Pairwise Markov random fields and segmentation of textured images

Wojciech Pieczynski, Abdel-Nasser Tebbache
Département Signal et Image
Institut National des Télécommunications
9, rue Charles Fourier, 91000 Evry, France
Wojciech.Pieczynski@int-evry.fr

Abstract. The use of random fields, which allows one to take into account the spatial interaction among random variables in complex systems, becomes a frequent tool in numerous problems of statistical mechanics, spatial statistics, neural network modelling, and others. In particular, Markov random field based techniques can be of exceptional efficiency in some image processing problems, like segmentation or edge detection. In statistical image segmentation, that we address in this work, the model is generally defined by the probability distribution of the class field, which is assumed to be a Markov field, and the probability distributions of the observations field conditional to the class field. Under some hypotheses, the a posteriori distribution of the class field, i.e. conditional to the observations field, is still a Markov distribution and the latter property allows one to apply different bayesian methods of segmentation like Maximum a Posteriori (MAP) or Maximum of Posterior Mode (MPM). However, in such models the segmentation of textured images is difficult to perform and one has to resort to some model approximations. The originality of our contribution is to consider the markovianity of the couple (class field, observations field). We obtain a different model; in particular, the class field is not necessarily a Markov field. However, the posterior distribution of the class field is a Markov distribution, which makes possible bayesian MAP and MPM segmentations. Furthermore, the model proposed makes possible textured image segmentation with no approximations.

Key words: hidden Markov fields, pairwise Markov fields, bayesian image segmentation, textured images.

1. Introduction

The aim of this paper is to propose a novel Markov Random Field model and describe some situations in which its use is more relevant that the use of the classic, Hidden Markov Random Field (HMRF) models. The main difference with the classical HMRF models ([2, 3, 5, 11–13]) is that the prior distribution of the hidden class field is not necessarily markovian. In the context of textured images segmentation, the advantage of the new model with respect to the classical HMRF model is that textured images can be segmented without any model approximation. More precisely, for a set of pixels S one classically considers two random fields $X = (X_s)_{s \in S}$ and $Y = (Y_s)_{s \in S}$. The unobserved field X is markovian, and one observes the field Y, which can be seen as a "noisy version" of X. So X is "hidden" and such a model is often called a "Hidden Markov Random Field" model. Estimating X from $Y = y$ requires being able to simulate realisations of X according to its distribution conditioned on $Y = y$ (its posterior distribution), which is possible when the latter distribution is markovian. In the pioneering papers [2, 3, 5] addressing this problem, the markovianity of the posterior distribution of X is ensured by the hypothesis according to which the random variables $(Y_s)_{s \in S}$ are independent conditionally to X. The drawback is that such a hypothesis is rather strong and, in particular, does not allow one to take into account possible textures of the considered classes. This problem has been described by Derin et al., who proposed a genuine "hierarchical" Markov model to solve it [4, 6, 7, 10]. Although the model proposed is well adapted to numerous situations, the posterior distribution of X is not markovian, and thus using bayesian estimations of X necessitates some model approximations. To avoid this, we propose in this paper a new model, briefly mentioned in [21], that we call a "Pairwise Markov Random Field" model (PM RF), consisting of assuming that the couple $Z = (X, Y)$ is markovian. This is not necessarily an HMRF because X is not necessarily markovian. So neither X nor Y are markovian in a PMRF, but they are simulable ($Z = (X, Y)$ is simulable). Furthermore, the distribution of Y conditional to X is markovian, and the distribution of X conditional to Y is markovian too. The first allows one to model textures, and the second allows one to apply bayesian MAP and MPM segmentations. As a consequence, textured images - which, moreover, can be corrupted by potentially correlated noise - can be segmented without any model approximation.

The organisation of the paper is as follows. The next section is devoted to the classical Hidden Markov Random Field model, in which the difficulty of modelling textured images is brought forth. The new PMRF model is introduced in the third section and some calculations in the gaussian case are specified. The fourth section is devoted to the presentation of some samplings of PMRF, and some segmentations using MPM of synthetic images so obtained. The fifth section contains some conclusions and perspectives.

2. Classical Hidden Markov Random Field (HMRF) model

2.1. General framework

Let S be the set of pixels, with $N = \mathrm{Card}(S)$, and let $X = (X_s)_{s \in S}$, $Y = (Y_s)_{s \in S}$ be two random fields. Each X_s takes its values in a finite set of classes $\Omega = \{\omega_1, \omega_2\}$ and each Y_s takes its values in $\mathbb{R}$. The field X is unobserved and the problem is to estimate his realizations from the observed field Y. Before looking for estimation methods, one has to define the distribution of (X, Y). The classical way of defining this distribution is to define the distribution of X and the distributions of Y conditional on X. In the HMRF context considered here, the field X a Markov one. In a general manner, X is said to be markovian with respect to a neighbourhood V_s, whose form is independent of the position of s in S and will be denoted by V, if its distribution can be written as:

$$P(X = x) = \gamma \exp(-U(x)) \tag{1}$$

with

$$U(x) = \sum_{e \in E} \Psi_e(x_e), \tag{2}$$

where E is the set of cliques (a clique being a subset of S which is either a singleton or a set of pixels mutually neighbours with respect to V), x_e the restriction of x to e, and Ψ_e a function, which depends on e and which takes its values in $\mathbb{R}$. In order to define the distributions of $Y = (Y_s)_{s \in S}$ conditional on $X = (X_s)_{s \in S}$, one classically assumes that the two following conditions hold:

(i) the random variables (Y_s) are independent conditional to X;

(ii) the distribution of each Y_s conditional to X is its distribution conditional to X_s. Due to theses hypotheses, all the distributions of Y conditional to X are defined, for k classes, by k distributions on $\mathbb{R}$. To be more precise, let f_i denote the density, with respect to the Lebesgue measure on $\mathbb{R}$, of the distribution of Y_s conditional to $X_s = \omega_i$. Then we have

$$P(Y = y \mid X = x) = \prod_s f_{x_s}(y_s), \tag{3}$$

where f_{x_s} is the density of the distribution of Y_s conditional to $X_s = x_s$. So, given that $P(X = x, Y = y) = P(Y = y \mid X = x)P(X = x)$ and putting $f_{x_s}(y_s) = \exp\left(\log(f_{x_s}(y_s))\right)$, we have

$$P(X = x, Y = y) = \gamma \exp - \left(\sum_{e \in E} \Psi_e(x_e) - \sum_{s \in S} \log f_{x_s}(y_s) \right). \tag{4}$$

Let us notice that as X is discrete and Y is continuous, the function (1) is a probability density with respect to the counting measure on Ω^N, the functions (3) are probability densities with respect to the Lebesgue measure on $\mathbb{R}^N$, and so the function (4) is a probability density with respect to the product of the counting measure on Ω^N by the Lebesgue measure on $\mathbb{R}^N$. As such, the notation $P(X = x, Y = y)$ is slightly abusive, because it is not a probability; however, once admitted, it significantly simplifies subsequent expressions. So, according to (4) the couple (X, Y) is then a markovian field, and the distribution of X conditional to $Y = y$ is still a distribution of a Markov field. This allows one to simulate realizations of X according to its posterior distribution and thus apply different segmentation methods like Maximum A Posteriori (MAP, [2,3]):

$$\hat{s}_{MAP}(y) = \arg\max_{x \in \Omega^N} \; P(X = x \mid Y = y) \tag{5}$$

or Maximum Posterior Mode (MPM [5]) :

$$\forall s \in S, \; \hat{s}_{MPM}(s, y) = \arg\max_{x_s \in \Omega} \; P(X_s = x_s \mid Y = y). \tag{6}$$

Remark 2.1. MAP and MPM are two bayesian methods, corresponding to two cost functions. A cost function $L : \Omega^N \times \Omega^N \to \mathbb{R}^+$ models the gravity of errors: $L(x_1, \widehat{x})$ is the cost of assuming that the value of x is $\widehat{x}$, when the real value is x_1. A bayesian strategy is then a strategy which minimises, with respect to $\widehat{s} : \mathbb{R}^N \to \Omega^N$, the expectation $E((L(X, \widehat{s}(Y)))$. In other words, a bayesian strategy minimises, in the long run, the average of costs of errors made. The bayesian strategy MAP corresponds then to $L_1(x, \widehat{x}) = 1_{[x \neq \widehat{x}]}$, and the bayesian strategy MPM corresponds to $L_1(x, \widehat{x}) = \sum_{s \in S} 1_{[x_s \neq \widehat{x}_s]}$. The success of the hidden Markov fields model in image segmentation is due mainly to the fact that MAP and MPM are not computable in the general case, but they are computable, via approximations, when using the hidden Markov random field model.

2.2. Gaussian model

Let us consider a simple gaussian case. The random field $X = (X_s)_{s \in S}$ is a Markov field with respect to four nearest neighbours and each X_s takes its values in the set of two classes $\Omega = \{\omega_1, \omega_2\}$. Denoting by $t \leftrightarrow u$ the fact that the pixels t and u are neighbours, we have:

$$U(x) = \sum_{t \leftrightarrow u} \varphi_1(x_t, x_u) + \sum_{s \in S} \varphi_2(x_s) \tag{7}$$

and

$$f_{x_s}(y_s) = \frac{1}{\sqrt{2\pi\sigma_{x_s}^2}} \exp\left(-\frac{(y_s - m_{x_s})^2}{2\sigma_{x_s}^2}\right) \tag{8}$$

which gives

$$P(Y = y \mid X = x) = \exp\left(\sum_{s \in S} -\frac{1}{2}\log(\pi\sigma_{x_s}^2) - \frac{(y_s - m_{x_s})^2}{2\sigma_{x_s}^2}\right) \tag{9}$$

and

$$P(X = x, Y = y) = \gamma \exp\left(-\sum_{s \in S}(\varphi_2(x_s) - \frac{1}{2}\log(\pi\sigma_{x_s}^2)) - \frac{(y_s - m_{x_s})^2}{2\sigma_{x_s}^2} - \sum_{t \leftrightarrow u} \varphi_1(x_t, x_u)\right), \tag{10}$$

where φ_1 is any function from Ω^2 to $\mathbb{R}^2$, and φ_2 is any function from Ω to $\mathbb{R}$. In practice, the random variables (Y_s) are not, in general, independent conditionally on X. In particular, (9) is too simple to allow one to take texture into account. For instance, if we consider that texture is a gaussian Markov random field realization ([1, 12, 13]), (9) should be replaced with:

$$P(Y = y \mid X = x) = \gamma(x) \exp\left(-\sum_{t \leftrightarrow u} a_{x_t x_u} y_t y_u - \sum_{s \in S}(a_{x_s x_s} y_s^2 + b_{x_s} y_s)\right) \tag{11}$$

In other words, Y is markovian and gaussian conditionally to X. Such a model is very satisfying in the case of textured classes; in fact, this make possible to model situations in which each class ω_i can have its own texture modelled with $a_{\omega_i \omega_i}$.

The trouble is that, except in some particular cases, the product of (1) by (11) is not a Markov distribution. In fact, denoting by $\Gamma(x)$ the covariance matrix of the gaussian distribution (conditional on $X = x$) of the random vector $Y = (Y_s)_{s \in S}$, we have

$$\gamma(x) = \frac{1}{(\sqrt{2\pi})^n \det \Gamma(x)}, \tag{12}$$

which can not be written, in the general case, as a Markov distribution with respect to x. Thus here X is markovian, Y is markovian conditional to X, but neither (X, Y) nor X conditional to Y, are markovian. This is uncomfortable, because the markovianity of the posterior distribution of X is required to apply different segmentation methods. Even if this drawback can be circumvented by some model approximations, as in [6, 10], it remains a problem in general.

3. Pairwise Markov Random Fields and textured image segmentation

This section is devoted to the new PMRF model. In order to do a clear link with the previous section, we start with the particular gaussian case and then present the general case.

3.1. Gaussian case

In order to circumvent the difficulty due to the possibly non markovian form of $\gamma(x)$ given by (12) of the previous section, we propose to directly assume that the couple (X, Y) is markovian with respect to four nearest neighbours. More precisely, we put

$$P(X = x, Y = y) = \gamma \exp\left(-\sum_{t \leftrightarrow u} \varphi((x_t, y_t), (x_u, y_u)) - \sum_{s \in S} \varphi^*((x_s, y_s)) \right) =$$

$$\gamma \exp\left(-\sum_{t \leftrightarrow u} (\varphi_1(x_t, x_u) + a_{x_t x_u} y_t y_u) - \sum_{s \in S} (\varphi_2(x_s) + a_{x_s x_s} y_s^2 + b_{x_s} y_s) \right). \tag{13}$$

We notice immediately that in the PMRF given by (13) the distribution of Y conditioned on X is a Markov field and, what is more, is it exactly of the same shape as the distribution (11). So, the "approximate" HMRF model obtained by multiplying (1) by (11) gives the same distribution of Y conditioned on X as the PMRF. The difference is that in the "approximate" HMRF model the distribution of X is a Markov distribution given by (1), and in the PMRF the distribution of X is a different, and not necessarily markovian, distribution. In fact, the distribution of X in the PMRF (13), which is its marginal distribution, is given by

$$P(X = x) = \gamma\sqrt{(2\pi)^N \det(\Gamma(x))} \exp\left(-\sum_{t\leftrightarrow u}\varphi_1(x_t, x_u) - \sum_{s\in S}\varphi_2(x_s)\right). \qquad (14)$$

Finally, the model proposed differs from the HMRF model defined with (1) and (11) by:

(i) The marginal distribution of X (its prior distribution) is markovian in the classical HMRF model and is not necessarily markovian in the proposed PMRF model;

(ii) The posterior distribution of X is not necessarily markovian in the classical HMRF model and is markovian in the PMRF model. Furthermore, as a consequence, we have:

(iii) when dealing with the problem of segmentation of textured images, the proposed model makes possible the use of MAP or MPM without any model approximation.

Remark 3.1. The fact that the distribution of X is not necessarily a markovian distribution in a PMRF could be felt as a drawback; in fact, in bayesian context of work, one is used to being able to write the prior distribution. This does not seem to us to be a serious drawback for at least two reasons. First, the markovianity of X is rarely established in real world images of classes and it is rather arbitrarily assumed. Furthermore, the main goal of this assumption is to ensure the posterior markovianity of X, which makes possible different bayesian segmentations. In other words, the markovianity of X appears rather as a sufficient condition of the applicability of the bayesian methods than as a necessary one. Second, even in the very simple classical HMRF given by (4), the distribution of Y is not necessarity a Markov distribution. So, the non markovianity of X in PMRF should not seem more strange, a priori, that the non markovianity of Y in classical HMRF models.

Remark 3.2. The existence of the distribution (13) is not ensured for every φ_1, φ_2, $a_{x_s x_t}$, $a_{x_s x_s}$, a, b_{x_s}, b. To ensure its existence, one has to verify that for every fixed $x = (x_s)_{s\in S}$, the energy with respect to $y = (y_s)_{s\in S}$ is a gaussian energy; i.e., for every fixed $x = (x_s)_{s\in S}$,$P(X = x, Y = y)$ given by (13) integrable with respect to y.

Let us return to the distribution defined by (13) and let us specify how to simulate realizations of (X, Y). Let $s \in S$ and let t_1, t_2, t_3, t_4 be the four nearest neighbours of the pixel s. When using the classical Gibbs sampler, important is to be able to simulate (X_s, Y_s) according to its distribution conditional to

$$[(X_{t_1}, Y_{t_1}), ..., (X_{t_4}, Y_{t_4})] = [(x_{t_1}, y_{t_1}), ..., (x_{t_4}, y_{t_4})]. \qquad (15)$$

So, let us calculate the density of this distribution. We will show that it is of the separable form (which also depends, of course, on (x_{t_1}, y_{t_1}), (x_{t_2}, y_{t_2}), (x_{t_3}, y_{t_3}), and (x_{t_4}, y_{t_4}) which are omitted):

$$h(x_s, y_s) = p(x_s) f_{x_s}(y_s), \qquad (16)$$

where p is a probability on the set of classes and for each class x_s, the function f_{x_s} is a gaussian density corresponding to this class. Of course, (16) makes simulation very easy: one first simulates x_s according to p, and then y_s is simulated according to f_{x_s}. We have :

$$\begin{aligned} &P((X_s,Y_s)=(x_s,y_s) \mid (X_{t_1},Y_{t_1}),...,(X_{t_4},Y_{t_4})=(x_{t_1},y_{t_1}),...,(x_{t_4},y_{t_4})) \\ &\qquad \propto \exp\left(-\sum_{i=1,\dots,4}\varphi((x_s,y_s),(x_{t_i},y_{t_i}))-\varphi^*((x_s,y_s))\right) = \\ &\exp\left(-\sum_{i=1,\dots,4}(\varphi_1(x_s,y_{t_i})+a_{x_s x_{t_i}}y_s y_{t_i})-(\varphi_2(x_s)+a_{x_s}y_s^2+b_{x_s}y_s)\right) \end{aligned} \tag{17}$$

Omitting temporarily the dependence on x_s, let us put :

$$\begin{aligned} \alpha &= \varphi_2(x_s)+\sum_{i=1,\dots,4}\varphi_1(x_s,y_{t_i}) \\ \beta &= (b_{x_s}+\sum_{i=1,\dots,4}a_{x_s x_{t_i}}y_{t_i}y_s) \\ \delta &= a_{x_s} \end{aligned} \tag{18}$$

the conditional density above is equal to

$$\begin{aligned} &\exp-(\alpha+\beta y_s+\delta y_s^2)=\exp-\left(\frac{(y_s+\frac{\beta}{2\delta})^2}{\delta^{-1}}-(\frac{\beta}{2\delta})^2\delta+\alpha\right)= \\ &\sqrt{\pi\delta^{-1}}\exp\left(-(\frac{\beta}{2\delta})^2\delta+\alpha\right)\frac{1}{\sqrt{\pi\delta^{-1}}}\exp\left(-\frac{(y_s+\frac{\beta}{2\delta})^2}{\delta^{-1}}\right)= \\ &\qquad \sqrt{\pi\delta^{-1}}\exp\left(-(\frac{\beta}{2\delta})^2\delta+\alpha\right)f_{x_s}(y_s), \end{aligned} \tag{19}$$

where f_{x_s} is a gaussian density with the mean $-\frac{\beta}{2\delta}$ and the variance $\frac{1}{2\delta}$. Finally, recalling (18), the density (16) is given by the gaussian densities defined by means M_{x_s} and variances $\sigma_{x_s}^2$:

$$\begin{aligned} M_{x_s} &= -\frac{b_{x_s}+\sum_{i=1,\dots,4}a_{x_s x_{t_i}}y_{t_i}}{2a_{x_s}} \\ \sigma_{x_s} &= \frac{1}{2a_{x_s}} \end{aligned} \tag{20}$$

and p the probabilty given on the set of classes with:

$$p(x_s)=\frac{\sqrt{(a_{x_s})^{-1}}\exp\left(\frac{(b_{x_s}+\sum_{i=1,\dots,4}a_{x_s x_{t_i}}y_{t_i})^2}{2a_{x_s}}+\varphi_2(x_s)+\sum_{i=1,\dots,4}\varphi_1(x_s,x_{t_i})\right)}{\sum_{\omega\in\Omega}\sqrt{(a_\omega)^{-1}}\exp\left(\frac{(b_\omega+\sum_{i=1,\dots,4}a_{\omega x_{t_i}}y_{t_i})^2}{2a_\omega}+\varphi_2(\omega)+\sum_{i=1,\dots,4}\varphi(\omega,x_{t_i})\right)}. \tag{21}$$

So, (20) and (21) allow one to simulate realisations of (X_s, Y_s) according to its distribution conditional on $[(X_{t_1}, Y_{t_1}), ..., (X_{t_4}, Y_{t_4})] = [(x_{t_1}, y_{t_1}), ..., (x_{t_4}, y_{t_4})]$, which affords using the Gibbs sampler to simulate realisations of the Markov field (X, Y).

3.2. General framework

The generalisation of the simple PMRF described above is immediate. Let S be a set of pixels with $N = \text{card}\,(S)$. One may consider k classes $\Omega = \{\omega_1, ..., \omega_k\}$, m sensors (each $Y_s = (Y_s^1, ..., Y_s^m)$ takes its values in $\mathbb{R}^m$), and a set of cliques C corresponding to a neighbourhood system. The random field $Z = (Z_s)_{s \in S}$, with $Z_s = (X_s, Y_s)$, is called a multisensor Pairwise Markov Random Field if its distribution may be written as

$$P(Z = z) = \gamma \exp\left(-\sum_{c \in C} \varphi_c(z_c)\right). \tag{22}$$

In particular, the three-sensor PMRF may be used in segmenting colour images, a problem in which statistical methods take growing importance, [16].

Remark 3.3. Let us notice that the conditions (i) and (ii) of the paragraph 2.1 are not necessary to obtain simultaneously prior and posterior markovianity of X. For instance, let X be a Markov field with respect to some neighbourhood, and let $B = (B_s)_{s \in S}$ a "noise" field, markovian with respect to the same neighbourhood as X. Let us assume that:

(i) the fields X and B are independent;
(ii) for each $s \in S$, we have $Y_s = F(X_s, B_s)$ and $B_s = G(X_s, Y_s)$; i.e., F can be inverted for evey fixed x_s.

Than we can state :

$$P(Y = y \mid X = x) = \left| \prod_{s \in S} \frac{\partial G}{\partial y_s}(x_s, y_s) \right| P(B = G(x, y)) \tag{23}$$

and

$$P(X = x, Y = y) = P(X = x) \left| \prod_{s \in S} \frac{\partial G}{\partial y_s}(x_s, y_s) \right| P(B = G(x, y)), \tag{24}$$

which implies that (X, Y) is a Markov field with respect to the same neighbourhood as the Markov fields X and B, and so the posterior distribution of X is still markovian with respect to the same neighbourhood. Such models can take correlated noise into account; however, they can not model really different textures because of the simplicity of the hypothesis (ii), according to which $Y = F(X, B)$ is defined "point by point" from the fields X and B : $Y_s = F(X_s, B_s)$ for each $s \in S$. In particular, all the textures would have the same correlation coefficient. When considering more complex F, the determinant in (23),(24) would no longer be $\prod_{s \in S} \frac{\partial G}{\partial y_s}(x_s, y_s)$, and so the markovianity would not be ensured in general.

In gaussian case, the model above generalises the simple model described in the paragraph 2.2. Taking the same distribution for X, let B be gaussian, centred, markovian with respect to four nearest neighbours, and verifying $E(B_s)^2 = 1$. Considering $Y_s = F(X_s, B_s) = m_{x_s} + \sigma_{x_s} B_s$ (and thus $B_s = G(X_s, Y_s) = \frac{Y_s - m_{x_s}}{\sigma_{x_s}}$) we obtain a more general model than the model described in the paragraph 2.2, where the field B was a white gaussian noise.

4. Visual examples

This section is devoted to some simulation results. We show that different class images and different noise correlations can be obtained with the PMRF model, and some segmentation results are presented. We have chosen rather noisy cases to show that the high segmentation power of classical hidden Markov fields is preserved.

The PMRFs presented in fig. 2 are markovian with respect to eight nearest neighbours. Assuming that ϕ_c are null for cliques containing more than two pixels, the general form of the distribution of (X, Y) is (see (22) of the previous section):

$$P(X = x, Y = y) = \gamma \exp\left(-\sum_{t \leftrightarrow u} \phi_{(t,u)}[(x_t, y_t), (x_u, y_u)] - \sum_{s \in S} \phi_s[(x_s, y_s)]\right) \tag{25}$$

and we take the following particular gaussian form:

$$\begin{aligned} \phi_{(t,u)}[(x_t, y_t), (x_u, y_u)] &= \frac{1}{2}(a_{x_t x_u} y_s y_t + b_{x_t x_u} y_s + c_{x_t x_u} y_t + d_{x_t x_u}) \\ \phi_s[(x_s, y_s)] &= \frac{1}{2}(\alpha_{x_s} y_s^2 + \beta_{x_s} y_s + \gamma_{x_s}). \end{aligned} \tag{26}$$

Let us note that the neighbors $t \leftrightarrow u$ can form four different cliques, which are specified in fig. 1.

Fig. 1. Four possible forms of cliques.

So we have, in the general case, four different functions $\phi_{(t,u)}$, that we will call of the form $1, 2, 3$, and 4, and denote ϕ_1, ϕ_2, ϕ_3, and ϕ_4, respectively.

Otherwise, it is useful to have some information about the distribution of Y conditional to $X = x$. In fact, we have seen that it was a gaussian distribution and knowing some parameters can be useful in adding or subtracting noise in simulations. More precisely, let Σ_x be the covariance matrix of the gaussian distribution of Y conditional to $X = x$ and let $Q_x = (q_{x_t, x_u})_{t,u \in S} = \Sigma_x^{-1}$. We have:

$$P(Y = y \mid X = x) \propto \exp\left(-\frac{1}{2}(y - m_x)^t Q_x (y - m_x)\right). \tag{27}$$

By expanding (27) and comparing with (26) one may specify different relations, and in particular (28) and (29) below.

$$m_{x_s} = -\frac{\beta_{x_s}}{2\alpha_{x_s}}, \qquad \sigma^2_{x_s} = \frac{1}{\alpha_{x_s}}. \tag{28}$$

The relation (28) may be used to modulate the level of the noise (notice that the noise level also depends on the noise correlation, which is not calculable from (26), and on the distribution of X).

Let us specify the functions defining $\phi_{t,u}$ and ϕ_s for the three couples of images (Image 1, Image 2),(Image 4, Image 5), and (Image 7, Image 8) in fig. 2. Given the following relations

$$\begin{gathered} a_{x_t x_u} = 2q_{x_t,x_u},\ b_{x_t x_u} = -2q_{x_t,x_u} m_{x_u x_t},\ c_{x_t x_u} = -2q_{x_t,x_u} m_{x_t}, \\ d_{x_t x_t} = 0.4 m_{x_t} m_{x_u} + \varphi_{(t,u)}(x_t, x_u),\ \alpha_{x_s} = 1,\ \beta_{x_s} = -2m_{x_s},\ \gamma_{x_s} = m^2_{x_s}. \end{gathered} \tag{29}$$

and labelling $m_1 = -0.25,\ m_2 = 0.25$ for the three couples, all we have to specify is $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ and q_{x_t,x_u} for each couple. For (Image 1, Image 2) we take:

$$\varphi_1(x_t, x_u) = \varphi_2(x_t, x_u) = \varphi_3(x_t, x_u) = \varphi_4(x_t, x_u) = \begin{cases} 1 & \text{if } x_s = x_t \\ -1 & \text{otherwise} \end{cases}$$

and

$$q_{x_t,x_u} = -0.2.$$

The parameters taken for (Image 4, Image 5) are:

$$\varphi_1(x_t, x_u) = \varphi_2(x_t, x_u) = \varphi_3(x_t, x_u) = \begin{cases} 1 & \text{if } x_s = x_t \\ -1 & \text{otherwise} \end{cases}$$

$$\varphi_4(x_t, x_u) = 5\varphi_1(x_t, x_u)$$

and

$$q_{x_t,x_u} = -0.1.$$

Finally the last couple (Image 7, Image 8) is sampled with:

$$\varphi_2(x_t, x_u) = \varphi_3(x_t, x_u) = \varphi_4(x_t, x_u) = \begin{cases} 1 & \text{if } x_s = x_t \\ -1 & \text{otherwise.} \end{cases}$$

$$\varphi_1(x_t, x_u) = 5\varphi_2(x_t, x_u)$$

and

$$q_{x_t,x_u} = \begin{cases} -0.1 & \text{if } x_t = x_u = 1 \\ 0 & \text{otherwise} \end{cases}$$

The estimated covariances and the error rates in MPM segmentations are given in tab. 1, and the segmentation results are presented in fig. 2.

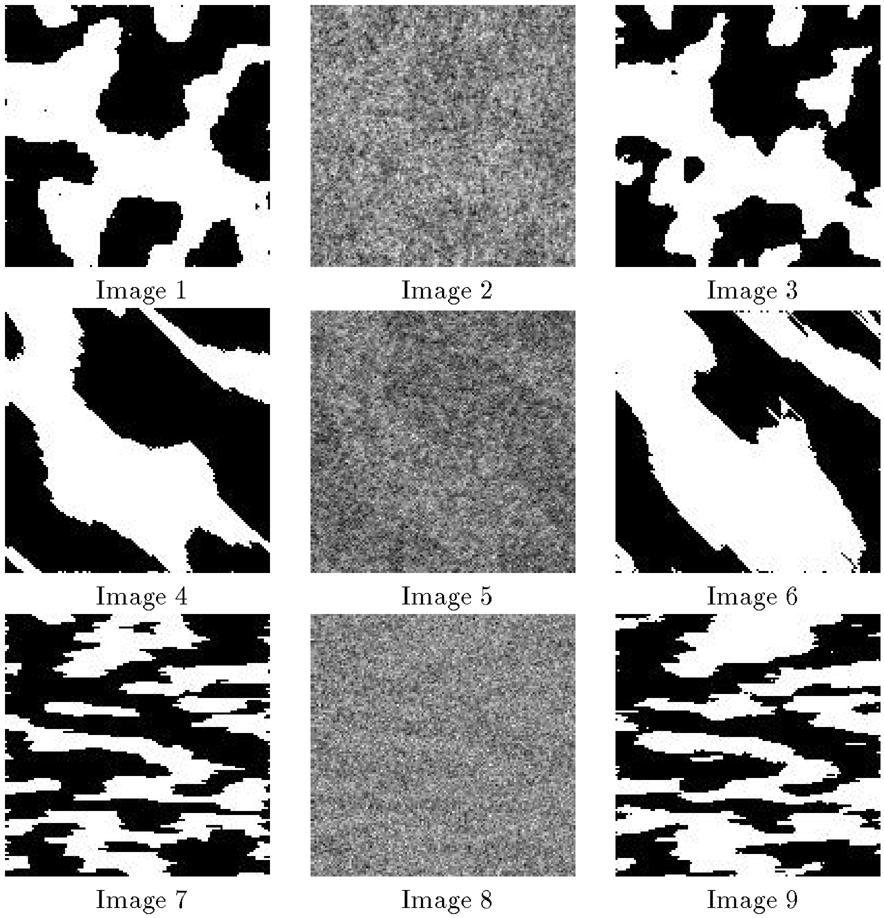

Fig. 2. Three realizations of Pairwise Markov Fields (Image 1, Image 2), (Image 4, Image 5),(Image 7, Image 8), and the MPM segmentations of Image 2 (giving Image 3), Image 5 (giving Image 6), and Image 8 (giving Image 9), respectively. The different parameters are specified in tab. 1.

5. Conclusions

We proposed in this paper a novel model called Pairwise Markov Random Field (PMRF). A random field of classes x and a random field of observations Y form a PMRF when the pairwise random field $Z = (X, Y)$ is a Markov field. Such a model is different from the classical Hidden Markov Random Field (HMRF); in particular, in the PMRF the

	Images 1,2,3	Images 4,5,6	Images 7,8,9
ρ_{11}	0.25	0.1	0.11
ρ_{21}	0.25	0.1	0.0
τ	13 %	11 %	10 %

Tab. 1. ρ_{11}, ρ_{22t} : the estimated covariances inter-class (neighbouring pixels).
τ : the error rate of wrongly classified pixels with MPM. The number of iterations in MPM is 25×30 (the posterior marginals are estimated from 20 realizations, each realization being obtained after 50 iterations of the Gibbs Sampler).

random field x is not necessarily a Markov field.

The PMRF allows one to deal with the statistical segmentation of textured images which can be, in addition, corrupted with correlated noise. Contrary to the use of hierarchical models [4], this can be done in the framework of the model, without any approximations. Roughly speaking, in the Hierarchical HMRF the prior distribution of X is markovian and its posterior distribution is not markovian; and in PMRF the prior distribution of x is not markovian but its posterior distribution is. When using a bayesian method of segmentation like MPM or MAP, we have to make some approximations when using Hierarchical HMRF, unlike PMRF. Furthermore, the distributions of Y conditional to X, which model different textures and different possibly correlated noises, can be strictly the same in the both Hierarchical HMRF and PMRF models.

We have presented different simulations of PMRF and different results of the Bayesian MPM segmentation of the observation fields. The cases presented are rather noisy and the results show that the well known efficiency of the HMRF is preserved when using PMRF, at least in the simple cases presented.

We mainly dealt with textured images segmentation, but, of course, PMRF can by used in any classification problem in which the observations are spatially dependent and in which inside of each class the observations are possibly correlated, with possibly different correlations attached with different classes.
As perspectives for further work, we may put forth the following.

(i) Parameter estimation from Y, allowing unsupervised segmentation. Given that the simulations of X according to its posterior distribution are feasible, we may propose the use of Iterative Conditional Estimation (ICE, [9]) for parameter estimation. This method has given good results, even in more complex situations where the form of the noise corresponding to each class is not known, [14, 15], or still in hierarchical models, [18, 20]. To apply ICE, one needs an estimator from complete data (X, Y) and the choice of the stochastic gradient [8], with the difference that it would be applied to (X, Y) instead of X, for this estimator could be a good one. In fact, the use of the stochastic gradient with ICE has already given good results in the context of fuzzy Markov random fields [17], which is more complex that the classical context of hidden Markov random fields;

(ii) Another direction of research could be the extension of the "pairwise" notion to other hidden Markov models. This has been recently one concerning the classical Hidden Markov Chains (HMC) by introducing Pairwise Markov Chain (PMC) model [22]. So, one could view the possibilities of applying the "pairwise" notion to different general hidden Markov models on networks, [19];

(iii) We have mainly introduced the PMRF models because of the lack of the markovianity of $\gamma(x)$ defined by (12), or rather because of the lack of the certainty that $\gamma(x)$ is markovian. However, we have not specified under what conditions $\gamma(x)$ is not markovian, at least with respect to a given neighbourhood. Even if the PMRF models allow us to avoid searching theses conditions, it would undoubtedly be interesting from the theoretical point of view.

Acknowledgement

We thank Alain Hillion, Directeur Scientifique de l'Ecole Nationale Supérieure des Télécommunications de Bretagne, and Jean-Marc Boucher, Professor at Ecole Nationale Supérieure des Télécommunications de Bretagne, for numerous discussions which greatly helped the writing of this paper. We also thank the anonymous reviewers, whose remarks allowed us to improve its readability.

References

1983

[1] Cross G. R., Jain A. K.: Markov random field texture models. IEEE Trans. PAMI, 5(1), 25-39.

1984

[2] Geman S., Geman D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. on PAMI, 6(6), 721-741.

1986

[3] Besag J.: On the statistical analysis of dirty pictures. J. Royal Statistical Society, Series B, 48, 259-302

1987

[4] Derin H., Elliot H.: Modelling and segmentation of noisy and textured images using Gibbs random fields. IEEE Trans. on PAMI, 9(1), 39-55.

[5] Marroquin J., Mitter S., Poggio T.: Probabilistic solution of ill-posed problems in computational vision. J. American Statistical Association, 82, 76-89.

1988

[6] Kelly P. A., Derin H., Hart K.D.: Adaptive segmentation of speckled images using a hierarchical random field model. IEEE Trans. ASSP, 36(10), 1628-1641.

1989

[7] Lakshmanan S., Derin H.: Simultaneous parameter estimation and segmentation of Gibbs random fields. IEEE Trans. PAMI, 11, 799-813.

[8] Younes L.: Estimation and annealing for Gibbsian fields. Annales de l'Institut Henri Poincaré, 24(2), 269-294,.

1992

[9] Pieczynski W.: Statistical image segmentation. MG&V, 1(1/2), 261-268,.

[10] Won C. S., Derin H.: Unsupervised segmentation of noisy and textured images using Markov random fields. CVGIP: GMIP, 54(4), 308-328,.

1993
[11] Chellapa R., Jain A. (Ed.): Markov Random Fields, Theory and Application, AP, San Diego,.
[12] Cressie N. A. C.: Statistics for Spatial Data. Wiley, Series in Probability and Mathematical Statistics.
1995
[13] Guyon X.: Random Fields on Network. Modeling, Statistics, and Applications. S-V, Probabily and its Applications, NY.
1997
[14] Delignon Y., Marzouki A., Pieczynski W.: Estimation of generalized mixture and its application in image segmentation. IEEE Trans. IP, 6(10), 1364-1375.
[15] Giordana N., Pieczynski W.: Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation. IEEE Trans. PAMI, 19(5), 465-475.
[16] Yamazaki T., Gingras D.: Image classification using spectral and spatial information based on MRF models. IEEE Trans. IP, 4(9), 1333-1339.
1997
[17] Salzenstein F., Pieczynski W.: Parameter estimation in hidden fuzzy markov random fields and image segmentation. GMIP, 59(4), 205-220.
1999
[18] Kato Z., Zerubia J., Berthod M.: Unsupervised parallel image classification using markovian models. PR, 32, 591-604.
[19] Serfozo R.: Introduction to Stochastic Networks. S-V, NY.
2000
[20] Mignotte M., Collet C., Pérez P., Bouthémy P.. Sonar image segmentation using an unsupervised hierarchical MRF model. IEEE Trans. IP, 9(7), 1216-1231,.
[21] Pieczynski W., Tebbache A.-N.: Pairwise Markov random fields and their application in textured images segmentation. Proc. IEEE Southwest Symp. on Image Analysis and Interpretation (SSIAI'2000), 2-4 April, Austin, Texas, USA, 106-110.
[22] Pieczynski W.: Pairwise Markov chains and bayesian unsupervised fusion. Proc. 3rd Int. Conf.on Information Fusion, FUSION 2000, 1, July 10th-13th, Paris, France, MoD4-24 - MoD4-31.

Wojciech Pieczynski received the Doctorat d'Etat degree from Pierre et Marie Curie University, Paris, France, in 1986. He is currently Professor and Head of the Statistical Image Processing Group, Institut National des Télécommunications, Evry, France. His research interests include mathematical statistics, stochastic processes, and statistical image processing.

Abdel-Nasser Tebbache received the Diplôme d'Etudes Approfondies degree from Pierre et Marie Curie University, Paris, France, in 1997. He is currently completing the PhD degree at the Institut National des Télécommunications, Evry, and at the Ecole Nationale Supérieure de Télécommunications de Bretagne, Brest, France. His research interests include bayesian classification, hidden Markov models parameter estimation, and sonar image segmentation.