Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Image and Vision Computing 24 (2006) 61–69
image and vision COMPUTING
www.elsevier.com/locate/imavis

Multisensor triplet Markov fields and theory of evidence

Wojciech Pieczynski *, Dalila Benboudjema

Département CITI, Institut National des Telecommunications (GET/INT), 9 Rue Charles Fourier, 91000 Evry, France

Received 20 October 2003; received in revised form 20 September 2005; accepted 24 September 2005

Abstract

Hidden Markov Fields (HMF) are widely applicable to various problems of image processing. In such models, the hidden process of interest X is a Markov field, which must be estimated from its observable noisy version Y. The success of HMF is due mainly to the fact that X remains Markov conditionally on the observed process, which facilitates different processing strategies such as Bayesian segmentation. Such models have been recently generalized to 'Pairwise' Markov fields (PMF), which offer similar processing advantages and superior modeling capabilities. In this generalization, one directly assumes the Markovianity of the pair (X,Y). Afterwards, 'Triplet' Markov fields (TMF) have been proposed, in which the distribution of (X,Y) is the marginal distribution of a Markov field (X,U,Y), where U is an auxiliary random field. So U can have different interpretations and, when the set of its values is not too complex, X can still be estimated from Y. The aim of this paper is to show some connections between TMF and the Dempster–Shafer theory of evidence. It is shown that TMF allow one to perform the Dempster–Shafer fusion in different general situations, possibly involving several sensors. As a consequence, Bayesian segmentation strategies remain applicable.

Keywords: Multisensor hidden Markov fields; Pairwise Markov fields; Triplet Markov fields; Bayesian classification; Dempster–Shafer fusion; Theory of evidence; Statistical unsupervised non-stationary image segmentation

1. Introduction

Hidden Markov Fields (HMF) are widely used in various image-processing problems (see [8,18,37], among others). They consist in considering two stochastic processes $X=(X_s)_{s\in S}$ and $Y=(Y_s)_{s\in S}$, in which $X=x$ is unobservable and must be estimated from the observed $Y=y$. The qualifier 'hidden Markov' means that the hidden process X has a Markov distribution. When the distributions $p(y|x)$ of Y conditional on $X=x$ are simple enough, the pair (X,Y) retains the Markovian structure, and likewise for the distribution $p(x|y)$ of X conditional on $Y=y$. The Markovianity of $p(x|y)$ is crucial if one's purpose is to estimate $X=x$ from $Y=y$. HMF have then been generalized to pairwise Markov fields (PMF [21]) and triplet Markov fields (TMF [1,24]), which are more general and still allow one to recover $X=x$ from $Y=y$. In particular, PMF have been proposed to solve the texture modeling problem, which is rather difficult and needs approximations when using HMF [17]. Considering that the pair (X,Y) is a PMF consists in assuming that the pair (X,Y) is a Markov field, which ensures the Markovianity of $p(y|x)$ and the Markovianity of $p(x|y)$. Such a model is not necessarily a HMF because X is not necessarily a Markov field. Thus compared to HMF, the Markovianity of $p(y|x)$ allows better modeling, and the Markovianity of $p(x|y)$ allows the same processing properties. The generalization of PMF to TMF consists in introducing a third process $U=(U_s)_{s\in S}$ and considering that (X,U,Y) is a Markov process. The process U needs not have a physical existence and the problem remains the same as above: estimating $X=x$ from $Y=y$. The Triplet models are more general than the Pairwise models because the distribution of (X,Y), which is a marginal distribution of (X,U,Y), is not necessarily a Markov distribution. However, the classical processing methods remain workable in the TMF context, as long as the set of values of U is not too complex. As specified in [1], this third random field U can have different meanings. In particular, it can model the fact that the field X may not be stationary, which seems to present encouraging perspective [2].

The aim of this paper is to propose new applications of TMF to the problem of Dempster–Shafer fusion (DS fusion) in a Markov field context. Although DS fusion is now well known and widely used in various situations [5,9,11,19,32–34,38], its use in the Markov field context is very rare; only a few papers deal with this kind of models [3,6,10,15,31,36]. To be more precise, we show how the use of TMF allows one to simultaneously integrate the possibly evidential aspects of

* Corresponding author. Tel.: +33 1 6076 4425; fax: +33 1 6076 4433.
E-mail address: wojciech.pieczynski@int-evry.fr (W. Pieczynski).

0262-8856/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2005.09.012

the prior information and the possibly evidential aspects of the sensors. The first integration can be of interest when the hidden scene is non-stationary, and the second one enables one to extend different classical multisensor images analysis (see [14,29,35], among others).

Let us mention that similar extensions have been proposed in the case of Markov chains. In fact, hidden Markov chains (HMC) have been generalized to pairwise Markov chains (PMC [23]), and triplet Markov chains (TMC [25]). In particular, evidential priors have been introduced in HMC [7,28], resulting in an improvement in the efficiency of the unsupervised segmentation of non-stationary chains [12]. Moreover, still more complex models, including partially Markov chains or semi-Markov chains, have been recently proposed in [26,27]. All these models can be possibly extended to Markov trees, and some first ideas are presented in [13]. However, although these very general ideas have inspired the main ideas of the present paper, the Markov field models—and the related problems such as, for instance, parameter estimation—are very different from the Markov chain models and the related problems. Therefore, below we will concentrate on Markov fields with no further consideration of Markov chains.

The organization of the paper is as follows. The basic notions and calculations relating to the theory of evidence are recalled in Section 2, while Section 3 is devoted to the introduction of evidential priors in the context of Markov fields. The use of evidential sensors is discussed in Section 4, and the general model, including evidential priors and evidential sensors, is specified in Section 5. Section 6 contains conclusions and some future prospects.

2. Theory of evidence

Let us consider a finite set of classes $\Omega=\{\omega_1,\ldots\omega_k\}$, and its power set $P(\Omega)=\{A_1,\ldots A_q\}$, with $q=2^k$. A function M from $P(\Omega)$ to [0,1] is called a 'basic belief assignment' (bba) if $M(\varnothing)=0$ and $\sum_{A\in P(\Omega)} M(A)=1$. A bba M defines a 'plausibility' function Pl from $P(\Omega)$ to [0,1] by $\mathrm{Pl}(A)=\sum_{A\cap B\neq\varnothing} M(B)$, and a 'credibility' function Cr from $P(\Omega)$ to [0,1] by $\mathrm{Cr}(A)=\sum_{B\subset A} M(B)$. For a given bba M, the corresponding plausibility function Pl and credibility function Cr are linked by $\mathrm{Pl}(A)+\mathrm{Cr}(A^c)=1$. So, each of them defines the other. Conversely, Pl and Cr can be defined by some axioms, and each of them defines an unique corresponding bba M. More precisely, Cr is a function from $P(\Omega)$ to [0,1] verifying $\mathrm{Cr}(\varnothing)=0$, $\mathrm{Cr}(\Omega)=1$, and $\mathrm{Cr}\left(\bigcup_{j\in J} A_j\right)\geq\sum_{I\neq\varnothing^I\subset J}(-1)^{|I|+1}\mathrm{Cr}\left(\bigcap_{j\in I} A_j\right)$, and Pl is a function from $P(\Omega)$ to [0,1] verifying analogous conditions, with $\leq$ instead of $\geq$ in the third one. A credibility function Cr verifying such conditions is also the credibility function defined by the bba $M(A)=\sum_{B\subset A}(-1)^{|A-B|}\mathrm{Cr}(B)$.

Finally, each of the three functions M, Pl, and Cr can be defined in an axiomatic way, and each of them defines the two others. Furthermore, when M is null outside singletons, the corresponding Pl and Cr are equal and become a classical probability. Thus a probability is obtained for a particular M, which will be called 'probabilistic' in the following.

When two bbas M_1, M_2 represent two pieces of evidence, we can combine—or fuse—them using the so-called 'Dempster–Shafer combination rule', or 'Dempster–Shafer fusion' (DS fusion) which gives $M=M_1\oplus M_2$ defined by:

$$M(A)=(M_1\oplus M_2)(A)=\begin{cases}\dfrac{1}{1-H}\displaystyle\sum_{(B_1,B_2)\in\Omega^2/B_1\cap B_2=A} M_1(B_1)M_2(B_2), & \text{for } A\neq\varnothing,\\ 0, & \text{for } A=\varnothing\end{cases}\tag{2.1}$$

Let us notice that the constant

$$H=1-\sum_{A\subset\Omega}\left[\sum_{(B_1,B_2)\in\Omega^2/B_1\cap B_2=A\neq\varnothing} M_1(B_1)M_2(B_2)\right]=\sum_{(B_1,B_2)\in\Omega^2/B_1\cap B_2=\varnothing} M_1(B_1)M_2(B_2)$$

has an intuitive meaning and can be interpreted as the degree of conflict between the two pieces of evidence modeled by the bbas M_1 and M_2.

In the following, we will use the proportionality symbol ' ∝ ' which is very practical to manipulate the DS fusion. Therefore, (2.1) will be written as:

$$M(A)=(M_1\oplus M_2)(A)\propto\sum_{B_1\cap B_2=A} M_1(B_1)M_2(B_2)\tag{2.2}$$

knowing that $A\neq\varnothing$ and $M(A)=(M_1\oplus M_2)(A)$ is obtained by dividing the r.h.s. of (2.1) by the sum

$$\sum_{A\subset\Omega}\left[\sum_{(B_1,B_2)\subset\Omega^2/B_1\cap B_2=A\neq\varnothing} M_1(B_1)M_2(B_2)\right]$$

As mentioned above, we will say that a bba M is 'probabilistic' when, being null outside singletons, it defines a probability and we will say that it is an 'evidential' bba when it is not probabilistic. As can be seen easily, when either M_1 or M_2 is probabilistic, the fusion result M is probabilistic.

In particular, one can see that the classical calculus of the posterior probability is a Dempster–Shafer fusion (DS fusion) of two probabilistic bbas. For example, let us consider two random variables X and Y taking their values in $\Omega=\{\omega_1,\omega_2\}$ and R, respectively. Let M_0 be the law of X (which is a probability on Ω, and thus also a bba null outside singletons), and let $p(y|x=\omega_1)$, $p(y|x=\omega_2)$ be the distributions of Y conditional on $X=\omega_1$ and ω_2, respectively. For observed $Y=y$, let M_1 the probability on Ω defined by $M_1(\omega_1)=p(y|x=\omega_1)/(p(y|x=\omega_1)+p(y|x=\omega_2))$ and $M_1(\omega_2)=p(y|x=\omega_2)/(p(y|x=\omega_1)+p(y|x=\omega_2))$ (which can be written $M_1(x)\propto p(y|x)$, where $p(y|x)$ is the distribution of Y conditional on $X=x$). Then a very simple calculus shows that the posterior distribution of X, i.e. its distribution conditional on $Y=y$, is the Dempster–Shafer fusion of M_0 with M_1. This simple fact opens numerous perspectives of extension of the posterior distribution of X: if we replace either M_0 or M_1 by a bba which is not null outside singletons, then the fusion M_0 and M_1 gives a probability which is an extension of the posterior distribution of X. Such a fusion result can then be used, in a strictly same manner as the classical one, in different Bayesian techniques for estimating X from $Y=y$. Studying these different extensions in the Markov fields context is the very aim of the paper.

Let us consider the following two examples, which briefly describe how the DS fusion can be used in the image analysis context.

Example 2.1. Let us consider the problem of satellite or airborne optical image segmentation into two classes $\Omega=\{\omega_1, \omega_2\}$ 'forest' and 'water'. Thus we have two random fields $X=(X_s)_{s\in S}$ and $Y=(Y_s)_{s\in S}$, each X_s taking its values in $\Omega=\{\omega_1, \omega_2\}$ and each Y_s taking its values in R. Let us assume that there are clouds and let us denote by ω_c the 'clouds' class. Thus, in the classical probabilistic context, there are three conditional distributions on R:$p(y_s|x_s=\omega_1)$, $p(y_s|x_s=\omega_2)$ and $p(y_s|x_s=\omega_c)$. This classical model admits the following 'evidential' interpretation. If we are only interested on the classes $\Omega=\{\omega_1,\omega_2\}$, the class ω_c brings no information about them, and thus ω_c models the ignorance and is assimilated to Ω. Finally, the classical probability q^{y_s} defined on $\{\omega_1,\omega_2,\omega_c\}$ by $q^{y_s}(\omega_1)\propto p(y_s|x_s=\omega_1)$, $q^{y_s}(\omega_1)\propto p(y_s|x_s=\omega_2)$, and $q^{y_s}(\omega_1)\propto p(y_s|x_s=\omega_c)$ is interpreted as a bba M_2 defined on $P(\Omega)=\{\omega_1, \omega_2,\Omega\}$ by $M_2(\{\omega_1\})\propto p(y_s|x_s=\omega_1)$, $M_2(\{\omega_2\})\propto p(y_s|x_s=\omega_2)$, and $M_2(\Omega)\propto p(y_s|x_s=\Omega)$. In other words, $M_2(\Omega)$ models the ignorance attached with the fact that one cannot see through clouds. An interesting result states that when M_1 is a Markov probabilistic field distribution, its DS fusion with M_2 remains a Markov distribution, which generalizes the classical calculus of Markov posterior distribution [3]. Furthermore, when clouds disappear, $M(\Omega)$ becomes null, and such a 'generalized' Markov model becomes a classical HMF. So, the generalization we obtain embeds HMF particular case. Different Bayesian processing methods can then be based on the fused Markov distribution obtained this way. Roughly speaking, when the prior distribution is a Markov probabilistic bba, its fusion with a more general evidential bba does not pose problem and Bayesian processing remains workable in this more general context.

Example 2.2. Let us consider the problem of segmentation of an observed bi-sensor image $Y=(Y^1,Y^2)=(y^1,y^2)$ into three classes. So, we have $\Omega=\{\omega_1,\omega_2,\omega_3\}$. Imagine that the first sensor Y^1 is sensitive to $\omega_1,\omega_2,\omega_3$ and $\{\omega_1,\omega_2,\omega_3\}$ (to fix ideas, it is an optical satellite sensor and there are clouds, modeled as in Example 2.1). Imagine that the second sensor Y^2 is sensitive to $\{\omega_1,\omega_2\}$ and ω_3 (to fix ideas, it is an infrared satellite sensor and ω_1, ω_2 are 'hot', while ω_3 is 'cold'). Let $\Lambda_1=\{\{\omega_1\},\{\omega_2\},\{\omega_3\},\{\omega_1,\omega_2,\omega_3\}\}$, $\Lambda_2=\{\{\omega_1,\omega_2\},\{\omega_3\}\}$, and $\Lambda_{1,2}=\{\{\omega_1\},\{\omega_2\},\{\omega_3\},\{\omega_1,\omega_2\}\}$. As above, y^1 defines a probability M^1 on Λ_1 and, in a similar way, y^2 defines a probability M^2 on Λ_2. As M^1 and M^2 are also bbas on $P(\Omega)$, we can fuse them, and the result $M^3=M^1\oplus M^2$ is a probability on $\Lambda_{1,2}$. Finally, M^3 can be fused with a probabilistic Markov field, resulting in a Markov distribution which extends the classical posterior distribution.

3. Triplet Markov fields with evidential priors

Let S be the set of pixels, and X,Y two random fields defined on S as specified in Section 1. Thus for each $s\in S$, the variables X_s and Y_s take their values in $\Omega=\{\omega_1,\ldots,\omega_k\}$ and R, respectively. The problem is to estimate X from Y, with immediate application to image segmentation. Considering a triplet Markov field (TMF), consists in introducing a third random field $U=(U_s)_{s\in S}$, where each U_s takes its values in $\Lambda=\{\lambda_1,\ldots,\lambda_m\}$, and in assuming that $T=(X,U,Y)$ is a Markov field. The distribution of $T=(X,U,Y)$ is then a classical Gibbs one:

$$p(t) = \gamma \exp\left[-\sum_{c\in C}\varphi_c(t_c)\right] \tag{3.1}$$

Classically, (3.1) means that the distribution of T verifies $p(t_s|t_r,r\neq s)=p(t_s|t_r,r\in W_s)$, where $(W_s)_{s\in S}$ is some neighborhood system (W_s is the set of neighbors of s). C is the set of cliques associated with $(W_s)_{s\in S}$ (a clique is either a singleton, or a set of mutually neighbors pixels).

In this model, X and Y are interpreted as usual, as for U, it can have physical meaning or not. More precisely, there are at least three situations in which U models some reality: (i) the noise densities are unknown, and are approximated by Gaussian mixtures; (ii) there are subclasses in at least one class among $\omega_1,\ldots,\omega_k$; and (iii) U models different stationarities of the field X (see [1] for (i) and (ii), and [2] for (iii)).

Important is that (X,U) can be classically simulated according to the Markov distribution $p(x,u|y)$, which enables to implement their different Bayesian estimations methods. For example, the classical Maximum Posterior Mode (MPM) method gives $\hat{x}=(\hat{x}_s)_{s\in S}$ such that for each $s\in S$ $p(\hat{x}_s|y)=\max_{x_s\in\Omega}p(x_s|y)$. This estimate can be computed once the posterior marginal distributions $p(x_s|y)$ are known: as $p(x_s,u_s|y)$ can be classically estimated from a simulated sample, $p(x_s|y)$ is given by $p(x_s|y)=\sum_{u_s\in\Lambda}p(x_s,u_s|y)$. Of course, U can be estimated similarly if it is of interest [2].

We now deal with the main point of this paper, i.e. we want to explain how the classical computing of the posterior distribution in hidden Markov fields can be extended to DS fusion when the hidden data become 'evidential'. Let us consider the classical hidden Markov field, with $p(x)=\gamma\exp\left[-\sum_{c\in C}\varphi_c(x_c)\right]$ and $p(y|x)=\prod_{s\in S}p(y_s|x_s)$. Then the distribution of (X,Y) is defined by

$$p(x,y) = \gamma \exp\left[-\sum_{c\in C}\varphi_c(x_c) + \sum_{s\in S}\mathrm{Log}[p(y_s|x_s)]\right] \tag{3.2}$$

The key point is to use the observation that $p(x|y)$ given by (3.2) can be seen as the result of the DS fusion of $p(x)=\gamma\exp\left[-\sum_{c\in C}\varphi_c(x_c)\right]$ with the probability distribution $q^y(x)$ (y is fixed) obtained by normalizing $p(y|x)$. More precisely, putting

$$q^y(x) = \prod_{s\in S}\left[\left((p(y_s|x_s))/\sum_{\omega\in\Omega} p(y_s|x_s=\omega)\right)\right]$$

(we will write $q^y(x) \propto \prod_{s\in S} p(y_s|x_s)$ in the following), we have $p(x|y)=(p\oplus q^y)(x)$ (where p is the Markov distribution of X). This provides numerous possibilities of extension of $p(x|y)$, by replacing in $p\oplus q^y$ either p or q^y by a mass function. In this section we shall show that when the Markov probability distribution p is replaced in $p\oplus q^y$ by a 'Markov' mass function M, the $M\oplus q^y$ can be seen as a marginal distribution of a particular TMF. An important consequence of this observation is that $M\oplus q^y$ can be used to estimate X from Y. Now, let us consider a set of classes Ω and the power set $P(\Omega)$. Let $n=\mathrm{Card}(S)$ and M^0 be a bba defined on $[P(\Omega)]^n$ by

$$M^0(A) = \gamma \exp\left[-\sum_{c\in C}\psi_c(A_c)\right] \quad (3.3)$$

where C is the set of cliques corresponding to a given neighborhood defining the Markovianity, $A=(A_s)_{s\in S}$, and $A_c=(A_s)_{s\in c}$. Such a bba will be called 'evidential Markov field' (EMF). We see that an EMF extends the classical Markov field, the latter being obtained when M^0 is null outside $\{\{\omega_1\},\ldots,\{\omega_k\}\}^n$. Both EMF M^0 and the probability distribution q^y (given by $p(y|x)=\prod_{s\in S} p(y_s|x_s)$) define a 'hidden' EMF (HEMF). We have the following result.

Proposition 3.1. Let M^0 be an EMF defined on $[P(\Omega)]^n$ by (3.3), and M^1 a probability over Ω^n defined from the observed field $Y=y\in R^n$ by $M^1(x)=q^y(x)\propto\prod_{s\in S} p(y_s|x_s)$. Let $\Lambda=P(\Omega)$ and $\Delta\subset\Omega\times\Lambda$ such that $(\omega,\lambda)\in\Delta$ if and only if $\omega\in\lambda$.

Then the probability distribution $M=M^0\oplus M^1$ is the marginal distribution $p(x|y)=\sum_{u\in[P(\Omega)]^n} p(x,u|y)$, where $p(x,u|y)$ is a Markov distribution obtained from the TMF $T=(X,U,Y)$, the distribution of which is defined on $(\Delta\times R)^n$ by (3.1), with

$$\varphi_c(t_c) = \varphi_c(x_c,u_c,y_c) = \begin{cases}\psi_c(u_c), & \text{for Card}(c)>1,\\ \psi_c(u_c)-\mathrm{Log}(p(y_s|x_s)), & \text{for } c=\{s\}\end{cases} \quad (3.4)$$

Proof. We have

$$M(x) = (M^0\oplus M^1)(x) \propto \sum_{x\in u}\exp\left[-\sum_{c\in C}\psi_c(u_c)\right]\prod_{s\in S} p(y_s|x_s)$$
$$= \sum_{x\in u}\exp\left[-\sum_{c\in C}\psi_c(u_c) + \sum_{s\in S}\log(p(y_s|x_s))\right]$$

In the sum above $x=(x_s)_{s\in S}$ is fixed and $u=(u_s)_{s\in S}$ varies in $[P(\Omega)]^n$ in such a way that $x_s\in u_s$ for each $s\in S$, which means that $u=(u_s)_{s\in S}$ varies in such a way that $(x,u)\in\Delta^n$. This means that $M(x)$ can be written as

$$M(x) \propto \sum_{u/(x,u,y)\in\Delta^n\times R^n}\exp\left[-\sum_{c\in C}\psi_c(u_c) + \sum_{s\in S}\log(p(y_s|x_s))\right]$$
$$= \sum_{u/(x,u,y)\in\Delta^n\times R^n}\exp[\varphi_c(x_c,u_c,y_c)]$$

which completes the proof.

□

Example 3.1.. One possible application of the HEMF above is inspired by the successful use of the similar hidden evidential Markov chains in the situations where the unknown process has unknown parameters and is not stationary [12]. Thus, let us consider the problem of segmenting an observed image $Y=y$ into two classes $\Omega=\{\omega_1,\omega_2\}$, and let us assume that the hidden class image $X=x$ is strongly heterogeneous and can hardly be modeled by a stationary Markov field. As HEMF is a particular TMF, the parameter estimation method proposed in [1] can be applied and thus unsupervised segmentation is workable. One can then compare two unsupervised methods: the classical HMF based method, and the new HEMF one.

Two examples of unsupervised segmentation performed with the classical HMF and the proposed HEMF are presented in Fig. 1. We consider two classes $\Omega=\{\omega_1,\omega_2\}$, and assume that both HMF and HEMF are Markovian with respect to the four nearest neighbors. Therefore, we have an EMF on $(\Lambda)^n$, with $\Lambda=P(\Omega)=\{\{\omega_1\},\{\omega_2\},\{\omega_1,\omega_2\}\}=\{\lambda_1,\lambda_2,\lambda_3\}$, the distribution of which is given by (3.3). We then consider a simple energy function given by the following potential functions. For horizontal cliques $c=(t,s)$, we take $\psi_c(\lambda_1,\lambda_2)=-\alpha_{1H}$, $\psi_c(\lambda_1,\lambda_3)=-\alpha_{2H}$, $\psi_c(\lambda_2,\lambda_3)=-\alpha_{3H}$, and $\psi_c(\lambda_1,\lambda_1)=\psi_c(\lambda_2,\lambda_2)=\psi_c(\lambda_3,\lambda_3)=0$. The same is done for vertical cliques, with α_{IV}, α_{2V} α_{3V} instead of α_{1H}, α_{2H}, and α_{3H}. Moreover, ψ_c is null for the cliques singletons.

According to Proposition 3.1, we have to consider the triplet Markov field defined by

$$\varphi_c(t_c) = \varphi_c(x_c,u_c,y_c) = \begin{cases}\psi_c(u_c), & \text{for } c=(s,t),\\ -\log(p(y_s|x_s)), & \text{for } c=\{s\}\end{cases}$$

where (x_s,u_s) are in $\Delta=\{(\omega_1,\{\omega_1\}),(\omega_1,\{\omega_1,\omega_2\}),(\omega_2,\{\omega_2\}),(\omega_2,\{\omega_1,\omega_2\})\}=\{\delta_1,\delta_2,\delta_3,\delta_4\}$. Thus, putting $v_s=(x_s,u_s)$, we have a standard hidden Markov field (V,Y) and for each $s\in S$, the probability $p(v_s|y)$ on $\Delta=\{\delta_1,\delta_2,\delta_3,\delta_4\}$ can be estimated by using the Gibbs sampler. The estimates obtained this way enable us to compute $p(x_s=\omega_1|y)=p(v_s=\delta_1|y)+p(v_s=\delta_2|y)$ and $p(x_s=\omega_2|y)=p(v_s=\delta_3|y)+p(v_s=\delta_4|y)$, which are then used to perform the Bayesian MPM segmentation. Concerning the classical hidden Markov field used, the potential functions are $\psi_c(\omega_1,\omega_2)=-\alpha_H$ for horizontal cliques, $\psi_c(\omega_1,\omega_2)=-\alpha_V$ for vertical cliques, $\psi_c(\omega_1,\omega_1)=\psi_c(\omega_2,\omega_2)=0$ for horizontal and vertical cliques, and ψ_c is null for the cliques singletons. The estimates of all parameters in both models are presented in Table 1, and the different images are presented in Fig. 1.

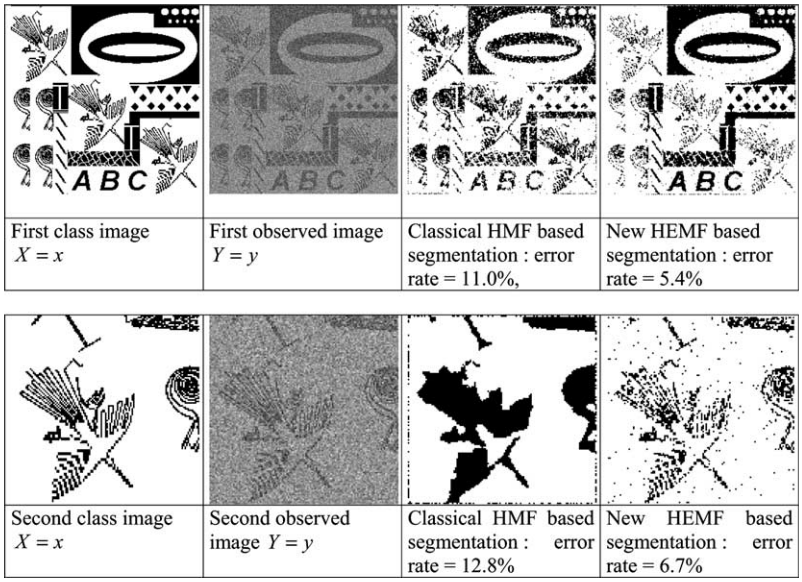

Fig. 1. Two class images corrupted by Gaussian noise $N_1(0,1)$ and $N_2(2,1)$, and unsupervised segmentation results based on classical HMF and new HEMF. The estimates, obtained with ICE, are presented in Table 1.

Let us mention a more general model, in which both $p(x)$ and $p(y|x)$ are Markov fields (see examples of such models in [1]). In particular, we can chose a model generalizing the classical model (3.2), in which $p(y|x)=\prod_{s\in S}p(y_s|x_s)$ is replaced with $p(y|x)=\gamma\exp\left[-\sum_{c\in C}\chi_c(x_c,y_c)\right]$. As above, we can then consider

$$M^1(x)=q^y(x)=\frac{\exp\left[-\sum_{c\in C}\chi_c(x_c,y_c)\right]}{\sum_{x\in\Omega^n}\exp\left[-\sum_{c\in C}\chi_c(x_c,y_c)\right]}$$

$$\propto\exp\left[-\sum_{c\in C}(x_c,y_c)\right]$$

which here is a Markov field.

Then we have the following result:

Proposition 3.2. Let M^0 be an EMF defined on $[P(\Omega)]^n$ by (3.3), and M^1 a classical Markov field defined over Ω^n from the observed process $Y=y\in R^n$ by $M^1(x)\propto\exp\left[-\sum_{c\in C}\chi_c(x_c,y_c)\right]$. Let $\Lambda=P(\Omega)$ and $\Delta\in\Omega\times\Lambda$ such that $(\omega,\lambda)\in\Delta$ if and only if $\omega\in\lambda$.

Then the probability distribution $M=M^0\oplus M^1$ is the conditional distribution $p(x|y)=p(x,y)/p(y)$, where $p(x,y)$ is a TMF. More precisely, $T=(X,U,Y)$ is a Markov field defined on $(\Delta\times R)^n$ by (3.1), with $\varphi_c(t_c)=\varphi_c(x_c,u_c,y_c)=\psi_c(u_c)+\chi_c(x_c,y_c)$.

The proof is analogous to the proof of Proposition 3.1 above.

4. Triplet Markov fields with evidential sensors

In this section we will consider that the prior information on the hidden field X is given by a probabilistic classical Markov distribution $p(x)=\gamma\exp\left[-\sum_{c\in C}\varphi_c(x_c)\right]$, and the information provided by the observations (one sensor Y_s or numerous sensors $Y_s^1,\ldots,Y_s^m$) can be 'evidential'. Let $\Omega=\{\omega_1,\ldots,\omega_k\}$ and $\Lambda=P(\Omega)$ as above. We will say that a sensor Y_s^j 'is sensitive' to

Table 1
Parameters estimates and error ratios of the unsupervised segmentation methods based on classical hidden Markov fields (HMF) and hidden evidential Markov fields (HEMF)

Parameters	Real	Image 1		Image 2	
		HMF	HEMF	HMF	HEMF
μ_1	0.00	-0.3	0.02	0.88	0.02
μ_2	2.00	1.86	2.03	2.01	2.02
σ_1	1.00	0.85	1.00	1.36	0.94
σ_2	1.00	1.05	0.97	0.99	0.67
α_{1H}, α_{1V}			0.25, 0.31		0.29, 0.33
α_{2H}, α_{2V}			0.38, 0.44		0.47, 0.55
α_{3H}, α_{3V}			0.01, 0.01		0.02, 0.02
α_H, α_V		0.65, 0.82		0.41, 0.50	
Error ratio		11.0%	5.40%	12.8%	6.7%

See Fig. 1 for visual results.

$B \in \Lambda$ if the knowledge t hat $x_s \in B$ modifies its probability law (see also Examples 2.1 and 2.2).

4.1. The case of one sensor

We first recall the extension of the classical HMF (3.2) which has been proposed in [3], and then we propose two new extensions: one for the HMF with spatially correlated noise, and another for the PMF. Therefore, let $\Omega = \{\omega_1, \ldots, \omega_k\}$, $\Lambda = P(\Omega)$ as above, and let $\Lambda_1 \subset \Lambda$ such that each Y_s is sensitive to the elements of Λ_1.

The classical HMF (3.2) will be called below 'HMF-IN' (IN for 'independent noise'), and we will call HMF a PMF such that the hidden field is a Markov one. Therefore, there are three models of increasing generality: HMF-IN, HMF, and PMF. The interest of such distinction will appear below.

So, the extension of HMF-IN, proposed in [3] and validated by different applications, is the following. In HMF-IN, the posterior distribution $p(x|y)$ is the DS fusion $p(x|y) = (p \oplus q^y)(x)$ of $p(x) = \gamma \exp\left[-\sum_{c\in C} \varphi_c(x_c)\right]$ with $q^y(x) \propto \prod_{s\in S} p(y_s|x_s)$. As the considered sensor is only sensitive to the elements of Λ_1, let us consider $q^{y,*}(x^*) \propto \prod_{s\in S} p(y_s|x_s^*)$, where $x^* = (x_s^*)_{s\in S}$ and each x_s^* varies in Λ_1. Then one can see that $p(x|y) = (p \oplus q^{y*})(x)$ remains a Markov field, written as

$$p(x|y) = \gamma\prime \exp\left[-\sum_{c\in C} \varphi_c(x_c) + \sum_{s\in S} \mathrm{Log}\left[\sum_{x_s \in x_s^*} p(y_s|x_s^*)\right]\right] \tag{4.1}$$

where the sum $\sum_{x_s\in x_s^*} p(y_s|x_s^*)$ is taken over x_s^*, x_s being fixed.

To consider the first new extension, let $p^*(x^*, y) = \gamma \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right]$ be an HEMF distribution on $(\Lambda_1)^n \times R^n$, which means that its marginal distribution $p^*(x^*)$ is a Markov one: $p^*(x^*) = \gamma' \exp\left[-\sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right]$. As a consequence, $p^*(y|x^*)$ is written

$$p^*(y|x^*) = \gamma\prime\prime \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c) + \sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right] \tag{4.2}$$

Then (4.2) defines

$$q^{y,*}(x^*) = \left(\frac{p^*(y|x^*)}{\sum_{x^*\in(\Lambda_1)^n} p^*(y|x^*)}\right) \propto p(y|x^*)$$

and one sees from (4.2) that $q^{y,*}(x^*) = \gamma(y)\exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c) + \sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right]$ is a Markov field. Then fusing $p(x) = \gamma \exp\left[-\sum_{c\in C} \varphi_c(x_c)\right]$ with $q^{y,*}(x^*)$ is a probability which extends the classical posterior $p(x|y)$ (which is obtained again for Λ_1 reduced to singletons). Therefore, we can state:

Proposition 4.1. Let us consider:

(i) a classical probabilistic Markov field M^0 on Ω^n, defined by $M^0(x) = \gamma \exp\left[-\sum_{c\in C} \varphi_c(x_c)\right]$, which classically models a prior information;
(ii) an HEMF distribution $p^*(x^*, y) = \gamma \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right]$, which means that the marginal distribution $p^*(x^*)$ is a Markov one: $p^*(x^*) = \gamma\prime \exp\left[-\sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right]$; and
(iii) the bba M^1 defined on $(\Lambda_1)^n$, for fixed $y \in R^n$, by $p^*(y|x^*)$ given by (4.2);
(iv) and the set $\Delta = \{(\omega, A) \in \Omega \times \Lambda_1 | \omega \in A\}$.

Then the DS fusion $(M^0 \oplus M^1)(x)$ is the probability $p(x|y)$, which is the marginal probability of $p(x, x^*|y)$ defined from the distribution of the TMF $T = (X, X^*, Y)$ given on $\Delta^n \times R^n$ by

$$p(x, x^*, y) \propto \exp\left[-\sum_{c\in C} \varphi_c(x_c) - \sum_{c\in C} \varphi_c^*(x_c^*, y_c) + \sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right] \tag{4.3}$$

As a consequence, different Bayesian segmentation techniques can be implemented.

Proof. We have

$$\begin{aligned}
(M^0 \oplus M^1)(x) &\propto \sum_{x\in x^*}\left[\exp\left[-\sum_{c\in C} \varphi_c(x_c)\right] \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c) + \sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right]\right] \\
&= \sum_{x\in x^*} \exp\left[-\sum_{c\in C}\left[\varphi_c(x_c) + \varphi_c^*(x_c^*, y_c) - \sum_{c\in C} \varphi_c^{\prime *}(x_c^*)\right]\right] \\
&= \sum_{x\in x^*} p(x, x^*, y)
\end{aligned}$$

which completes the proof. □

To consider the second new extension, let $p^*(x^*) = \gamma \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right]$ be a pairwise evidential Markov field (PEMF) distribution. $p^*(x^*, y)$ is not necessarily an HEMF and thus the form of the marginal distribution $p^*(x^*)$ is not necessarily known. However, $p^*(x^*|y) \propto \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right]$ is Markovian one. We can then develop all calculations above concerning the first new example with $p^*(x^*|y)$ instead of $q^{y,*}(x^*)$. Finally, we can state:

Proposition 4.2. Let us consider:

(i) a classical probabilistic Markov field M^0 on Ω^n defined by $M^0(x) = \gamma \exp\left[-\sum_{c\in C} \varphi_c(x_c)\right]$, and modeling some 'prior' information;
(ii) a PEMF distribution $p^*(x^*, y) = \gamma \exp\left[-\sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right]$;
(iii) the bba M^1 defined on $(\Lambda_1)^n$ by $p^*(x|y)$;
(iv) the set $\Delta = \{(\omega, A) \in \Omega \times \Lambda_1 | \omega \in A\}$.

Then the DS fusion $(M^0 \oplus M^1)(x)$ is the probability $p(x|y)$, which is the marginal probability of $p(x, x^*|y)$ defined from the distribution of the TMF $T = (X, X^*, Y)$ given on $\Delta^n \times R^n$ by

$$p(x, x^*, y) \propto \exp\left[-\sum_{c\in C} \varphi_c(x_c) - \sum_{c\in C} \varphi_c^*(x_c^*, y_c)\right] \tag{4.4}$$

As a consequence, different Bayesian segmentation techniques can be implemented.

Let us notice that an HEMF is also a PEMF, and thus when considering an HEMF we have the choice between two different ways of using some 'prior' information modeled by a classical probabilistic Markov field.

4.2. The case of numerous sensors

Let us consider r sensors: $Y_s=(Y_s^1,\dots Y_s^r)$. We take $r=2$ for the sake of simplicity (the extension to $r>2$ is immediate, see also the next section). So, we have $Y_s=(Y_s^1,Y_s^2)$ and $y_s=(y_s^1,y_s^2)$ denote its realization. We will assume that the sensors are independent, which means that in the classical case that we are going to generalize, the random fields Y^1, Y^2 are independent conditionally on X.

As above, let us consider the same $\Omega=\{\omega_1,\dots,\omega_k\}$, $\Lambda=P(\Omega)$, and $\Lambda_1\subset\Lambda$, $\Lambda_2\subset\Lambda$ such that each Y_s^1 is sensitive to the elements of Λ_1 and each Y_s^2 is sensitive to the elements of Λ_2. Furthermore, let $\Lambda_{1,2}\subset\Lambda$ be the set of sets $A\subset\Lambda$ such that there exists $A_1\subset\Lambda_1$ and $A_2\subset\Lambda_2$ verifying $A=A_1\cap A_2$.

As in the previous subsection, let us first recall the extension of the classical bi-sensor HMF-IN proposed in [3]. The first sensor produces $q^{y^1,*}(x^*)\propto\prod_{s\in S}p(y_s^1|x_s^*)$, the second produces $q^{y^2,*}(x^*)\propto\prod_{s\in S}p(y_s^2|x_s^*)$, and the information provided by the observation $(Y^1,Y^2)=(y_1,y_2)$ is given on $(\Lambda_{1,2})^n$ by

$$q^{(y^1,y^2),*}(x^*)=(q^{y^1,*}\oplus q^{y^2,*})(x^*)\propto\prod_{s\in S}\left[\sum_{x_s^{*,1}\cap x_s^{*,2}=x_s^*}p(y_s^1|x_s^{*,1})p(y_s^2|x_s^{*,2})\right]\tag{4.5}$$

Then $p(x|y)$ is given by a formula similar to (4.1), obtained by replacing $p(y_s|x_s^*)$ with $\sum_{x_s^{*,1}\cap x_s^{*,2}=x_s^*}p(y_s^1|x_s^{*,1})p(y_s^2|x_s^{*,2})$. We see that the sensors can be fused 'pixel by pixel', which is the reason why there is no need for triplet Markov fields when the extensions of HMF-IN are considered.

Let us now consider the new case of spatially correlated sensors. As above, we have two cases: (X^*,Y) is an HEMF with known Markovian $p(x^*)$, or (X^*,Y) is a PEMF with unknown $p(x^*)$. In the first case we apply (4.2) to each sensor to get $q^{y^{1,*}}(x)$ and $q^{y^2,*}(x^*)$, and we have the following proposition.

Proposition 4.3. Let M^0 be the Markov field on Ω^n defined by $M^0(x)=\gamma\exp\left[-\sum_{c\in C}\varphi_c(x_c)\right]$, and let M^1, M^2 be bba's defined on $(\Lambda_1)^n$ and $(\Lambda_2)^n$, for fixed $y\in R^n$, *by* $q^{y^1,*}(x^*)$ *and* $q^{y^2,*}(x^*)$. *Let* $\Delta=\{(\omega,A_1,A_2)\in\Omega\times\Lambda_1\times\Lambda_2|\omega\in A_1\cap A_2\}$.

Then the DS fusion $(M^0\oplus M^1\oplus M^2)(x)$ is the probability $p(x|y)$ defined by the TMF $T=(X,X^{*,1},X^{*,2},Y)$ whose distribution on $\Delta^n\times R^n$ is given by

$$p(x,x^{*,1},x^{*,2},y)\propto\exp\left[-\sum_{c\in C}\varphi_c(x_c)-\sum_{c\in C}\varphi_c^{*,1}(x_c^{*,1},y_c)+\sum_{c\in C}\varphi_c'^{*,1}(x_c^{*,1})-\sum_{c\in C}\varphi_c^{*,2}(x_c^{*,2},y_c)+\sum_{c\in C}\varphi_c'^{*,2}(x_c^{*,2})\right]\tag{4.6}$$

as a consequence, different Bayesian segmentation techniques can be implemented.

The proof is similar to the proof of Proposition 4.1.

This second case is quite similar to the second case of the previous subsection: we obtain a proposition analogous to Proposition 4.2 above, with the only difference that in $p^*(x^*|y)$ we have $y=(y^1,y^2)$.

5. General model

We assumed in the previous sections that either the prior information, or the observation provided by sensors, were probabilistic and modeled by some Markov fields. This enabled us to see how different classical models can be successively extended to more and more complex—or simply different—situations. This section is devoted to an ultimate extension:

Let us consider r sensors: $Y=(Y_1,\dots,Y^r)$, and $Y^i=(Y_s^i)_{s\in S}$ for $i=1,\dots,r$. Let M^0 be a prior EMF defined on $(\Lambda_0)^n$ by $M^0(x^*)=\gamma_0\exp\left[-\sum_{c\in C}\varphi_c^0(x_c^*)\right]$, and for $i=1,\dots,r$ let M^i be the EMF defined on $(\Lambda_i)^n$, for fixed y^i, by a PEMF $M^i(x^{*,i})=\gamma_i\exp\left[-\sum_{c\in C}\varphi_c^i(x_c^{*,i},y^i)\right]$. Let us consider Λ the set of all $A\subset\Omega$ such that there exists at least one $(A_0,A_1,\dots,A_r)\in\Lambda_0\times\Lambda_1\times\cdots\times\Lambda_r$ for which $A=A_0\cap A_1\cap\cdots\cap A_r$. Then the DS fusion gives $M=M_0\oplus M_1\oplus\cdots\oplus M^r$ defined on Λ^n by

$$M(x^*)\propto\sum_{(x^{*,0},x^{*,1},\dots,x^{*,r})}1_{[x^*=x^{*,0}\cap x^{*,1}\cap\dots\cap x^{*,r}]}\exp\left[-\sum_{c\in C}\varphi_c^i(x_c^{*,0})-\sum_{1\le i\le r}\left[\sum_{c\in C}\varphi_c^i(x_c^{*,i},y^i)\right]\right]\tag{5.1}$$

and thus, as in the particular cases above, M can be interpreted as a marginal distribution defined on Λ^n of the EMF defined on $(\Delta)^n=\Lambda^n\times(\Lambda_0)^n\times(\Lambda_1)^n\times\cdots\times(\Lambda_r)^n$ by

$$M'(x^*,x^{*,0},x^{*,1},\dots,x^{*,r})\propto1_{[x^*=x^{*,0}\cap x^{*,1}\cap\cdots\cap x^{*,r}]}\exp\left[-\sum_{c\in C}\varphi_c^i(x_c^{*,0})-\sum_{1\le i\le r}\left[\sum_{c\in C}\varphi_c^i(x_c^{*,i},y^i)\right]\right]\tag{5.2}$$

Let us notice that, on the contrary of the two previous sections, the result $M(x^*)$ is not necessarily a probability measure. However, it is still possible to perform statistical segmentation using the so-called 'maximum of plausibility' principle. More precisely, once $M(x_s^*)$ is estimated, we compute the plausibility of each $x_s=\omega\in\Omega$ by $\mathrm{Pl}(x_s=\omega)=\sum_{\omega\in x_s^*}M(x_s^*)$, and then the estimated $\hat{x}=(\hat{x}_s)_{s\in S}$ is given by $\hat{x}_s=\arg\max_\omega\mathrm{Pl}(x_s=\omega)$.

Example 5.1. Let us consider an example illustrating the interest of the general model (5.2) and its use in a concrete situation. Imagine a satellite image representing a scene containing a river (ω_1), a sea (ω_2), urban area (ω_3), and forest (ω_4). Imagine that we have some prior knowledge about the probabilistic distribution of 'water', which is $\{\omega_1,\omega_2\}$, and 'land', which is $\{\omega_3,\omega_4\}$, and that this distribution is a Markov field. Thus we have $\Lambda_0=\{\{\omega_1,\omega_2\},\{\omega_3,\omega_4\}\}=\{\lambda_1^0,\lambda_2^0\}$ and M^0 is an EMF defined on $(\Lambda_0)^n$, where n is the number of pixels. Furthermore, there are two sensors: an optical sensor Y^1, and an infrared sensor Y^2. The optical sensor cannot see any difference between the river (ω_1) and the sea (ω_2), and there are some clouds hiding part of the scene. Thus this sensor is sensitive to $\Lambda_1=\{\{\omega_1,\omega_2\},\{\omega_3\},\{\omega_4\},\Omega\}=\{\lambda_1^1,\lambda_2^1,\lambda_3^1,\lambda_4^1\}$. The infrared sensor mainly detects temperature differences and can only detect a difference between the urban area and other classes; thus it is sensitive to $\Lambda_2=\{\{\omega_3\},\{\omega_1,\omega_2,\omega_4\}\}=\{\lambda_1^2,\lambda_2^2\}$. The EMFs M^1 and M^2 can then be searched as follows. The observation $Y^1=y^1$ is considered as the observation of a classical probabilistic hidden or pairwise Markov field (U^1,Y^1), where $U^1=(U_s^1)_{s\in S}$ and each U_s^1 takes its values in Λ_1. Let us assume that (U^1,Y^1) is pairwise Markov field, with possibly unknown distribution of U^1. The distribution of (U^1,Y^1) is written $p(u^1,y^1)=\gamma^1\exp\left[-\sum_{c\in C}\varphi_c^1(u_c^1,y_c^1)\right]$, and the corresponding parameters can be estimated from $Y^1=y^1$ by the method proposed in [1]. Once these parameters estimated, we have Markovian $p(u^1|y^1)$, which is M^1. Then M^2 is found in a similar way. Given the forms of Λ_0, Λ_1, and Λ_2 we can see that $\Lambda=\{\{\omega_1,\omega_2\},\{\omega_3\},\{\omega_4\}\}=\{\lambda_1,\lambda_2,\lambda_3\}$. Finally, according to (5.2), we have to consider the subset $\Delta\subset\Lambda\times\Lambda_0\times\Lambda_1\times\Lambda_2$ such that $(A,A_0,A_1,A_2)\in\Delta$ is equivalent to $A=A_0\cap A_1\cap A_2$. Recalling that

$$\begin{aligned}\Lambda_0&=\{\{\omega_1,\omega_2\},\{\omega_3,\omega_4\}\}=\{\lambda_1^0,\lambda_2^0\};\\ \Lambda_1&=\{\{\omega_1,\omega_2\},\{\omega_3\},\{\omega_4\},\Omega\}=\{\lambda_1^1,\lambda_2^1,\lambda_3^1,\lambda_4^1\};\\ \Lambda_2&=\{\{\omega_3\},\{\omega_1,\omega_2,\omega_3\}\}=\{\lambda_1^2,\lambda_2^2\}\end{aligned}\tag{5.3}$$

We see that Δ is the set of the following elements: $(\lambda_1,\lambda_1^0,\lambda_1^1,\lambda_2^2)$, $(\lambda_1,\lambda_1^0,\lambda_4^1,\lambda_2^2)$, $(\lambda_2,\lambda_2^0,\lambda_2^1,\lambda_1^2)$, $(\lambda_2,\lambda_2^0,\lambda_4^1,\lambda_1^2)$, $(\lambda_3,\lambda_2^0,\lambda_3^1,\lambda_2^2)$.

Finally, we have an EMF M' defined on Δ^n with (5.2). Then sampling realizations of M', we estimate $M(x_s^*)$, which is a bba on $\Lambda=\{\{\omega_1,\omega_2\},\{\omega_3\},\{\omega_4\}\}=\{\lambda_1,\lambda_2,\lambda_3\}$. The plausibility on $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$ is computed for each $x_s=\omega\in\Omega$ by $\text{Pl}(x_s=\omega_1)=\text{Pl}(x_s=\omega_2)=M(x_s^*=\{\omega_1,\omega_2\})$, $\text{Pl}(x_s=\omega_3)=M(x_s^*=\{\omega_3\})$, and $\text{Pl}(x_s=\omega_4)=M(x_s^*=\{\omega_4\})$, which is then used to estimate $\hat{x}=(\hat{x}_s)_{s\in S}$ by $\hat{x}_s=\text{argmax}_\omega\,\text{Pl}(x_s=\omega)$. Concerning the case of correlated sensors, the adaptation of the modeling proposed in [22] to the general models (5.1) and (5.2) does not pose particular difficulties. Moreover, extending to such new models of different classical parameter estimation methods, as for example the method proposed in [20], could be viewed.

6. Conclusion and perspectives

The aim of this paper was to study the different possibilities of using the Dempster–Shafer theory of evidence in multisensor Markov fields context. Using the recent Triplet Markov fields (TMF) model, we showed how different Dempster–Shafer fusions, which generalize the classical calculation of the posterior distributions, can be performed. The latter allow one to propose Bayesian segmentation methods, which are then workable in more general settings. Moreover, different model parameters can be estimated with the general 'Iterative Conditional Estimation' (ICE [1]), whose relationship to the well known 'Expectation-Maximization' method (EM [16]) is described in [4]. Some examples of real situations in which the new models are of interest have been provided, likewise some experiments involving unsupervised image segmentation.

Hyperspectral data analysis [14,35], which at present is an active field of investigation, or even 3D Markov models for image analysis ([30], among others) could possibly be areas in which the fusion techniques proposed in this paper could be applied.

References

[1] D. Benboudjema, W. Pieczynski, Unsupervised image segmentation using triplet Markov fields, Computer Vision and Image Understanding 99 (3) (2005) 476–498.
[2] D. Benboudjema, W. Pieczynski, Segmenting non stationary images with triplet Markov fields, International Conference on Image Processing (ICIP 2005), Geneva, Italy, September 11–14, 2005.
[3] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, W. Pieczynski, Multisensor images segmentation using Dempster-Shafer fusion in Markov fields context, IEEE Transactions on Geoscience and Remote Sensing 39 (8) (2001) 1789–1798.
[4] J.-P. Delmas, An equivalence of the EM and ICE algorithm for exponential family, IEEE Transaction on Signal Processing 45 (10) (1997) 2613–2615.
[5] T. Denoeux, Reasoning with imprecise belief structures, International Journal of Approximate Reasoning 20 (1) (1999) 79–111.
[6] S. Foucher, M. Germain, J.-M. Boucher, G.B. Benié, Multisource classification using ICM and Dempster–Shafer theory, IEEE Transactions on Instrumentation and Measurement 51 (2) (2002) 277–281.
[7] L. Fouque, A. Appriou, W. Pieczynski, An evidential markovian model for data fusion and unsupervised image classification, Proceedings of 3rd International Conference on Information Fusion, FUSION 2000, July 10th–13th, Paris, France, vol. 1, 2000, pp. TuB4-25–TuB4-31.
[8] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions and the Bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (1984) 721–741.
[9] J. Guan, D.A. Bell, Evidence Theory and Its Applications, North-Holland, Amsterdam, 1991.
[10] Zhu Hongwei, O. Basir, An adaptive fuzzy evidential nearest neighbor formulation for classifying remote sensing images, IEEE Transactions on Geoscience and Remote Sensing 43 (8) (2005) 1874–1889.
[11] F. Janez, A. Appriou, Theory of evidence and non-exhaustive frames of discernment: plausibilities correction methods, International Journal of Approximate Reasoning 18 (1–2) (1998) 1–19.
[12] P. Lanchantin, W. Pieczynski, Unsupervised restoration of hidden non stationary Markov chain using evidential priors, IEEE Transactions on Signal Processing 53 (8) (2005) 3091–3098.
[13] P. Lanchantin, W. Pieczynski, Chaînes et arbres de Markov évidentiels avec applications à la segmentation des processus non stationnaires, Traitement du Signal 22 (1) (2005) 15–26.

[14] D. Landgrebe, Hyperspectral image data analysis, IEEE Signal Processing Magazine 19 (1) (2002) 17–28.
[15] S. Le, S. Le Hégarat-Mascle, I. Bloch, D. Vidal-Madjar, Introduction of neighborhood information in evidence theory and application to data fusion of radar and optical images with partial cloud cover, Pattern Recognition 31 (11) (1998) 1811–1823.
[16] G.J. McLachlan, T. Krishnan, EM algorithm and extensions, Series in Probability and Statistics, Wiley, New York, 1997.
[17] D. Melas, S.P. Wilson, Double Markov random fields and bayesian image segmentation, IEEE Transactions on Signal Processing 50 (2) (2002) 357–365.
[18] P. Pérez, Markov random fields and images, CWI Quarterly 11 (4) (1998) 413–437.
[19] S. Perrin, E. Duflos, P. Vanheeghe, A. Bibaut, Multisensor fusion in the frame of evidence theory for landmines detection, IEEE Transactions on Systems Man and Cybernetics 34 (4) (2004) 485–498 Part C.
[20] W. Pieczynski, J. Bouvrais, C. Michel, Estimation of generalized mixture in the case of correlated sensors, IEEE Transactions on Image Processing 9 (2) (2000) 308–311.
[21] W. Pieczynski, A.-N. Tebbache, Pairwise Markov random fields and segmentation of textured images, Machine Graphics & Vision 9 (3) (2000) 705–718.
[22] W. Pieczynski, Unsupervised Dempster–Shafer fusion of dependent sensors, Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI'2000), Austin, Texas, United States, 2–4 April 2000, pp. 247–251.
[23] W. Pieczynski, Pairwise Markov chains, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (5) (2003) 634–639.
[24] W. Pieczynski, D. Benboudjema, P. Lanchantin, Statistical image segmentation using Triplet Markov fields, SPIE's International Symposium on Remote Sensing, September 22–27, Crete, Greece, 2002.
[25] W. Pieczynski, C. Hulard, T. Veit, Triplet Markov Chains in hidden signal restoration, SPIE's International Symposium on Remote Sensing, September 22–27, Crete, Greece, 2002.
[26] W. Pieczynski, Fusion de Dempster–Shafer dans les chaînes triplet partiellement de Markov—Dempster-Shafer fusion in triplet partially Markov chains, Comptes Rendus de l'Académie des Sciences—Mathématique, Série I 339 (11) (2004) 797–802.
[27] W. Pieczynski, Modeling non stationary hidden semi-Markov chains with triplet Markov chains and theory of evidence, Statistical Signal Processing (SSP2005), Bordeaux, France, July 2005.
[28] W. Pieczynski, Triplet markov chains and theory of evidence, International Journal of Approximate Reasoning, Submitted for publication.
[29] J.A. Richards, Analysis of remotely sensed data: the formative decades and the future, IEEE Transactions on Geoscience and Remote Sensing 43 (3) (2005) 422–432.
[30] S. Ruan, C. Jaggi, J. Xue, J. Fadili, D. Bloyet, Brain tissue classification of magnetic resonance images using partial volume modeling, IEEE Transactions on Medical Imaging 19 (12) (2000) 1179–1187.
[31] A. Sarkar, A. Banerjee, N. Banerjee, S. Brahma, B. Kartikeyan, M. Chakraborty, K.L. Majumder, Landcover classification in MRF context using Dempster–Shafer fusion for multisensor imagery, IEEE Transactions on Image Processing 14 (5) (2005) 634–645.
[32] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
[33] P. Smets, Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem, International Journal of Approximate Reasoning 9 (1993) 1–35.
[34] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191–234.
[35] S.M. Schweizer, J.M.F. Moura, Efficient detection in hyperspectral imagery, IEEE Transactions on Image Processing 10 (4) (2001) 584–597.
[36] F. Tupin, H. Maitre, I. Bloch, A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors, IEEE Transactions on Geoscience and Remote Sensing 37 (3) (1999) 1327–1343.
[37] G. Winkler, Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction, Springer, Berlin, 2003.
[38] R.R. Yager, J. Kacprzyk, M. Fedrizzi, Advances in the Dempster–Shafer Theory of Evidence, Wiley, New York, 1994.