UNSUPERVISED SEGMENTATION OF NONSTATIONARY PAIRWISE MARKOV CHAINS USING EVIDENTIAL PRIORS

Mohamed El Yazid Boudaren [1, 2], *Emmanuel Monfrini* [2], *Wojciech Pieczynski* [2]

1. Ecole Militaire Polytechnique, Laboratoire Mathématiques Appliquées, Algiers 16111, ALGERIA
2. Institut Mines-Telecom, TELECOM SudParis, Dpt. CITI, UMR 5157, Evry 91011, FRANCE

ABSTRACT

Hidden Markov models have been widely used to solve some inverse problems occurring in image and signal processing. These models have been recently generalized to pairwise Markov chains, which present higher modeling capabilities with comparable computational complexity. To be applicable in the unsupervised context, both models assume the data of interest stationary. When these latter are actually stationary, the models yield satisfactory results thanks to some Bayesian techniques such as MPM and MAP. However, when the data are nonstationary, they fail to establish an appropriate link with the data and the obtained results are quite poor. One interesting way to overcome this drawback is to use the Dempster-Shafer theory of evidence by introducing a mass function to model the lack of knowledge of the a priori distributions of the hidden data to be recovered. It has been shown that the use of such theory in the hidden Markov chains context yields significantly better results than those provided by the standard models. The aim of this paper is to apply the same theory in the pairwise Markov chains context to deal with nonstationary data hidden with correlated noise. We show that MPM restoration of data remains workable thanks to the triplet Markov models formalism. We also provide the corresponding parameters estimation in the unsupervised context. The new evidential model is then assessed through experiments conducted on synthetic and real images.

Index Terms— Hidden Markov chains, pairwise Markov chains, triplet Markov chains, theory of evidence, nonstationary data

1. INTRODUCTION

Let us consider two sequences of random variables: $X = (X_n)_{n=1}^{N}$, an unobservable process that takes its values from a finite set of classes $\Omega = \{\omega_1, .., \omega_K\}$, and $Y = (Y_n)_{n=1}^{N}$, with $Y_n \in \mathbb{R}$, an observable noisy process obtained from X in some way. The aim then is to recover X from Y. Realization of such processes will be denoted by lowercase letters. We will consider this example all along the paper to describe the different models and their corresponding formalisms. The estimator that will be used in this framework is the MPM estimator which is given by the following formula:

$$[\hat{x} = \hat{s}_{MPM}(y)] \Leftrightarrow \left[\hat{x}_n = \underset{\omega}{\operatorname{argmax}}\, p(x_n = \omega | y)\right] \quad (1)$$

Hidden Markov chains (HMCs) have been extensively used to solve various inverse problems occurring in a wide range of fields covering signal and image processing [1-3], and communications [3] among others. Let us also mention [4-5] as pioneering papers.

According to these models, the probability of observing the whole data is given by:

$$p(x, y) = p(x_1)p(y_1|x_1)\prod_{n=2}^{N} p(x_n|x_{n-1})p(y_n|x_n) \quad (2)$$

The MPM restoration can then be achieved thanks to the possibility of recursive computations of forward probabilities $\alpha_n(x_n) = p(y_1, .., y_n, x_n)$ and backward probabilities $\beta_n(x_n) = p(y_{n+1}, .., y_N|x_n)$. The posterior distributions required to achieve the MPM estimation of (1) are then given by:

$$p(x_n = \omega|y) = \alpha_n(\omega)\, \beta_n(\omega) \quad (3)$$

Moreover, when the distributions $p(x_n|x_{n-1})$ do not depend on n, the MPM restoration can be achieved in the unsupervised context thanks to some parameters estimation algorithms such as Expectation- Maximization (EM) [6] and Iterative Conditional Estimation (ICE) [7].

However, these algorithms may become inefficient when the distributions $p(x_n|x_{n-1})$ depend on n. We deal then with nonstationary data.

To overcome this drawback, *Lanchantin* and *Pieczynski* [8] introduce an evidential mass function to model the varying $p(x_n|x_{n-1})$. In fact, the nonstationary aspect of the *a priori* distributions has been assimilated to an imprecision or a lack of knowledge within these latter. The corresponding model is called evidential hidden Markov chain and it turned out that it outperforms the standard HMC model. Evidential Markov models were also used in [9, 10].

Let us notice that another way of dealing with varying *a priori* distributions is to apply the switching hidden Markov chains [11]. However, the use of this model subsumes the knowledge of the number of the stationary parts constituting the data.

HMCs have been generalized to pairwise Markov chains (PMCs) [7]. In PMC, one directly assumes the Markovianity of $Z = (X, Y)$, (2) becomes:

$$p(z) = p(z_1) \prod_{n=2}^{N} p(z_n | z_{n-1}), \quad (4)$$

and the HMC can then be seen as a particular PMC where $p(z_n|z_{n-1}) = p(x_n|x_{n-1})p(y_n|x_n)$, whereas in general PMC we have $p(z_n|z_{n-1}) = p(x_n|x_{n-1}, y_{n-1})$ $p(y_n|x_{n-1}, y_{n-1}, x_n)$. The superiority of PMCs over HMCs relies in the fact that he process X is no longer necessarily Markovian. On the other hand, MPM restoration and parameters estimation according to PMC paradigm remain workable in similar manner as in HMC context.

The aim of this paper is to propose an evidential pairwise Markov chain (EPMC) to model nonstationary data corrupted with correlated noise, which generalizes the evidential hidden Markov chain [8].

The remainder of the paper is organized as follows: section 2 describes the proposed EPMC model and provides its corresponding restoration and parameters estimation algorithms. In section 3, the model performance is assessed through experiments conducted on synthetic and real nonstationary images. Finally, we end the paper with some concluding remarks and possible future improvements.

2. EVIDENTIAL PAIRWISE MARKOV CHAINS

In this section, we describe the proposed EPMC and its corresponding formalism. For this purpose, we first introduce the nonstationary PMC model and show how the new model can take into account the nonstationary aspect of the data to be modeled.

2.1. Nonstationary PMC

Let us consider the example described in the previous section. we can write:

$$p(z_n, z_{n+1}) = p(x_n, x_{n+1})\, p(y_n, y_{n+1} | x_n, x_{n+1}) \quad (5)$$

When we deal with the stationary PMC, the distributions $p(x_n, x_{n+1})$ do not depend on n, and the model is fully defined through the distribution:

$$p(z_1, z_2) = p(i, j)\, f_{i,j}(y_1, y_2), \quad (6)$$

where $p(i, j)$ is a probability on Ω^2, and $f_{i,j}(y_1, y_2)$ are distributions in $\mathbb{R}^2$.

Let us assume now that the distributions $p(x_n, x_{n+1})$ depend on n. (6) becomes:

$$p(z_n, z_{n+1}) = p_n(i, j)\, f_{i,j}(y_n, y_{n+1}) \quad (7)$$

Hence, the parameters estimation procedures, such as EM and ICE, applied on such data considered as stationary, will provide a common value $p^*(i, j)$ of the distribution $p_n(i, j)$.

Another alternative is to use an evidential mass function to model the lack of precision in p^*like authors did in the hidden Markov chains case [8].

2.2. Dempster- Shafer theory of evidence

In this section, we give an overview about the so called theory of evidence introduced by Dempster in the 1960s and reformulated by Shafer in the 1970s [12]. Let us consider a frame of discernments $\Omega = \{\omega_1, \omega_2\}$ and let us consider the set of all the subsets of this latter $P(\Omega) = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \Omega\}$. A mass function is m is a function from $P(\Omega)$ to $\mathbb{R}^+$ that fulfills that following:

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in P(\Omega)} m(A) = 1 \end{cases} \quad (8)$$

Let us now consider a family of probability distributions $(p_\theta)_{\theta \in \Theta}$ defined on $\Omega = \{\omega_1, \omega_2\}$, and let us define the following "lower" probability $\breve{p}(\omega_n) = inf_{\theta \in \Theta}\, p_\theta(\omega_n)$. Hence, m defined by $m(\{\omega_1\}) = \breve{p}(\omega_1)$, $m(\{\omega_2\}) = \breve{p}(\omega_2)$ and $m(\{\omega_1, \omega_2\}) = 1 - \breve{p}(\omega_1) - \breve{p}(\omega_2)$ is a mass function and the latter quantity models the "lack of precision" in the exact probability p.

A mass m defined on $P(\Omega^N)$ is said to be an evidential Markovian chain (EMC) if it vanishes outside $[P(\Omega)]^N$ and if it can be written:

$$m(A_1, A_2, .., A_N) = m(A_1)m(A_2|A_1), .., m(A_N|A_{N-1}) \quad (9)$$

Hence, it can be seen as a generalization of the standard Markov chain and will be used instead of it within the PMC formalism to take the nonstationary aspect of the model into account.

2.3. Evidential PMC

Let m_1 be an EMC whose distribution is given by (9) and let $m_2(x_1, x_2, .., x_N) = \frac{p(y_1, y_2|x_1, x_2) \ldots p(y_{N-1}, y_N|x_{N-1}, x_N)}{p(y_2|x_2) \ldots p(y_{N-1}|x_{N-1})}$ be the Bayesian distribution defined by the observation $y = (y_1, y_2, \ldots, y_N)$. Let us consider the triplet $T = (U, X, Y)$ defined on $P(\Omega) \times \Omega \times \mathbb{R}$, and whose distribution is given by:

$$p(t_1, \ldots, t_N) \propto q_1(t_1, t_2) q_2(t_2, t_3) \ldots q_{N-1}(t_{N-1}, t_N) \quad (10)$$

where for $3 \leq n \leq N$, we have:

$$\begin{cases} q_1(t_1, t_2) = 1_{[x_1 \in u_1]} 1_{[x_2 \in u_2]} m_1(u_1) m_1(u_2|u_1) \frac{p(y_1, y_2|x_1, x_2)}{p(y_2|x_2)} \\ q_{n-1}(t_{n-1}, t_n) = 1_{[x_n \in u_n]} m_1(u_n|u_{n-1}) \frac{p(y_{n-1}, y_n|x_{n-1}, x_n)}{p(y_n|x_n)} \end{cases}$$

Such a model is a particular triplet Markov chain (TMC) [13]. The Dempster-Shafer fusion of m_1 and m_2 gives the posterior distribution $p(x|y)$ defined by $p(x, y)$, which is, itself, the marginal distribution of the TMC $T = (U, X, Y)$. According to the TMC formalism, the distributions $p(x_n, u_n|y)$ are workable, and the computational complexity of the processing is linear to the size of the data [13].

2.4. MPM restoration of EPMCs

Let us consider the model $T = (U, X, Y)$ defined previously. We define the generalized Forward probabilities $\alpha(x_n, u_n) = p(y_1, \dots, y_n, x_n, u_n)$ and Backward probabilities $\beta(x_n, u_n) = p(y_{n+1}, \dots, y_N|y_n, x_n, u_n)$ that can be computed iteratively as follows:

$$\alpha_1(x_1, u_1) = p(y_1, x_1, u_1),$$
$$\alpha_n(x_n, u_n) = \sum_{x_{n-1}, u_{n-1}} \alpha_{n-1}(x_{n-1}, u_{n-1}) p(t_n|t_{n-1}) \quad (11)$$

$$\beta_N(x_N, u_N) = 1,$$
$$\beta_n(x_n, u_n) = \sum_{x_{n+1}, u_{n+1}} \beta_{n+1}(x_{n+1}, u_{n+1}) p(t_{n+1}|t_n) \quad (12)$$

where

$$p(t_n, t_{n+1}) \propto 1_{[x_n \in u_n]} 1_{[x_{n+1} \in u_{n+1}]} m(u_n, u_{n+1}) p(y_n, y_{n+1}|x_n, x_{n+1})$$

Notice that the joint mass m is used instead of the transition mass m_1 for the sake of simplicity.

The posterior distributions are then given by:

$$p(x_n, u_n|y) \propto \alpha_n(x_n, u_n) \beta_n(x_n, u_n) \quad (13)$$

The posterior marginal distribution needed to achieve the data segmentation can then be derived as follows:

$$p(x_n|y_n) = \sum_{u_n} p(x_n, u_n|y_n) \quad (14)$$
$$p(u_n|y_n) = \sum_{x_n} p(x_n, u_n|y_n) \quad (15)$$

On the other hand, the posterior distributions $\psi_n(x_n, u_n, x_{n+1}, u_{n+1}) = p(x_n, u_n, x_{n+1}, u_{n+1}|y)$ needed for the parameters estimation procedure are computed as follows:

$$\psi_n(x_n, u_n, x_{n+1}, u_{n+1}) = \alpha_n(x_n, u_n) p(t_{n+1}|t_n) \beta_{n+1}(x_n, u_n) \quad (16)$$

2.5. Parameters estimation

In this paragraph we show how the set of parameters of the model can be estimated for the correlated Gaussian noise case from the only observation Y. For this purpose, we propose to use an adapted version of the EM algorithm. Let us consider an EPMC $T = (U, X, Y)$ as defined in (10). Such model is said to be "Gaussian" if, additionally, the distributions $p(y_{n-1}, y_n|x_{n-1}, x_n)$ are of Gaussian form. Considering K classes $\Omega = \{\omega_1, \dots, \omega_K\}$, we have to estimate, the following parameters: $2K^2$ means $\mu^1_{kk'}, \mu^2_{kk'}$, $2K^2$ standard deviations $\sigma^1_{kk'}, \sigma^2_{kk'}$, and K^2 correlation coefficients $\rho_{kk'}$, with $1 \leq k, k' \leq K$, defining the Gaussian noise densities $p(y_{n-1}, y_n|x_{n-1}, x_n)$ and the $(2^K - 1)^2$ elements $m_{ij} = m(u_n, u_{n+1})$ of the stationary evidential mass function m defined on $[P(\Omega)]^2$. The estimation procedure runs as follows:

- Consider an initial set of parameters $\Theta^0 = (\mu^1_{1..K}, \mu^2_{1..K}, \sigma^1_{1..K}, \sigma^2_{1..K}, \rho, m)^0$.
- For each iteration $q + 1$, compute Θ^{q+1} from Θ^q and in two steps:
  - E-step: Calculate $\alpha^q(x_n, u_n)$ and $\beta^q(x_n, u_n)$ and then derive $\psi^q(x_n, u_n, x_{n+1}, u_{n+1})$.
  - M- step: Compute Θ^{q+1} as follows:

$$(\mu^1_{kk'})^{q+1} = \frac{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n = \omega_k, u_n, x_{n+1} = \omega_{k'}, u_{n+1}) y_n}{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j)} \quad (17)$$

$$(\mu^2_{kk'})^{q+1} = \frac{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n = \omega_k, u_n, x_{n+1} = \omega_{k'}, u_{n+1}) y_{n+1}}{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j)} \quad (18)$$

$$[(\sigma^1_{kk'})^{q+1}]^2 = \frac{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n = \omega_k, u_n, x_{n+1} = \omega_{k'}, u_{n+1}) \left(y_n - (\mu^1_{kk'})^{q+1}\right)^2}{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j)} \quad (19)$$

$$[(\sigma^2_{kk'})^{q+1}]^2 = \frac{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n = \omega_k, u_n, x_{n+1} = \omega_{k'}, u_{n+1}) \left(y_{n+1} - (\mu^2_{kk'})^{q+1}\right)^2}{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j)} \quad (20)$$

$$[(\rho_{kk'})^{q+1}]^2 = \frac{\sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n = \omega_k, u_n, x_{n+1} = \omega_{k'}, u_{n+1}) \left(y_n - (\mu^1_{kk'})^{q+1}\right)\left(y_{n+1} - (\mu^2_{kk'})^{q+1}\right)}{(\sigma^1_{kk'})^{q+1} (\sigma^2_{kk'})^{q+1} \sum_{n=1}^{N-1} \sum_{u_n, u_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j)} \quad (21)$$

$$m_{ij}^{q+1} = \frac{1}{\#i\#j} \sum_{n=1}^{N-1} \sum_{x_n, x_{n+1}} \psi_n^q(x_n, u_n = i, x_{n+1}, u_{n+1} = j) \quad (22)$$

where #i denotes the cardinal of the set i.

3. EXPERIMENTS

In this section, we assess our model against the former Markov models. For this purpose, we present two series of experiments. In the first one, we deal with synthetic data sampled according to a nonstationary PMC. In the second one, we consider a real nonstationary class-image that we corrupt in some random manner such that a correlated noise is introduced. For all experiments one-dimensional data are transformed from and to images via Hilbert-Peano scan [7].

3.1. Unsupervised segmentation of nonstationary PMC

Let us consider the nonstationary PMC $Z = (X, Y)$ described in section 2.1 with $\Omega = \{\omega_1, \omega_2\}$ and $\mathrm{N} = 16384$. Let us assume that the nonstationary distributions $p_n(i,j)$ are governed by two different matrices Q and L:

$$Q = \begin{bmatrix} 0.48 & 0.02 \\ 0.02 & 0.48 \end{bmatrix}, L = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$$

For a given value of s we have:

- $p_n(i,j) = q_{ij}$ for $2rs + 1 \leq n \leq (2r+1)s, r \in \mathbb{N}$.
- $p_n(i,j) = l_{ij}$ for $(2r+1)s + 1 \leq n \leq (2r+2)s, r \in \mathbb{N}$.

The data are sampled considering different values for s and the following Gaussian noise parameters:

$$\mu^1 = [\mu^2]^T = \begin{bmatrix} -5 & -3 \\ 3 & 5 \end{bmatrix}, \ \sigma^1 = [\sigma^2]^T = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}.$$

For the correlation coefficient, we considered 3 different data sets with a different value for ρ per each set:

- Set A: $\rho = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$,
- Set B: $\rho = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$,
- Set C: $\rho = \begin{bmatrix} 0.9 & 0.9 \\ 0.9 & 0.9 \end{bmatrix}$.

MPM restoration is then achieved according to the nonstationary PMC with real parameters Θ (used as a reference) and the standard HMC, evidential HMC and the proposed EPMC using parameters estimated with EM. The obtained results are summarized in TABLE I. For data set C and $\mathrm{s} = 4096$, the sampled data and their corresponding restoration results are depicted in Fig. 1.

TABLE I
Segmentation Error Ratios of Synthetic Data (%)

Set	s	$\tau_{\Theta-PMC}$	τ_{EM-HMC}	$\tau_{EM-EHMC}$	$\tau_{EM-EPMC}$
A	16	9.1	17.2	17.1	12.8
	64	9.2	16.8	16.2	10.6
	256	9.2	16.6	15.9	9.7
	1024	9.1	16.7	15.7	9.2
	4096	9	16.5	15.5	9
B	16	7.6	20.5	20.4	8.9
	64	7.3	20.7	19.9	7.7
	256	7.3	21	20.5	7.7
	1024	7.6	20.4	19.7	7.6
	4096	7.8	21.2	20.4	7.8
C	16	3	24.2	24.2	3.7
	64	2.9	22.9	22.6	3.2
	256	2.6	23	22.8	2.8
	1024	2.4	22.4	21.8	2.6
	4096	2.7	23	22	2.7

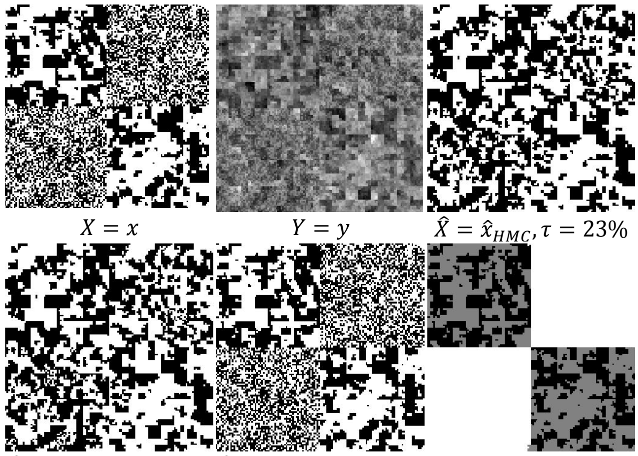

Fig.1. MPM restoration of synthetic data sampled according to a nonstationary PMC

Overall, the segmentation results show that the proposed model outperforms the classic ones. Furthermore, the mismatch between the standard models and the data is higher as the noise correlation is higher. On the other hand, the evidential PMC model, not only takes into account the data correlation, but rather benefits from the correlation as a feature to distinguish between the classes. Another important observation is that the EM-MPM restoration results according to the proposed EPMC are comparable to those obtained with the nonstationary PMC based on real parameters, especially for high values of s (1024 and 4096).

3.2. Unsupervised segmentation of nonstationary images corrupted with correlated noise

Let us consider the 128×128 "*Nazca bird*" nonstationary class-image (Fig. 2). We have then a realization of the hidden process X with $\Omega = \{\omega_1, \omega_2\}$ where ω_1 and ω_2 corresponds to black pixels and white ones respectively. Then, the image is corrupted with a correlated noise. The observed process is $Y_s = \sigma_{x_s} W_s + \mu_{x_s} + a \sum_{i=1}^{4} \left(\sigma_{x_{s_i}} W_s + \mu_{x_{s_i}} \right)$, where W is a white Gaussian noise with variance 1 and $s_1, .. s_4$ denote the four neighbors of the pixel s. The realizations of X and Y are converted to sequences via the Hilbert-Peano scan as in [7]. Hence, the data $Z = (X, Y)$ are of very complicated form and do not fit any of the Markov models. The interest of this experiment is to check whether the proposed model allow to satisfactorily restore the genuine image.

Experiments have been conducted considering different values for a, μ_{x_s} and σ_{x_s}. MPM restoration has been performed using EM procedure according to standard HMC, PMC, evidential HMC and evidential PMC. It turned out that the proposed model provides significantly better results.

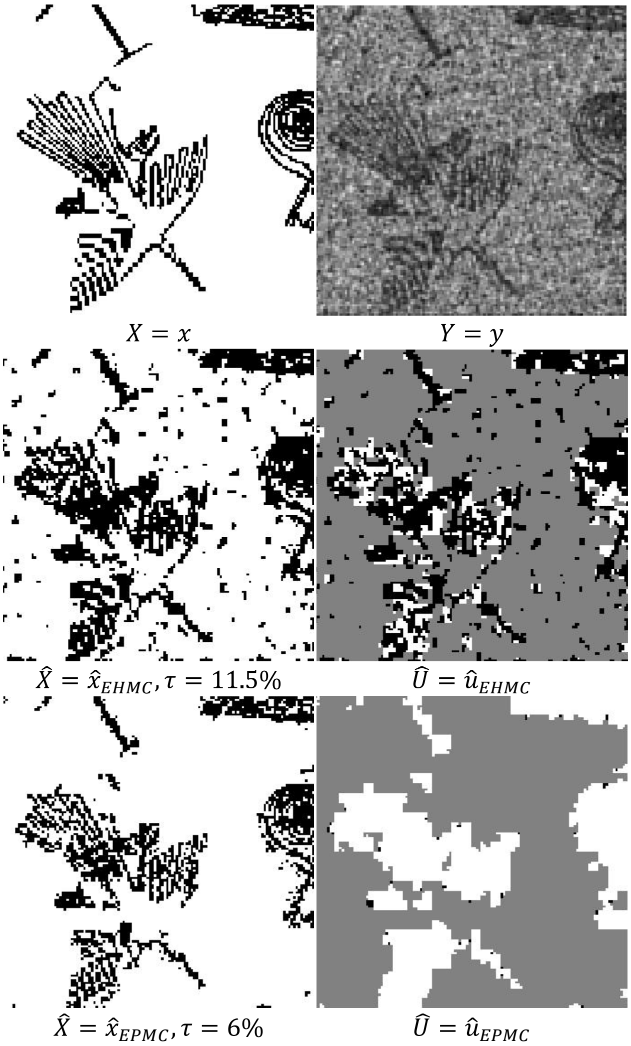

Fig.2. MPM restoration of a noisy nonstationary image

Figure 2 illustrates the restoration results for $a = 0.25$, $\mu_{\omega_1} = 0, \mu_{\omega_1} = 3$, $\sigma_{\omega_1} = 1$ and $\sigma_{\omega_2} = 2$. The gain in restoration error compared to the EHMC model is close to 50%. The restoration result of the auxiliary process U shows clearly how the evidential PMC is able to distinguish between image regions characterized with lot of details (e.g. wings and tail of the bird) and those where there is almost no exchange between the two classes. The evidential HMC, on the other hand, cannot make such a differentiation because it does not take advantage of correlation information. The restoration results based on HMC and PMC are poor given the nonstationary aspect of the data.

4. CONCLUSION

In this paper, we showed how the theory of evidence can be used within the pairwise Markov chains to model nonstationary data corrupted with correlated noise. We proposed an EM-like parameters estimation procedure and demonstrated through experiments that the novel model outperforms the former evidential hidden Markov chain in the sense that it allows to take correlated noise into account. As future work, we intend to extend the present formalism to pairwise Markov fields to better handle bi-dimensional data like images, and to multisensor nonstationary data [14].

5. REFERENCES

[1]. C. Raphael, "Automatic segmentation of acoustic musical signals using hidden Markov models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 (4), 1999.

[2]. M. Chen, A. Kundu, J. Zhou, "Off-line handwritten work recognition using a hidden Markov model type stochastic network", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (5), 1994.

[3]. O. Cappé *et al*, *Inference in hidden Markov models*, Springer, New York, 2005.

[4]. L.E. Baum, T. Petrie, G. Soules, N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains", *Annals of Mathematical Statistics* 41, 1970.

[5]. L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of IEEE*, vol. 77(2), pp. 257-286, 1989.

[6]. G. McLachlan and T. Krishnan, *EM Algorithm and Extensions*. Wiley, 1997.

[7]. S. Derrode and W. Pieczynski, "Signal and image segmentation using Pairwise Markov chains", *IEEE Trans. on Signal Processing*, Vol. 52, No. 9, pp. 2477-2489, (2004).

[8]. P. Lanchantin and W. Pieczynski, "Unsupervised restoration of hidden non stationary Markov chain using evidential priors", *IEEE Trans. on Signal Processing,* Vol. 53 (8), pp. 3091-3098, 2005.

[9]. H. Subaras, On evidential Markov chains, *Foundations of reasoning under uncertainty*, Studies in Fuzziness and Soft Computing, Vol. 249, Springer, pp. 247-264,2010.

[10]. Emmanuel Ramasso, "*Contribution of belief functions to HMM with an application to fault diagnosis*", IEEE International Workshop on Machine Learning and Signal Processing, Grenoble, France, Sept. 2-4, 2009.

[11]. P. Lanchantin, J. Lapuyade-Lahorgue and W. Pieczynski, "Unsupervised segmentation of randomly switching data hidden with non-Gaussian correlated noise", *Signal Processing*, Vol. 91(2), pp. 163-175, 2011.

[12]. G. Shafer. *A mathematical theory of evidence*, Princeton, 1976.

[13]. W. Pieczynski, "Multisensor triplet Markov chains and theory of evidence", *International Journal of Approximate Reasoning*, Vol. 45(1), pp. 1-16, 2007.

[14] M. Y. Boudaren, E. Monfrini, W. Pieczynski and A. Aïssani, "Dempster-Shafer fusion of multisensor signals in nonstationary Markovian context", EURASIP Journal on Advances in Signal Processing (To appear).