y_s takes its values from the set of real numbers R . The possible nonstationarity of the distribution $p(x, y)$ of the pairwise field (X, Y) is managed by introducing a third random field $U = (U_s)_{s \in S}$, each U_s taking its values from a finite set $\Lambda = \{\lambda_1, \dots, \lambda_m\}$. Thus, the idea is to interpret the m possible values of U_s as m different possible stationarities of $(X, Y) = (x, y)$. The triplet $T = (X, U, Y)$ being assumed Markovian, the posterior margins $p(x_s, u_s | y)$ are classically estimable and, thus, $p(x_s | y)$ is given by $p(x_s, y) = \sum_{u \in \Lambda} p(x_s, u_s | y)$. Having $p(x_s | y)$ enables us to perform MPM segmentation as in classical HMF.

To be more precise, let us consider a pairwise Markov field (PMF) $Z = (X, Y)$, which is Markovian with respect to the neighborhood system corresponding to a set of cliques C . Thus, the distribution of $Z = (X, Y)$ is given by $p(x, y) = \gamma \exp[-\sum_{c \in C} f_c(x_c, y_c)]$ and Z is currently said “to be stationary” when f_c does not depend on the position of c in S . Z is “nonstationary” when there can be some c such that f_c do depend on the position of c in S . Such a definition is quite satisfying; however, as the parameters in f_c does depend on the position of c in S , their estimation can be difficult to perform. As an alternative, we propose a different definition of the nonstationarity based on triplet Markov fields (TMF).

Definition. Let $Z = (X, Y) = (X_s, Y_s)_{s \in S}$ be a pairwise random field. We will say that Z has an “ m -Markov nonstationarity” (or is “ m -Markov nonstationary,” which will be denoted by

“ m -MNS”) if there exists a random field $U = (U_s)_{s \in S}$, each U_s taking its values in a finite set $\Lambda = \{\lambda_1, \dots, \lambda_m\}$, such that the triplet $T = (X, U, Y)$ is a stationary Markov field.

Let us note the following link between the two kinds of stationarity mentioned above. If Z is an m -MNS, $T = (X, U, Y)$ is a Markov field and, thus, we can say that $p(z|u)$ is a Markov nonstationary distribution. Therefore, the distribution of Z conditional on $U = u$ is “nonstationary” in the former “classical” meaning. Finally, we can say that an m -MNS field Z , is nonstationary (in the classical meaning) conditionally on U . However, an m -MNS field Z being not necessarily Markovian, is not nonstationary in the classical meaning.

Anyway, it is important that $T = (X, U, Y)$ is stationary and, thus, all of its parameters can be estimated from $Y = y$ (see [1]). As a consequence, nonstationary images can be segmented in an unsupervised way and, as shown below, such methods can significantly improve the efficiency of the classical method. Let us mention here a recent model, based on a tree structured MRF, which models the nonstationarity in a somewhat different way [31].

2.2 Experiments

In all of the experiments in this paper, we will consider the Markovianity with respect to four nearest neighbors, which is the simplest one. In fact, our aim is to study what the new models contribute with respect to the classical one rather than search for their best efficiency. Therefore, we chose to directly define rather simple energies so far to not obscure the message by too complex writing. However, more complex Markov fields can be viewed with no additional theoretical difficulties with respect to the classical hidden Markov fields. Of course, the number of parameters to be estimated and the different computation times can strongly increase when considering more complex models.

Therefore, let us consider a Markov field, $X = (x_s)_{s \in S}$, whose distribution is classically defined by the energy

$$W(x) = \sum_{(s,t) \in C_H} \alpha_H (1 - 2\delta(x_s, x_t)) + \sum_{(s,t) \in C_V} \alpha_V (1 - 2\delta(x_s, x_t)), \quad (2.1)$$

where C_H is the set of couples of pixels horizontally neighbors, C_V is the set of couples of pixels vertically neighbors, and $\delta(x_s, x_t)$ verifies $\delta(x_s, x_t) = 1$ for $x_s = x_t$ and $\delta(x_s, x_t) = 0$ for $x_s \neq x_t$. The field $X = (x_s)_{s \in S}$ is classically a Markov one with respect to four nearest neighbors. Furthermore, let us consider, classically, that the random variables $(Y_s)_{s \in S}$ are independent and Gaussian conditionally on $X = (X_s)_{s \in S}$. Assuming, in addition, that $p(y_s | x) = p(y_s | x_s)$ for each s and x_s , (X, Y) is a very classical HMF whose distribution is written

$$p(x, y) = \gamma \exp \left[-W(x) + \sum_{s \in S} \text{Log}(p(y_s | x_s)) \right]. \quad (2.2)$$

Let us imagine that this very classical model is not stationary and there are two different stationarities. Therefore, we propose to model this nonstationarity by the following m -MNS. As $m = 2$, let us put $\Lambda = \{a, b\}$ and let us consider that the Markov distribution of $T = (X, U, Y)$ is obtained by assuming that (X, U) is a Markov field and by considering that the distribution of Y conditional on (X, U)

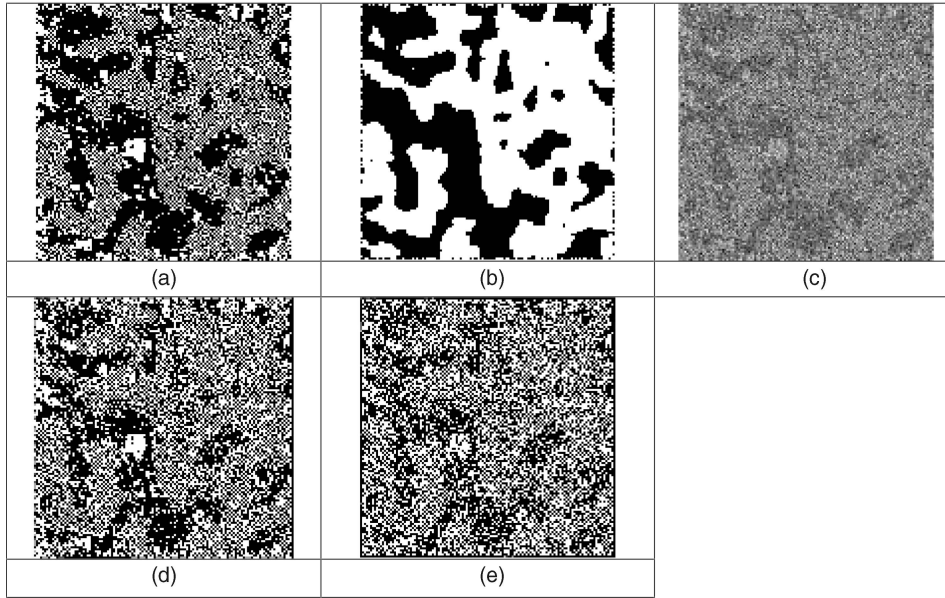


Fig. 1. (a) and (c) Simulated images using the new TMF model and (d) and (e) two MPM based on TMF and classical HMF unsupervised segmentation results. (a) $X = x$. (b) $U = u$ field. (c) $Y = y$ field. (d) New TMF-based MPM error ratio = 14.12 percent. (e) Classical HMF-based MPM error ratio = 18.6 percent.

verifies the classical condition $p(y|x, u) = \prod_{s \in S} p(y_s|x_s)$. Otherwise, the distribution of the Markov field (X, U) we propose to consider is defined by the energy

$$W(x, u) = \sum_{(s,t) \in C_H} \alpha_H^1 (1 - 2\delta(x_s, x_t)) - (\alpha_{aH}^2 \delta^*(u_s, u_t, a) + \alpha_{bH}^2 \delta^*(u_s, u_t, b)) (1 - \delta(x_s, x_t)) + \sum_{(s,t) \in C_V} \alpha_V^1 (1 - 2\delta(x_s, x_t)) - (\alpha_{aV}^2 \delta^*(u_s, u_t, a) + \alpha_{bV}^2 \delta^*(u_s, u_t, b)) (1 - \delta(x_s, x_t)), \quad (2.3)$$

with $\delta^*(u_s, u_t, a) = 1$ for $u_s = u_t = a$ and $\delta^*(u_s, u_t, a) = 0$, otherwise; $\delta(u_s, u_t, b) = 1$ for $u_s = u_t = b$ and $\delta^*(u_s, u_t, b) = 0$ otherwise.

Finally, let us define the distribution of the triplet (X, U, Y) by

$$p(x, u, y) = \gamma \exp \left[-W(x, u) + \sum_{s \in S} \log(p(y_s|x_s)) \right]. \quad (2.4)$$

Thus, (2.4) defines an m -MNS (with $m = 2$) and, conditionally on $U = u$, it defines a nonstationary Markov field of the kind (2.2). In fact, for fixed $U = u$, the distribution $p(x|u)$ is Markovian with the energy $\sum_{c \in C} f_c^u(x_c)$, where $c = (s, t)$ is either a horizontal or vertical clique. For horizontal cliques, we have $f_c^u(x_c) = \alpha_H^1 (1 - 2\delta(x_s, x_t)) - \alpha_{aH}^2 (1 - \delta(x_s, x_t))$ if $(u_s, u_t) = (a, a)$, $f_c^u(x_c) = \alpha_H^1 (1 - 2\delta(x_s, x_t)) - \alpha_{bH}^2 (1 - \delta(x_s, x_t))$ if $(u_s, u_t) = (b, b)$, and $f_c^u(x_c) = \alpha_H^1 (1 - 2\delta(x_s, x_t))$ if $u_s \neq u_t$. A similar formula being valid for vertical cliques, we see how f_c^u varies with c .

Both models HMF given by (2.2) and TMF given by (2.4) enable one to estimate $p(x_s|y)$. In HMF, this is classically done from (2.2) using the Gibbs sampler and, in (2.4), this is done in two steps: 1) estimate $p(x_s, u_s|y)$ by the Gibbs sampler and 2) calculate $p(x_s|y) = \sum_{u_s \in \Lambda} p(x_s, u_s|y)$. Therefore, the Bayesian Maximum Posterior Mode (MPM) can be used in both HMF and TMF given by (2.2) and (2.4), respectively.

Remark 2.1. Let us remark that, when $k = m$ (the number of classes is equal to the number of different stationarities), taking $x = u$ in the energy (2.3) again gives the energy $W(x)$ in (2.1). Therefore, when $x = u$, the model (2.1)-(2.2) can be seen as a particular case of the (2.3)-(2.4).

We give below four examples of using the TMF model given by (2.4) in unsupervised image segmentation. All parameters are estimated from $Y = y$ by a particular algorithm belonging to the so-called "ICE" family of methods, which is described in detail in [1].

The two first examples concern simulations with respect to the new model (2.3)-(2.4).

In each of the two cases considered, (X, U) is first simulated according to $p(x, u) = \gamma \exp[-W(x, u)]$ and then $X = x$ is corrupted with Gaussian white noise with means $m_1 = 0.$, $m_2 = 2.$, and variances $\sigma_1^2 = \sigma_2^2 = 1$. Fig. 1 corresponds to $\alpha_H^1 = \alpha_V^1 = 1.$, $\alpha_{aH}^2 = \alpha_{aV}^2 = 0.5$, $\alpha_{bH}^2 = \alpha_{bV}^2 = -0.4$, whereas Fig. 2 corresponds to $\alpha_H^1 = \alpha_V^1 = 1.$, $\alpha_{aH}^2 = 1.$, $\alpha_{aV}^2 = -0.3$, $\alpha_{bH}^2 = -0.3$, $\alpha_{bV}^2 = 1$. In both cases, the image $Y = y$ is segmented in two unsupervised Bayesian ways: the new one (TMF-based MPM) and the classical one (HMF-based MPM). We can see that the new method gives better results and, according to Fig. 2, the difference can be striking. Of course, as the data correspond to the new model, these results are not surprising; however, they show the following interesting behavior. As the classical model is known to be very robust, it was necessary to verify the existence of situations in which it is not sufficient to deal with nonstationary data obtained with the new model. In other words, these first experiments show that the new model will possibly be of interest in real situations.

The third example concerns a hand-drawn image x , noisy as above with $m_1 = 0.$, $m_2 = 1.6$, and $\sigma_1^2 = \sigma_2^2 = 1$. Such an image is neither TMF nor HMF and we can see, according to the results shown in Fig. 3, that the new method is of interest.

Finally, the fourth example concerns an application of the proposed model to a real image. Fig. 4 compares results

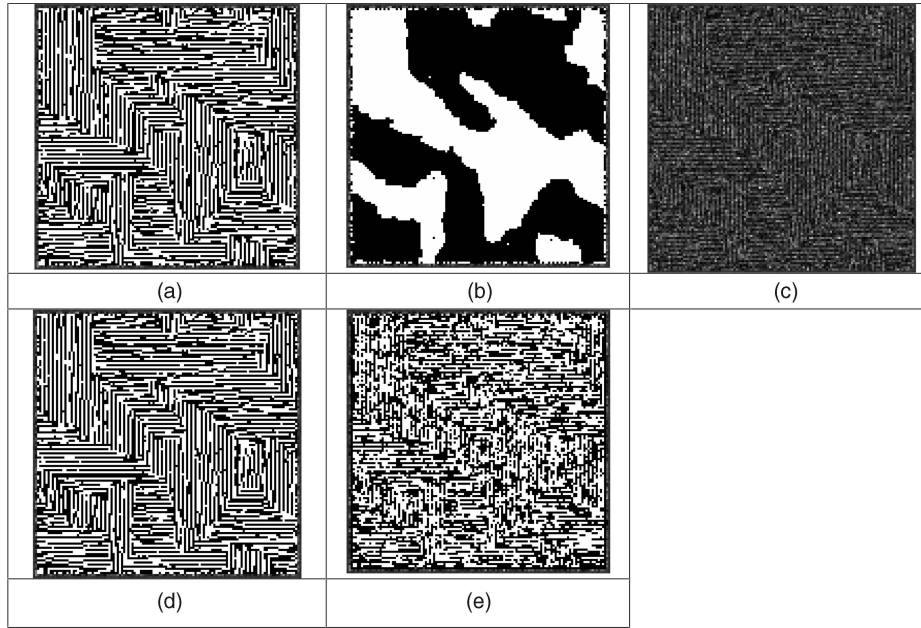


Fig. 2. Simulated images using the new TMF model and two MPM based on TMF and classical HMF unsupervised segmentation results. (a) $X = x$ field. (b) $U = u$ field. (c) $Y = y$ field. (d) New TMF-based MPM error ratio = 14.21 percent. (e) Classical HMF-based MPM error ratio = 18.6 percent.

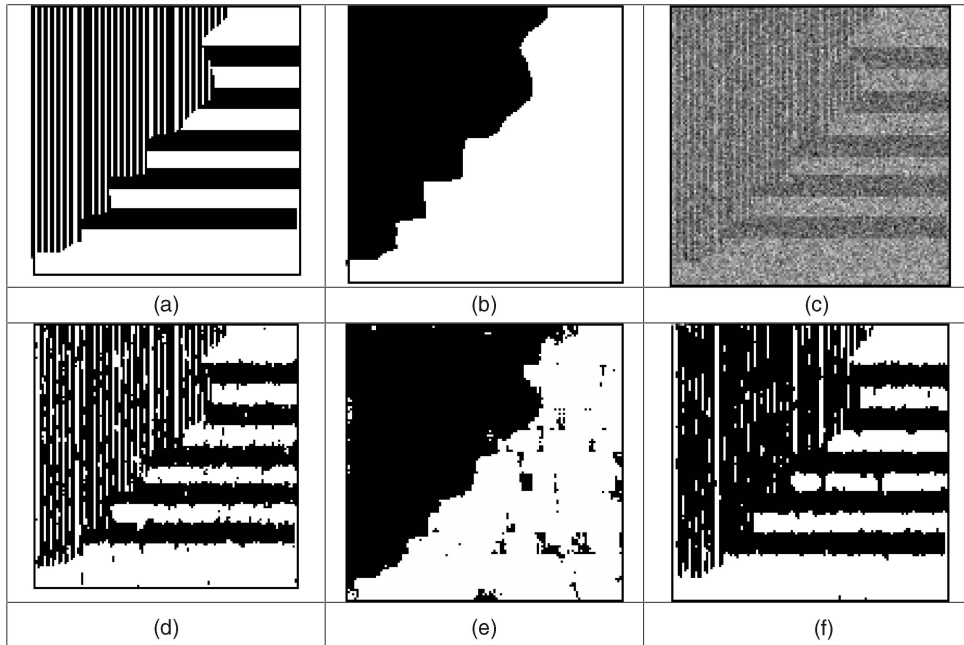


Fig. 3. (a) Hand-drawn class image x , (b) u corresponding to x , (c) its noisy version, (e) estimated u , (d) new method unsupervised segmentation result, and (f) classical method one. (a) Class image x . (b) True u . (c) Observed $Y = y$. (d) MPM based on TMF error ratio = 6.41 percent. (e) Estimated u . (f) MPM based on classical HMF error ratio = 10.37 percent.

obtained with the new TMF model-based unsupervised segmentation and the classical HMF-based one. This 430×430 radar image represents a part of The Netherlands and contains four classes. As we have no ground truth, it is difficult to draw rigorous conclusions; however, in some spots, we better recover some details like connections between canals, which are represented by the black lines in the image.

All of the segmented images presented above have been obtained after 100 MPM samples, each of which is performed by 20 Gibbs sampler iterations.

Remark 2.2. We focus in this paper on recovering $X = x$ from $Y = y$; however, $U = u$ can also have a physical interpretation and its estimation can be of interest. For example, let us imagine that there are two classes, “houses” and “trees.” Otherwise, half of the image $X = x$ is a “town” and the other half is “outside town.” Furthermore, let us assume that there are trees in the town and there are houses outside town. As the distribution of X is different in “town” and “outside town,” these two possibilities will be modeled, according to our model, by two stationarities, λ_1, λ_2 . Therefore, for a given pixel $s \in S$, we can say that x_s is “house” or

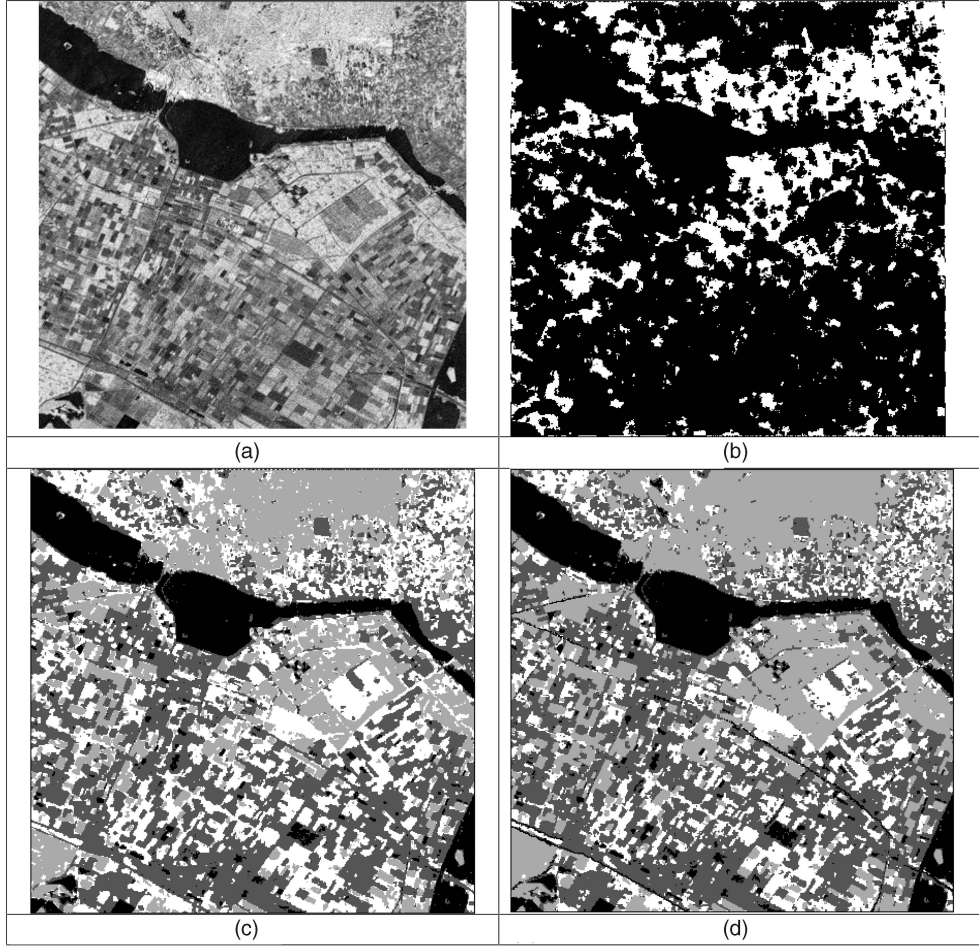


Fig. 4. Real image segmentation (four classes) with the new method and a classical one. (a) Flevoland. (b) Estimated $U = u$. (c) MPM based on HMF. (d) MPM based on TMF.

“trees” and u_s is “town” or “outside town” and both of them are of interest. As $p(u_s|y)$ is computable from $p(x_s, u_s|y)$ by $p(u_s|y) = \sum_{x_s \in \Omega} p(x_s, u_s|y)$, $U = u$ can be estimated from $Y = y$ by MPM. One such example is provided in Fig. 3.

3 NON-GAUSSIAN m -NONSTATIONARY MARKOV FIELD

The aim of this section is to propose a workable extension of the m -MNS field above in which the random variables (Y_s) would no longer be independent conditionally on X , the distributions $p(y_s|x_s)$ would no longer be necessarily Gaussian, and their form could vary with the class x_s . For example, for two classes $\Omega = \{\omega_1, \omega_2\}$, we wish to have two different forms for the distributions $p(y_s|x_s = \omega_1)$ and $p(y_s|x_s = \omega_2)$, as it can occur in real situations [6]. Moreover, we present an original parameter estimation method, which can be seen as a simultaneous extension of both methods proposed in [1] and [5], and present some experiments showing its interest in unsupervised segmentation of nonstationary images.

We consider two successive extensions. In the next section, we consider the stationary case and the associated m -MNS field is introduced in Section 3.2.

3.1 Hidden Markov Fields with Non-Gaussian Correlated Noise

Let us remark that when dealing with the parameter estimation problem, it is useless to consider U . In fact, the problem of parameter estimation is strictly the same for a pairwise Markov field (X, Y) than for a triplet Markov field (X, U, Y) . This comes from the fact that the TMF (X, U, Y) can be seen as a PMF (V, Y) , with $V = (X, U)$. So, we will consider (X, Y) in this section, but it remains valid for (V, Y) , with $V = (X, U)$.

The main idea consists of using a hidden Gaussian Markov field, in which (Y_s) are dependent conditionally on X and Gaussian and transform it to obtain any form of margins. More precisely, the model is inspired from the classical (2.2) model as follows.

We can say that the classical model (2.2) can be obtained in the following way:

1. take a Markov field X with the energy W ;
2. consider the set $(Y'_s)_{s \in S}$ of independent Gaussian variables zero-mean and with variance one;
3. for each $s \in S$, put $Y_s = \sigma_{x_s} Y'_s + m_{x_s}$.

In fact, one can show in a quite classical way that the distribution of (X, Y) , which verifies Points 1 and 2, is exactly of the form (2.2).

Let us assume that there are k classes $\Omega = \{\omega_1, \dots, \omega_k\}$ and we want to have k densities h_1, \dots, h_k , possibly of different forms, for the k distributions $p(y_s|x_s = \omega_1), \dots, p(y_s|x_s = \omega_k)$. Let G be the cumulative distribution function of Gaussian zero-mean variable with variance one and let H_1, \dots, H_k be the cumulative distribution functions corresponding to h_1, \dots, h_k .

Let us consider the following hidden Markov field (X, Y) :

4. take a Markov field X with the energy W ;
5. consider $Y' = (Y'_s)_{s \in S}$ Gaussian Markov random field, where each Y'_s zero-mean and, with variance one,

$$p(y') = \gamma \exp \left[-\frac{1}{2} \left(\sum_{s \in S} (y'_s)^2 + \sum_{(s,t)} -\beta_1 y'_s y'_t \right) \right] \quad (3.1)$$

$$= \gamma \exp[-W'(y')];$$

6. for each $s \in S$, put $Y_s = H_{x_s}^{-1} \circ G(Y'_s)$.

As it is well-known that the distribution of $Y_s = H_{x_s}^{-1} \circ G(Y'_s)$ is given by the cumulative distribution function H_i (or, equivalently, by the density h_i), we see that the model 1-3 above verifies the desired properties: The random variables $(Y_s)_{s \in S}$ are correlated conditionally on X , each $p(y_s|x_s = \omega_i)$ has the desired density h_i , and (X, Y) is a Markov field. We can also see that Points 4, 5, and 6 are a double extension of items 1, 2, and 3; in fact, items 1, 2, and 3 are obtained for $\beta_1 = 0$ in (3.1) and for H_1, \dots, H_k being Gaussian.

Moreover, one can classically show that the distribution of (X, Y) verifying items 4, 5, and 6 are

$$p(x, y) = \gamma \exp \left[-W(x) - W'(\varphi(x, y)) + \sum_{s \in S} \text{Log} \left| \frac{\partial(G^{-1} \circ H_{x_s}(y_s))}{\partial y_s} \right| \right], \quad (3.2)$$

where $\varphi(x, y) = y' = (Y'_s)_{s \in S}$ such that $Y'_s = G^{-1} \circ H_{x_s}(y_s)$ for each $s \in S$.

We see again how (3.2) is a double generalization of (2.2). In fact, when $(Y'_s)_{s \in S}$ are independent, Gaussian, zero-mean and, with variance equal to one, (3.1) becomes $p(y') = \gamma \exp[-\frac{1}{2}(\sum_{s \in S} (Y'_s)^2)]$ and (3.2) becomes

$$p(x, y) = \gamma \exp \left[-W(x) + \sum_{s \in S} \left(-\frac{1}{2} (G^{-1} \circ H_{x_s}(y_s))^2 + \text{Log} \left| \frac{\partial(G^{-1} \circ H_{x_s}(y_s))}{\partial y_s} \right| \right) \right]. \quad (3.3)$$

When, in addition, h_1, \dots, h_k are Gaussian, (3.3) gives (2.2).

3.2 Parameter Estimation

Let us now tackle the problem of items 4, 5, and 6 of model learning, which we propose to solve by a new “generalized” mixture estimation method. The word “generalized” means that, for each class ω_i , the form of h_i is not known, but it is known that this form belongs to a given set of forms $F^i = \{F_1^i, \dots, F_{j(i)}^i\}$. For example, we can know that h_1 is

either Gaussian or Gamma, h_2 is either Gamma or Beta, h_3 is Gaussian, Exponential, or Beta, etc. For each $i = 1, \dots, k$, the families $F_1^i, \dots, F_{j(i)}^i$ are parameterized by $\beta_1^i, \dots, \beta_{j(i)}^i$, respectively. Therefore, estimating the noise in a “generalized” mixture consists of solving a double problem: For each i , find the right family F_r^i in $\{F_1^i, \dots, F_{j(i)}^i\}$ and estimate β_r^i . Different methods of “generalized” mixtures estimation have been proposed in the case of independent (Y'_s) in [5], [15], [23], [27]; thus, the novelty here is to extend these methods to the case of a Markov $Y' = (Y'_s)_{s \in S}$.

Finally, to learn the items 4, 5, and 6 of the model, we have to estimate α defining the Markov distribution of X , β^1 defining the Markov distribution of Y' , find, for each i , the right family F_r^i in $\{F_1^i, \dots, F_{j(i)}^i\}$, and estimate β_r^i . All of this is done using the two following points:

1. Let us imagine that both X and Y are observable. Then α is estimated from $X = x$ by Derin et al.’s method as described in [1]. Otherwise, $Y' = y'$ can be calculated from $(X, Y) = (x, y)$ by $Y'_s = G^{-1} \circ H_{x_s}(y_s)$ (for each $s \in S$) and, thus, β^1 can be estimated from $Y' = y'$ by the method described in [1]. Finally, there are numerous methods enabling one to find, for each $i = 1, \dots, k$, the right family F_r^i in $\{F_1^i, \dots, F_{j(i)}^i\}$ and to estimate β_r^i . One possible general method is based on the Kolmogorov distance, as described in [15]. Another method, less general but simpler and faster, is based on the Pearson system. As it will be used in experiments in the next section, let us briefly recall its principle. We assume that all families F^i are equal and contain the eight distributions of the Pearson system. This system is the set of pdf f on R verifying

$$\frac{df(x)}{dx} = -\frac{x + a}{c_0 + c_1 x + c_2 x^2} f(x) \quad (3.4)$$

and contains some current distributions like Gaussian, Gamma, inverse Gamma, Beta of the first kind, and Beta of the second kind distributions. The first four moments of $p(y_s|x_s = \omega_i)$ give, by analytic formulas, both the right family F_r^i in F^i and the parameters β_r^i . The exact calculations, which can be seen in [15], are rather complicated and, thus, are not recalled here. It is important that the right family F_r^i in F^i and the parameters β_r^i can be estimated, for each $i = 1, \dots, k$, from $(X, Y) = (x, y)$. In fact, having $X = x$, let $S_i = \{s \in S | x_s = \omega_i\}$. As all $(y_s)_{s \in S_i}$ are realizations of Y_s according to $p(y_s|x_s = \omega_i)$, we can use them to estimate the first four moments, which gives F_r^i and the parameters β_r^i .

2. As X is not observable, we have to use the sole $Y = y$. We propose the following iterative method, which is an extension of the Iterative Conditional Estimation (ICE) based method used in [1]. The proposed method runs as follows:

- a. Initialize the searched parameters and the densities associated with the classes (which can be assumed, for example, to be Gaussian) by some simple method, specific to a given application.
- b. At each iteration q , the next parameters and densities associated with the classes are obtained

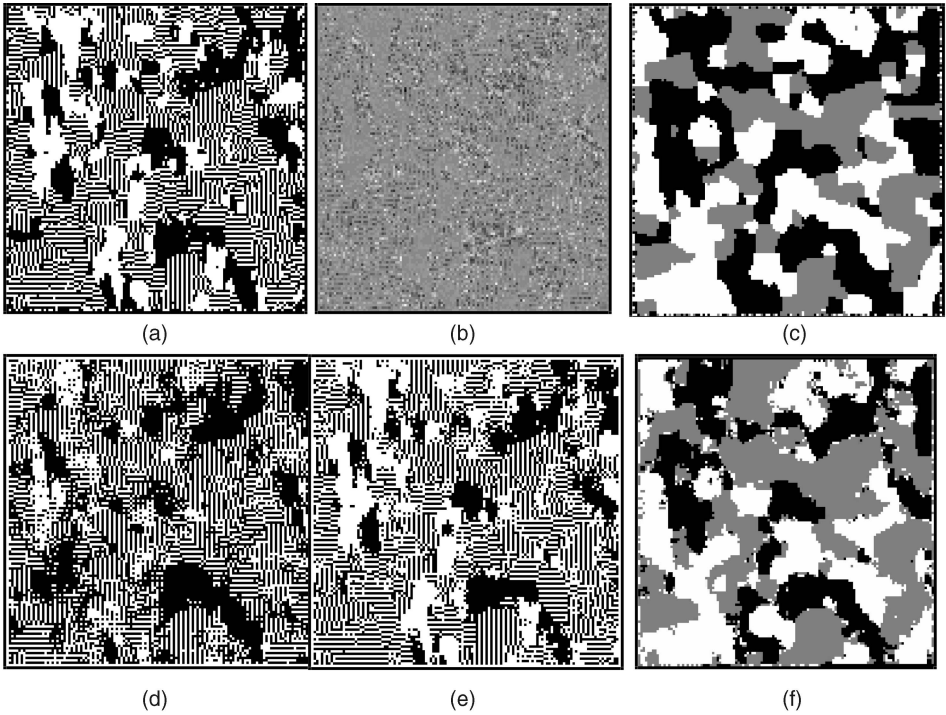


Fig. 5. Unsupervised segmentation using (d) Gaussian laws and (e) Beta and Gamma laws. (a) Class Image $X = x$. (b) Noisy Image $Y = y$. (c) $U = u$ field. (d) Gaussian error ratio = 19.14 percent. (e) Non-Gaussian error ratio = 8.81 percent. (f) estimated u .

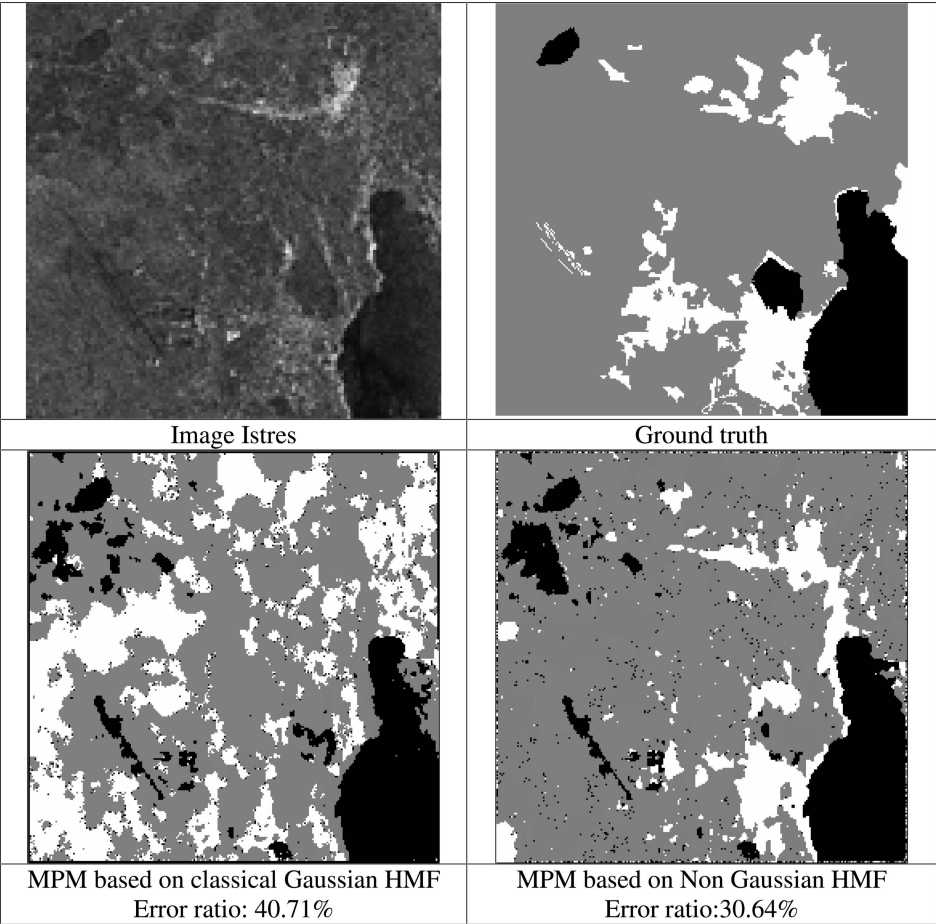


Fig. 6. Unsupervised segmentation results obtained with “Gaussian” and “non-Gaussian” (based on the Pearson system), classical HMF, and the new TMF-based unsupervised segmentation.

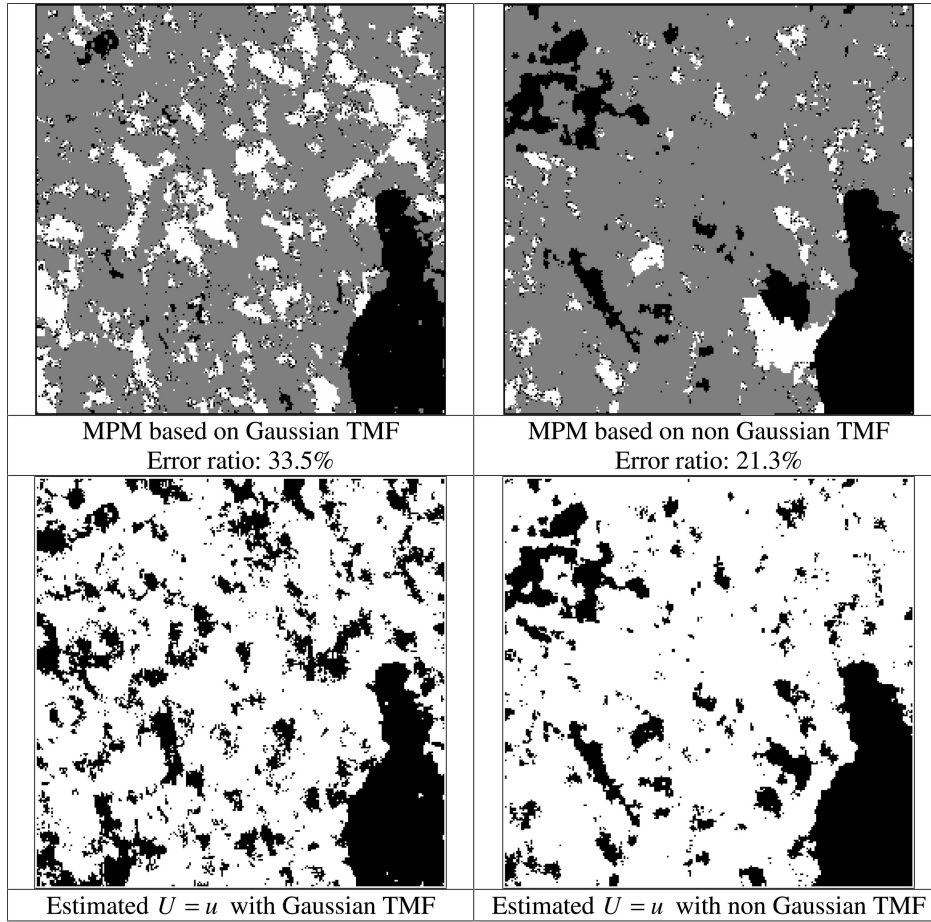


Fig. 6. (continued).

from $Y = y$ and current parameters and densities in the following way:

- Simulate x^{q+1} according to $p(x|y)$ based on the current parameters.
- Use x^{q+1} to estimate α , which gives the next α^{q+1} .
- For each $i=1, \dots, k$, put $S_i^{q+1} = \{s \in S / x_i^{q+1} = \omega_i\}$ and use $(y_s^{q+1})_{s \in S_i^{q+1}}$ to calculate the four moments of $p(y_s | x_s = \omega_i)$. Use the Pearson system to find $F_r^{i,q+1}$ and calculate the parameters $\beta_r^{i,q+1}$.
- Use the densities $h_i^{q+1}(y_s) = p(y_s | x_s = \omega_i)$ to calculate the cumulative functions H_i^{q+1} and $y_s^{q+1\prime} = G^{-1} \circ H_i^{q+1}(y_s)$. Use $y^{q+1\prime} = (y_s^{q+1\prime})_{s \in S}$ to estimate β^1 , which gives the next $\beta^{1,q+1}$.
- c. Stop the procedure when the estimates become steady, according to some criterion specific to a given application dealt with.

Let us remark that the use of the Pearson system has already been used in different hidden Markov models and gave satisfactory results [6], [8], [9], [10]. Therefore, the originality of the present paper lies in the fact that we take spatial correlations into account, which generalize the previous models. Its use is also quite practical because the four first moments give the form and the parameters of the searched density; however, in the general case, any families $F_1^i, \dots, F_{j(i)}^i$ (for $i = 1, \dots, k$) can be used. The calculation of

the four first moments is replaced by some other method, allowing one to chose $F_r^{i,q+1}$ from $S_i^{q+1} = \{s \in S / x_i^{q+1} = \omega_i\}$ and $(y_s^{q+1})_{s \in S_i^{q+1}}$. For example, $F_r^{i,q+1}$ can be chosen by minimizing the Kolmogorov distance to the histogram obtained from S_i^{q+1} , as described in [15].

Let us notice that the ICE method, first proposed in [33], is more flexible and may be better suited to the framework we are in than the well-known EM method [24]; in fact, ICE does not use the likelihood, which is difficult to handle in Markov field context (some relationships between ICE and EM can be seen in [7]). Other learning methods can be seen in [18], [30].

We do not present any experiments here as a more complex case, including nonstationary images, will be studied in the next section.

3.3 Parameter Estimation in m -MNS with Non-Gaussian Correlated Noise

Passing from the model (3.2) above to an m -MNS is similar to the method described in Section 2, which extends (2.2) to (2.4). Therefore, let us consider the following model, similar to the model items 4, 5, and 6 defined by (3.2), except item 4 which becomes: Item 4', take a Markov field (X, U) with the energy W .

Then, the distribution of the m -MNS related to items 4', 5, and 6 is

TABLE 1
Parameter Estimation and Unsupervised Segmentation Results Obtained with Classical HMF
(in Both the Gaussian and “Pearson” Cases) and New “Nonstationary” TMF (in Both the Gaussian and “Pearson” Cases)

Parameters	HMF		NS TMF	
	Gaussian	Pearson	Gaussian	Pearson
α	(0.55,0.76)	(0.55,0.76)	(0.58,0.68) (0.53,0.55) (0.63,0.69)	(0.58,0.68) (0.53,0.55) (0.63,0.69)
q				-0.08
ω_1 distribution's	$N_1(32.01,340.57)$	$B(4.83,5.80)$	$N_1(37.01,320.41)$	$B(5.13,6.10)$
ω_2 distribution's	$N_2(68.61,548.74)$	$\Gamma(10.7,5.50)$	$N_2(70.62,538.24)$	$\Gamma(11.37,6.10)$
ω_3 distribution's	$N_3(86.79,1827.16)$	$B(0.92,2.00)$	$N_3(89.46,1820.72)$	$B(1.01,2.10)$
Error ratio	40.71%	30.64%	33.50%	21.30%

$$p(x, u, y) = \gamma \exp \left[-W(x, u) - W'(\varphi(x, y)) \right. \\ \left. + \sum_{s \in S} \text{Log} \left| \frac{\partial(G^{-1} \circ H_{x_s}(y_s))}{\partial y_s} \right| \right]. \quad (3.5)$$

The extension of the parameter estimation procedure above to the m -MNS defined by (3.5) does not pose a particular problem. The difference is that α are now the parameters defining the Markov field (X, U) , instead of the Markov field X in the previous section. Thus, the two first items in item 2 are replaced by:

- simulate (x^{q+1}, u^{q+1}) according to $p(x, u|y)$ based on the current parameters;
- use (x^{q+1}, u^{q+1}) to estimate α , which gives the next α^{q+1} .

Remark 3.1. We assumed in (3.5) that $p(y_s|x_s, u_s) = p(y_s|x_s)$, but the general case in which $p(y_s|x_s, u_s)$ depends on both x_s and u_s does not pose particular problems. For k classes, there would be for an associated m -MNS, $k \times m$ distributions $p(y_s|x_s, u_s)$. This means that the third field U can model different stationarities of X , different stationarities of Y , or even, in the general case, different stationarities of both (X, Y) . Moreover, the dependence of $p(y_s|x_s, u_s)$, which depends on u_s , can have an intuitive justification: In the example in Remark 2.2, one can imagine that the aspect of trees is not the same in town and outside town. Another example could be the presence of clouds producing two kinds of illumination and, thus, two kinds of stationarities of $p(y_s|x_s, u_s)$.

3.4 Experiments

We present below two experiments: simulated image and real image segmentation. In the case of simulated image, we consider two classes and three different stationarities. In the case of real image, we consider three classes and two different stationarities.

The aim of the first experiment is to show that there exist situations in which the use of correct margins instead of incorrect Gaussian margin can improve the unsupervised segmentation accuracy. To be more precise, we simulate a nonstationary correlated hidden Markov field as described above such as $p(y_s|x_s = \omega_1) = \Gamma(1, 2)$ and $p(y_s|x_s = \omega_2) = \beta(2, 1)$. Then, we segment the noisy image

using the MPM method in two ways. On the one hand, we suppose that the margins of each class are Gaussian. On the other hand, we use the Pearson system to find the nature of the class margins and use them rather than Gaussian margins. A comparison between these two approaches is depicted in Fig. 5.

The real image considered, shown in Fig. 6, is a radar one. It is an image of Istres city, France, which can be divided into three classes: 1) Water (in black), 2) Forest and cultivation (in gray), and 3) Houses (in white). We dispose of the ground truth, also presented in Fig. 6.

Our aim is twofold:

1. Is there any interest in using the Pearson system instead of classical Gaussian distributions? According

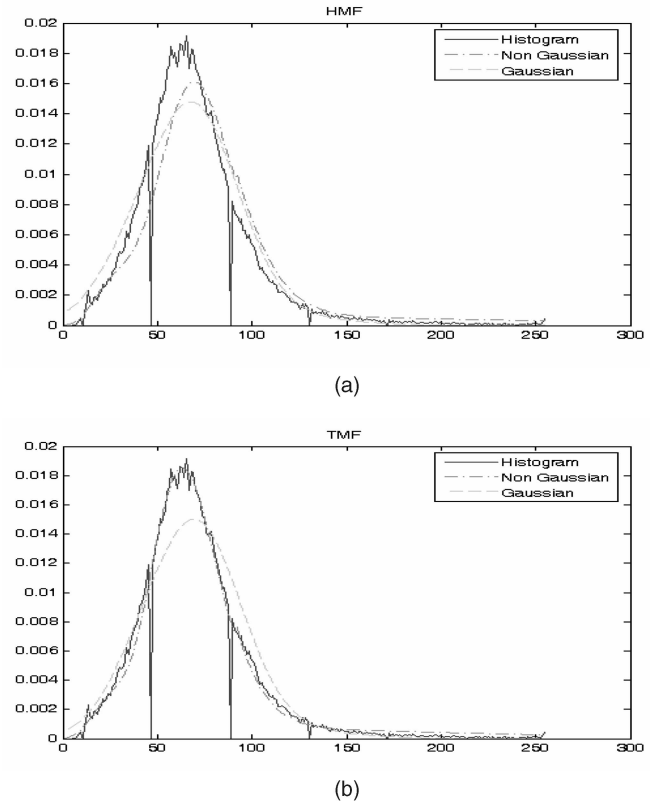


Fig. 7. Histogram of the image Istres compared to Gaussian and Pearson estimates in (a) the HMF case and (b) the TMF case.

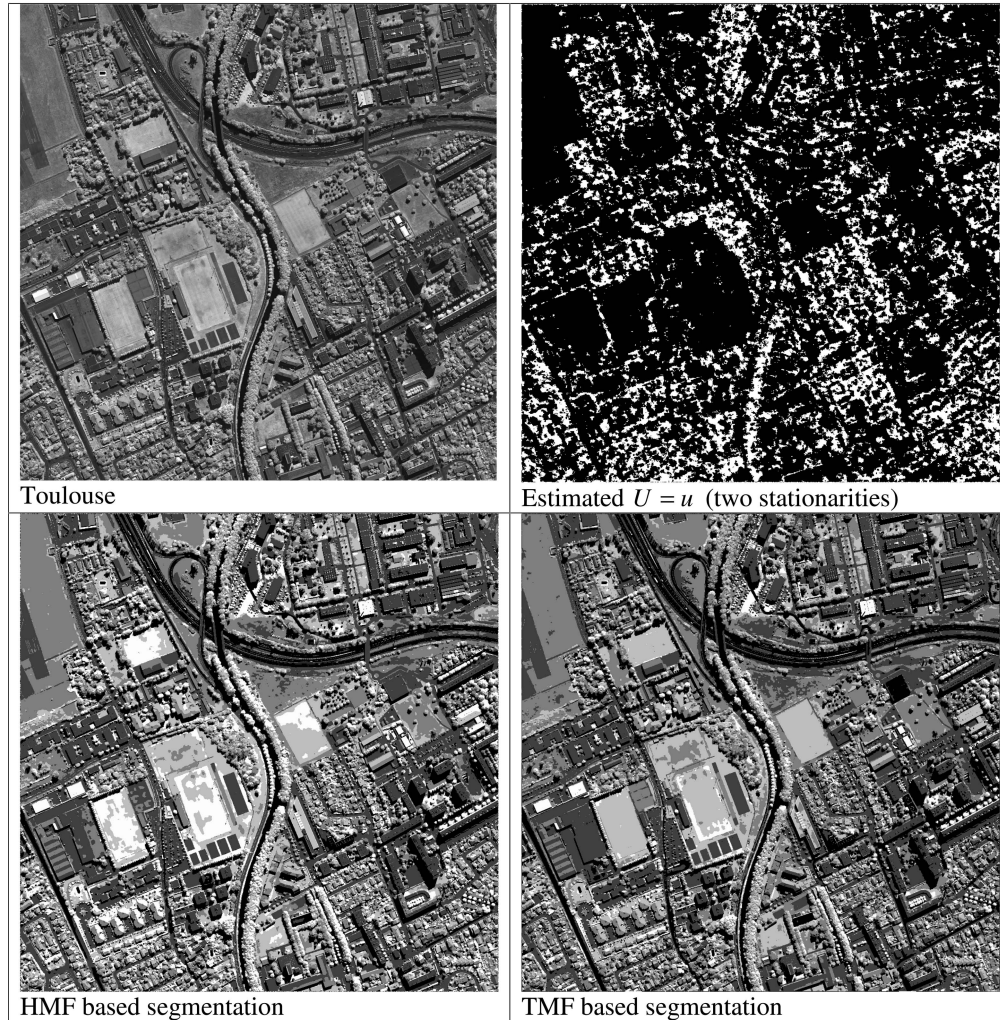


Fig. 8. Real airborne image of the city of Toulouse, estimated $U = u$, and HMF and TMF-based unsupervised segmentation results.

to Table 1 and Fig. 6, we see that using Pearson system improves the unsupervised segmentation results in the classical HMF-based case and in the new TMF case as well.

2. Is there any interest in using the new TMF instead of classical HMF? According to Table 1 and Fig. 6, we see that using the new TMF improves the unsupervised segmentation results in the classical Gaussian case and in the "Pearson system"-based case as well.

Finally, we see that simultaneously using the new TMF and the Pearson system gives the best results and the difference from those obtained with the classical Gaussian HMF is quite significant.

Finally, let us specify that the Gaussian triplet model used is the model given by (2.4), while the general triplet model is given by (3.5). In the former case, the parameters are estimated by the classical ICE, while, in the latter case, they are estimated by the "generalized" ICE above, based on the Pearson system. The estimates of the parameters obtained by these two methods are presented in Table 1. The unsupervised segmentation results obtained by both the "Gaussian" and "Pearson" methods are presented in Fig. 6 and we can see that the latter significantly improves the result: The error ratio passes from 33.5 percent to 21.3 percent. We also show the

estimated realizations of U , which models two different stationarities considered.

We have used 30 iterations in ICE and the posterior marginal distributions, used in the MPM segmentation, are estimated from 100 samples, each of which was sampled with 20 Gibbs sampler iterations. Concerning the computation time, unsupervised TMF-based segmentation takes about 50 seconds in the Gaussian case and about 350 seconds in the non-Gaussian one.

Let us notice that the error ratio varies very little when performing numerous times the same experiments. We performed the four unsupervised segmentations in Table 1 10 times and the means of these ratios for HMF Gaussian, HMF Pearson, TMF Gaussian, and TMF Pearson were 40.60 percent, 30.61 percent, 33.68 percent, and 21.28 percent, respectively, while the variance was about 0.01 for the four series of experiments.

Remark 3.2. Let us notice that one possible interpretation of different values of U can be the presence of different textures. This is particularly striking in the examples given in Fig. 3 and Fig. 5. In such cases, searching U can play a role comparable to searching X and, finally, the problem can be to find both X and U . Otherwise, different kinds of nonstationarities can be simultaneously present in a given

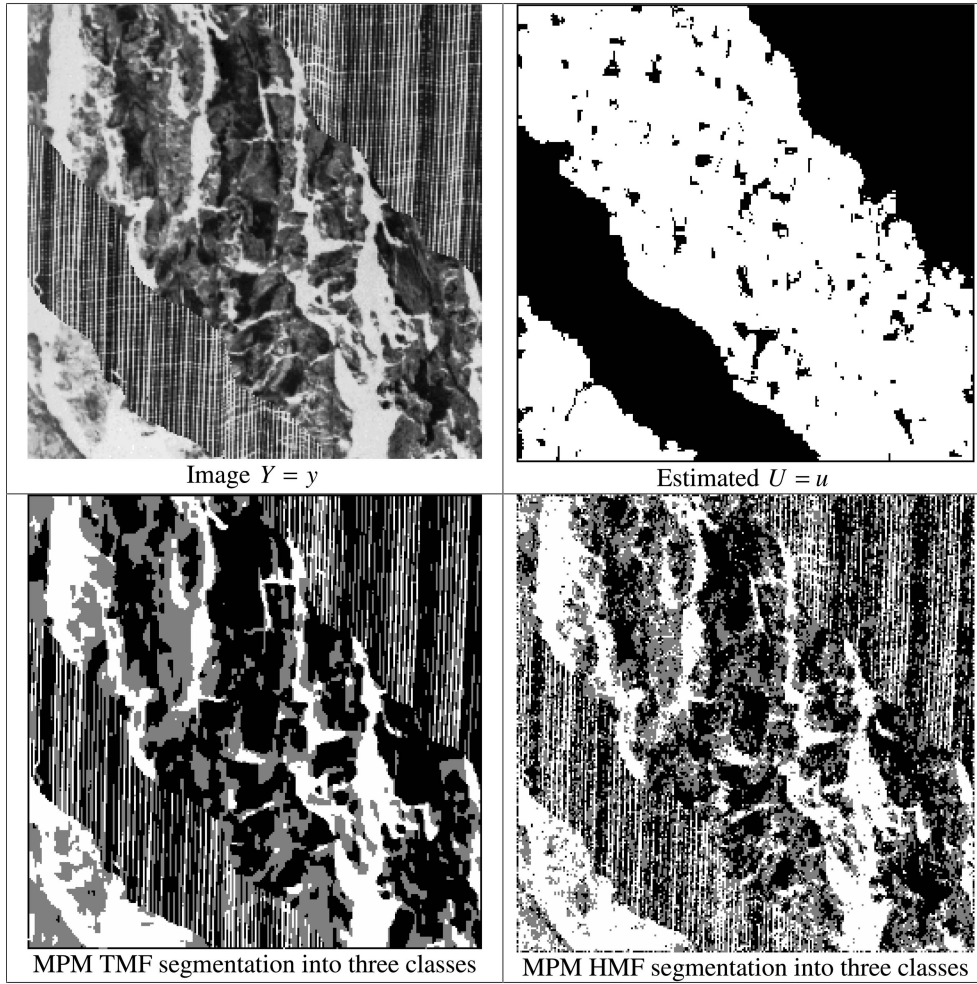


Fig. 9. Collage of Brodatz textures $Y = y$, estimated $U = u$, and TMF-based segmentation into three classes and the classical HMF one.

image and, thus, simultaneously modeled by numerous fields U_1, \dots, U_r . For example, we can have $U = (U_1, U_2)$, where U_1 models different textures and U_2 models the presence of clouds, as described in Remark 3.1 above.

Concerning the computation time, unsupervised TMF-based segmentation takes about 390 seconds in the Gaussian case and about 1,350 seconds in the non-Gaussian one. While using HMF, these times are about 350 and 1,060 seconds, respectively.

According to Fig. 7, we see that the estimated mixture based on Gaussian estimates encounters some difficulties in fitting the histogram and this is true in both the HMF and TMF cases. Using Pearson system gives a very good fit in the TMF case, while it slightly improves the Gaussian fit in the HMF case.

In Fig. 8, we present another example of segmentation of a real airborne image of the city of "Toulouse" into five classes. In the new TMF model, we consider two different stationarities and the estimated $U = u$ seems to indicate two kinds of areas really corresponding to two visually different stationarities. The differences between TMF and HMF-based segmentation results are not easy to clearly differentiate; however, some details, like the four bright squares, seem to be better represented when using TMF.

Finally, let us remark that the random field U can possibly be more interesting than the random field X . We

present in Fig. 9 an image $Y = y$ obtained by the collage of two Brodatz textures. On the one hand, the estimated $U = u$ can be seen as a textures classification result. On the other hand, the three classes segmentation provided by TMF is visually closer to the real image than the segmentation obtained with the classical HMF.

4 CONCLUSIONS

We proposed in this paper some extensions of the classical hidden Markov field model and some original related parameter estimation methods. The most general extension allows one to deal with nonstationary images and possibly non-Gaussian correlated noise. Moreover, the form of the noise corresponding to different classes can vary with the underlying class. Such a general model can then be learned by a new method proposed in the paper which is based on the Pearson system and the general Iterative Conditional Estimation (ICE) algorithm. The interest in the unsupervised statistical segmentation based on both new models and parameter estimation method is attested to via different experiments conducted on synthetic and real images.

As perspectives, let us make two points. We can notice that the proposed model can be seen as a particular case of a Gaussian copula adapted to a Markov field. In this context, we can thus envisage the adaptation of other kinds of copulas, whose general theory can be seen in [29], to the Markov field

models considered in this paper. In fact, the recent use of copulas in Markov chains [4] or multiband Markov trees [12] has provided encouraging results. Otherwise, multispectral or hyperspectral imagery, as in [28], among others, also offers further possibilities for development of the models and related processing proposed in this paper.

REFERENCES

- [1] D. Benboudjema and W. Pieczynski, "Unsupervised Image Segmentation Using Triplet Markov Fields," *Computer Vision and Image Understanding*, vol. 99, no. 3, pp. 476-498, 2005.
- [2] D. Benboudjema and W. Pieczynski, "Segmenting Non Stationary Images with Triplet Markov Fields," *Proc. IEEE Int'l Conf. Image Processing*, Sept. 2005.
- [3] D. Benboudjema and W. Pieczynski, "Unsupervised Segmentation of Non Stationary Images with Non Gaussian Correlated Noise Using Triplet Markov Fields and the Pearson System," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing*, May 2006.
- [4] N. Brunel and W. Pieczynski, "Unsupervised Signal Restoration Using Hidden Markov Chains with Copulas," *Signal Processing*, vol. 85, no. 12, pp. 2304-2315, 2005.
- [5] Y. Delignon, A. Marzouki, and W. Pieczynski, "Estimation of Generalized Mixture and Its Application in Image Segmentation," *IEEE Trans. Image Processing*, vol. 6, no. 10, pp. 1364-1375, 1997.
- [6] Y. Delignon and W. Pieczynski, "Modeling Non Rayleigh Speckle Distribution in SAR Images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1430-1435, 2002.
- [7] J.-P. Delmas, "An Equivalence of the EM and ICE Algorithm for Exponential Family," *IEEE Trans. Signal Processing*, vol. 45, no. 10, pp. 2613-2615, 1997.
- [8] S. Derrode, G. Mercier, J.-M. Le Caillec, and R. Garello, "Estimation of Sea-Ice SAR Clutter Statistics from Pearson's System of Distributions," *Proc. Int'l Geoscience and Remote Sensing Symp.*, 2001.
- [9] S. Derrode and W. Pieczynski, "Signal and Image Segmentation Using Pairwise Markov Chains," *IEEE Trans. Signal Processing*, vol. 52, no. 9, pp. 2477-2489, 2004.
- [10] S. Derrode and G. Mercier, "Unsupervised Multiscale Oil Slick Segmentation from SAR Images Using a Vector HMC Model," *Pattern Recognition*, vol. 40, no. 3, pp. 1135-1147, 2007.
- [11] F. Destremes and M. Mignotte, "A Statistical Model for Contours in Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 626-238, May 2004.
- [12] F. Flitti, C. Collet, and A. Joannic-Chardin, "Unsupervised Multiband Image Segmentation Using Hidden Markov Quadtree and Copulas," *Proc. Int'l Conf. Image Processing*, Sept. 2005.
- [13] S. Foucher, G.B. Benie, and J.-M. Boucher, "Multiscale MAP Filtering of SAR images," *IEEE Trans. Image Processing*, vol. 10, no. 1, pp. 49-60, 2001.
- [14] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.
- [15] N. Giordana and W. Pieczynski, "Estimation of Generalized Multisensor Hidden Markov Chains and Unsupervised Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 465-475, May 1997.
- [16] X. Guyon, "Random Fields on a Network," *Probability and Its Applications*, Springer-Verlag, 1995.
- [17] L. Hellequin, J.-M. Boucher, and X. Lurton, "Processing of High-Frequency Multibeam Echo Sounder Data for Seafloor Characterization," *IEEE J. Oceanic Eng.*, vol. 28, no. 1, pp. 78-89, 2003.
- [18] M.V. Ibáñez and A. Simó, "Parameter Estimation in Markov Random Field Image Modeling with Imperfect Observations, A Comparative Study," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2377-2389, 2003.
- [19] P. Lanchantin and W. Pieczynski, "Unsupervised Non Stationary Image Segmentation Using Triplet Markov Chains," *Proc. Advanced Concepts for Intelligent Vision Systems*, 2004.
- [20] P. Lombardo and A. Farina, "Coherent Radar Detection against K-Distributed Clutter with Partially Correlated Texture," *Signal Processing*, vol. 48, pp. 1-15, 1996.
- [21] A. Maffett and C. Wackermann, "The Modified Beta Density Function as a Model for Synthetic Aperture Radar Clutter Statistics," *IEEE Trans. Geoscience and Remote Sensing*, vol. 29, no. 2, pp. 277-283, 1991.
- [22] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic Solution of Ill-Posed Problems in Computational Vision," *J. Am. Statistical Assoc.*, vol. 82, pp. 76-89, 1987.
- [23] E. Monfrini and W. Pieczynski, "Estimation de Mélanges Généralisés dans les Arbres de Markov Cachés, Application à la Segmentation des Images de Cartons d'Orgue de Barbarie," *Traitement du Signal*, vol. 22, no. 2, pp. 135-147, 2005.
- [24] G.J. McLachlan and T. Krishnan, *EM Algorithm and Extensions*. Wiley, 1997.
- [25] D. Melas and S.P. Wilson, "Double Markov Random Fields and Bayesian Image Segmentation," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 357-365, 2002.
- [26] M. Mignotte, Ch. Collet, P. Pérez, and P. Bouthemy, "Three-Class Markovian Segmentation of High Resolution Sonar Images," *Computer Vision and Image Understanding*, vol. 76, no. 3, pp. 191-204, 1999.
- [27] G. Moser, J. Zerubia, and S.B. Serpico, "Dictionary Based Stochastic Expectation-Maximization for SAR Amplitude Probability Density Function Estimation," *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 188-200, 2006.
- [28] R. Neher and A. Sristava, "A Bayesian Framework for Labelling Terrain Using Hyperspectral Imaging," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1363-1374, 2005.
- [29] R.B. Nelsen, "An Introduction to Copulas," *Lecture Notes in Statistics*, vol. 139, 1998.
- [30] P. Pérez, "Markov Random Fields and Images," *CWI Quarterly*, vol. 11, no. 4, pp. 413-437, 1998.
- [31] G. Poggi, G. Scarpa, and J. Zerubia, "Supervised Segmentation of Remote Sensing Images Based on a Tree-Structured MRF Model," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 8, pp. 1901-1911, 2005.
- [32] W. Pieczynski, "Modèles de Markov en Traitement d'Images," *Traitement du Signal*, vol. 20, no. 3, pp. 255-278, 2003.
- [33] W. Pieczynski, "Statistical Image Segmentation," *Machine Graphics and Vision*, vol. 1, nos. 1/2, pp. 261-268, 1992.
- [34] W. Pieczynski and A.-N. Tebbache, "Pairwise Markov Random Fields and Segmentation of Textured Images," *Machine Graphics & Vision*, vol. 9, no. 3, pp. 705-718, 2000.
- [35] W. Pieczynski and D. Benboudjema, "Multisensor Triplet Markov Fields and Theory of Evidence," *Image and Vision Computing*, vol. 24, no. 1, pp. 61-69, 2006.
- [36] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre, "A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR Images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2046-2057, 2004.
- [37] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction*. Springer, 2003.



evidence in the Markov field context.



Dalila Benboudjema received the PhD degree in optimization and systems security from the Institut National des Télécommunications d'Evry in a joint program with Compiègne University of Technology, France, in 2005. She is currently a postdoctoral researcher in Signal and Image Processing (TSI) Department at the Ecole Nationale Supérieure des Télécommunications, Paris, France. Her research interests include statistical image processing and theory of

Wojciech Pieczynski received the Doctorat d'Etat degree in mathematical statistics from the Université Pierre et Marie Curie, Paris, France, in 1986. He is currently a professor at the Institut National des Télécommunications and head of the Communications, Image, and Information Processing (CITI) Department. His research interests include mathematical statistics, stochastic processes, and statistical image processing.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.