Unsupervised Segmentation of Switching Pairwise Markov Chains

Mohamed El Yazid Boudaren
Laboratoire Mathématiques Appliquées
Ecole Militaire Polytechnique
Algiers, Algeria
boudaren@gmail.com

Emmanuel Monfrini, Wojciech Pieczynski
CITI Dpt., Telecom SudParis
Institut Telecom
Evry, France

Abstract—**Pairwise Markov chains (*PMC*) have now shown their supremacy over hidden Markov chains (*HMC*) in unsupervised data segmentation since they allow one to deal with more complex processes structures. *HMC*s are particular cases of PMCs and these latter provide a gain in restoration accuracy within comparable computational complexity. On the other hand, the recent triplet Markov chains (*TMC*) have successfully substituted for classical *HMC*s to model data with some irregularities that these latter cannot handle. In fact, they provide an elegant formalism through the introduction of a third underlying process that permits to consider, for instance, regime switches or semi- Markovianity of the hidden process. The aim of this paper is to generalize the switching *HMC* to switching *PMC*. To validate the proposed model, we choose non stationary image segmentation as illustrative application field. Experimental results of synthetic and real images segmentation are provided.**

I. INTRODUCTION

The *HMC* is a pairwise stochastic process $Z = (X, Y) = (X_n, Y_n)_{n=1}^N$ in which, the unobservable process of interest X is a Markov chain that is to be recovered from an observable process Y that can be seen as its noisy version.

In such a model, when the classical noise assumptions hold, the distribution of Z is given by

$$p(z_n|z_{n-1}) = p(x_n|x_{n-1})p(y_n|x_n) \quad (1)$$

All along this paper, each X_n will take its values from a finite set of classes $\Omega = \{\omega_1, \dots, \omega_K\}$ and each Y_n will take its values from R.

HMC generalizes the so called mixture model since the hidden states $(X_n)_{n=1}^N$are related through a Markovian process rather than independent on each other, which provides a well-established theory to consider the spatial regularities of many data encountered in nature. Hence, hidden Markov models (chains, trees and fields) have been widely applied for a variety of problems that include image and signal processing [1]. *HMCs* have also been generalized to *PMCs* [2] which offer similar processing advantages and greater modeling potential.

However, when applied to non stationary data, both *HMCs* and *PMCs* fail to yield suitable segmentation results because of the lack of correspondence between the estimated stationary model and the data. To surmount this lack in the modeling, authors in [3] use an underlying discrete process $U = (U_n)_{n=1}^N$ within a *TMC* [4] to take into account the switches of the latent process *X*. In such a model, each one of the *M* stationary parts of the signal is modeled via a classical HMC and the switches between these different parts are assumed to be Markovian. The same thing happens when we consider a multi-textured image for instance; it is intuitive to consider a HMC per each texture provided that the number of textures *M* is known. Furthermore, the gain in using S-HMC is that the switches between different textures are governed by a Markov chain rather than independent on each other, which may serve as a regularization tool to prevent the "pepper and salt" aspect of the unobserved image which is being determined. However, the distributions $p(z|u)$ still however, follow a *HMC*. In this paper, we propose to consider $p(z|u)$ governed by a genuine *PMC* to make possible the modeling of more complex noise structures. Superiority of PMCs over HMCs has already been shown in previous works [2]. It hinges essentially on their ability to take noise correlation into account. Furthermore, in PMCs, one can easily model the fact that pixels near region boundaries inside the image may have different visual aspect than those located inside region.

Let us mention some previous works dealing with the nonstationarity of $p(x)$, which call on the introduction of the "time duration function" [5] followed by further generalizations to hidden semi-Markov chains [6]. Theory of evidence has also been successfully applied in Markovian context to model non stationary data [7]. The difference between these frameworks and the work proposed in [3], that we intend to extend, is that in our model, we have a precise information about nonstationarity: there is *M* different stationarities, and hence, *M* models (HMCs in [3] and PMCs here) are to be considered.

The goal of this work is to generalize the switching HMC proposed in [3] to switching PMC. To validate our novel model and its corresponding estimation and restoration processing, we choose unsupervised segmentation of images as illustrative application field.

The remainder of this paper is organized as follows: section 2 briefly reminds the formalism of HMC, *PMC* and *TMC*. Section 3 describes our new TMC and its computational developments. Section 4 is devoted to experiments achieved on synthetic and real images. Future improvements and conclusion are given in section 5.

II. Hidden, Pairwise and Triplet Markov Chains

In this section, we give an overview about three families of Markov chains with strictly growing degrees of generality: HMCs, PMCs and TMCs. Let us notice that HMCs, which are the most elementary ones, were extended in different directions. However, to our knowledge, in all these extensions, the hidden process remains Markovian and the resulting models are still hidden Markov models [8]. On the other hand, PMCs in which the hidden process is not Markovian exist and are therefore firmly more general. Similarly, TMCs form a family which is strictly more general than PMCs since TMCs that are not PMCs exist and were used to deal with several data irregularities that neither HMCs nor PMCs can handle [9].

A. Hidden Markov Chains

A HMC is a pairwise process $Z = (X, Y) = (X_n, Y_n)_{n=1}^{N}$ that considers X as a Markov chain which is to be recovered from its noisy version Y. Moreover, when the classical noise assumptions hold, the joint probability of Z is given by (1).

Therefore, *X* may be recovered from *Y* by means of some Bayesian decision rules such as maximal posterior mode (MPM) or maximum a posteriori (MAP) [1]. Throughout this paper, MPM will be adopted. Its corresponding formula is the following:

$$[\hat{x} = \hat{x}_{MPM}(y)] \Leftrightarrow [\hat{x}_n = argmax_{\ i}\, p(x_n = \omega_i|y)] \quad (2)$$

When the model parameters are known, the posterior distributions $p(x_n|y)$ required to perform MPM estimation are computed thanks to forward functions $\alpha_n(x_n) = p(y_1, \dots, y_n, x_n)$ and backward functions $\beta_n(x_n) = p(y_{n+1}, \dots, y_N|y_n, x_n)$ that can be computed in the following iterative way:

$$\alpha_1(x_1) = p(x_1)p(y_1|x_1);$$

$$\alpha_{n+1}(x_{n+1}) = \sum_{v_n} \alpha_n(x_n)p(x_{n+1}|x_n)p(y_{n+1}|x_{n+1}), \quad (3)$$

$$\beta_N(x_N) = 1;$$

$$\beta_n(x_n) = \sum_{v_{n+1}} \beta_{n+1}(x_{n+1})p(x_{n+1}|x_n)p(y_{n+1}|x_{n+1}), \quad (4)$$

The posterior margins can then be computed as follows:

$$p(x_n|y) \propto \alpha_n(x_n)\beta_n(x_n) \quad (5)$$

The indexing is then derived according to (2).

On the other hand, when the model parameters are unknown, several relatively quick algorithms can be used to find out these latter. We can cite for instance expectation-maximization algorithm (EM), its stochastic version (SEM) or iterated conditional estimation (ICE) [8].

B. Pairwise Markov Chains

In PMC, one directly assumes the Markovianity of $Z = (X, Y)$. Equation (1) becomes then

$$p(z_n|z_{n-1}) = p(x_n|x_{n-1}, y_{n-1})p(y_n|x_n, x_{n-1}, y_{n-1}) \quad (6)$$

The *HMC* can then be seen as a particular *PMC* where $p(x_n|x_{n-1}, y_{n-1}) = p(x_n|x_{n-1})$ and $p(y_n|x_n, x_{n-1}, y_{n-1}) = p(y_n|x_n)$. This shows the greater generality of *PMC* over *HMC* at the local level. On the other hand, at the global level, notice that the distribution of $p(y|x)$ is Markovian in *PMC* whereas, it is given by the independent noise formula $p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$ in *HMC*. As in *HMC*, the posterior marginal distributions $p(x_n|y)$ remain computable thanks to generalized forward and backward recursive functions [2].

C. Triplet Markov Chains

The model $Z = (X, Y)$ is said to be a *TMC* if there exists an underlying process $U = (U_1, \dots, U_N)$ with each U_n taking its values from a finite set $\Lambda = \{\lambda_1, \dots, \lambda_M\}$ such that $T = (U, X, Y)$ is a Markov chain. Let $V = (U, X)$. $T = (V, Y)$ is then a pairwise Markov chain (*PMC*). This makes the computation of the distributions $p(x_n|y)$, required to perform *MPM* restoration, affordable even when *Z* is not Markovian. This quickly shows the greater generality of *TMC* over *PMC*, which is itself more general than *HMC*. Hence, TMCs were used to model switches of the hidden process *X* [3], Dempster-Shäfer fusion in the Markovian context [4], etc.

III. Switching Pairwise Markov Chains

In this section, we develop our new *TMC* considered for unsupervised segmentation of switching data and its corresponding restoration processing. Throughout this paper, this model will be called switching *PMC* in contrast to the *TMC* proposed in [3] that will be called switching *HMC*.

A. Model Presentation

Let $X = (X_n)_{n=1}^{N}$ be an unobservable process that is to be estimated from its noisy version $Y = (Y_n)_{n=1}^{N}$. According to *TMC* formalism, one can set up an underlying process to model the regime switches of $Z = (X, Y)$. Let $U = (U_n)_{n=1}^{N}$ be such a process. $T = (U, X, Y)$ is then a *TMC* with transition probabilities given by

$$p(t_n|t_{n-1}) = p(u_n|u_{n-1}, z_{n-1})p(z_n|u_n, u_{n-1}, z_{n-1}) \quad (7)$$

The process *Z* is then called a switching *PMC* that generalizes the switching *HMC* governed by the simpler formula

$$p(t_n|t_{n-1}) = p(u_n|u_{n-1}, x_{n-1})p(x_n|u_n, u_{n-1}, x_{n-1})p(y_n|x_n) \quad (8)$$

In this paper, we consider *U* Markovian. We then deal with the particular *S-PMC* with transition probabilities given by

$$p(t_n|t_{n-1}) = p(u_n|u_{n-1})p(x_n|u_n, x_{n-1}, y_{n-1})p(y_n|x_n, x_{n-1}, y_{n-1}) \quad (9)$$

Dominance of *PMC* over *HMC* has been already shown in previous works [2]. This remains true when dealing with switching data because Z remains a *PMC*, only its parameters change along the signal.

Let us consider the distributions $p_m(i,j) = p(x_{n-1} = \omega_i, x_n = \omega_j | u_n = \lambda_m)$ and $f_{i,j}(y_{n-1}, y_n) = p(y_{n-1}, y_n | x_{n-1} = \omega_i, x_n = \omega_j)$, the distribution of $p(z|u)$ can then be expressed through

$$p(z_{n-1}, z_n|u_n = \lambda_m) = p_m(i,j)f_{i,j}(y_{n-1}, y_n) \quad (10)$$

The distribution of $p(z|u)$ can be equivalently determined in the classical way through the initial probabilities (11) and the transition probability (12).

$$p(z_1|u_2=\lambda_m)=\sum_{\omega_k\in\Omega} p_m(i,k)\, f_{i,k}(y_1) \quad (11)$$

where $f_{i,k}(y_1)=\int f_{i,k}(y_1,y_2)\,dy_2$.

$$p(z_n|z_{n-1},u_n=\lambda_m)=\frac{p(z_{n-1},z_n|u_n)}{p(z_{n-1}|u_n)}=\frac{p_m(i,j)f_{i,j}(y_{n-1},y_n)}{\sum_{\omega_k\in\Omega} p_m(i,k)f_{i,k}(y_{n-1})} \quad (12)$$

B. Switching Gaussian PMC Simulation

A switching *PMC* is said to be Gaussian if the noise densities $f_{i,j}(y_{n-1},y_n)$ are Gaussian. Such a model is specified through the M^2 distributions given by the transition matrix $A_{\lambda,\lambda'}=[a_{\lambda,\lambda'}]_{\lambda,\lambda'\in\Lambda}$ where $a_{\lambda,\lambda'}=p(u_n=\lambda'|u_{n-1}=\lambda)$, the MK^2 bi-dimensional densities $p_\lambda(i,j)$ given by the M transition matrices $(P_m)_{m=1}^{M}$ and the Gaussian noise densities given by the means $\mu_1^{i,j},\mu_2^{i,j}$, the standard deviations $\sigma_1^{i,j},\sigma_2^{i,j}$ and the correlation coefficient $\rho^{i,j}$. All the parameters being given, we can simulate a realization of a the corresponding switching *PMC* sequential as follows

1) The realization of U is simulated independently: u_1 is sampled by drawings from eigenvectors of $A_{\lambda,\lambda'}$ and the following realizations are obtained by drawings from the transition matrix $A_{\lambda,\lambda'}$.
2) The simulation of z_1 is obtained by drawings from the following distributions

$$p(x_1|u_1)=\sum_{\omega_k\in\Omega} p_m(k,i) \quad (13)$$

$$p(y_1|x_1,u_1)=\sum_{\omega\in\Omega} p_m(k,i)\, f_{k,i}(y_1) \quad (14)$$

Notice that $f_{k,i}(y_1)$ is a Gaussian density with mean $\mu_1^{k,i}$ and standard deviation $\sigma_1^{k,i}$. The simulation of y_1 is then obtained by a drawing from a Gaussian mixture.

3) The simulation of the next realizations of z are obtained by drawings from the following distributions

$$p(x_n|u_n,z_{n-1})=\frac{p_m(i,j)f_{i,j}(y_{n-1})}{\sum_{\omega_k\in\Omega} p_m(i,k)\, f_{i,k}(y_{n-1})} \quad (15)$$

$$p(y_n|x_n,z_{n-1})=\frac{f_{i,j}(y_{n-1},y_n)}{f_{i,j}(y_{n-1})} \quad (16)$$

It can be shown that $p(y_n|x_n,z_{n-1})$ is a Gaussian density with mean $\mu_2^{i,j}+\rho^{i,j}(\sigma_2^{i,j}/\sigma_1^{i,j})(y_{n-1}-\mu_1^{i,j})$ and standard deviation $\sigma_2^{i,j}\sqrt{1-(\rho^{i,j})^2}$. The simulation of y_n is then obtained by a drawing from a Gaussian mixture.

C. MPM Restoration

To achieve segmentation, one has to compute the posterior distributions $p(x_n|y)$. Since $T=(U,X,Y)$ is a *TMC*, $T=(V,Y)$ is then a *PMC* with $V=(U,X)$. Consequently, the marginal posterior distributions $p(v_n|y)$ are computable via the modified forward function $\alpha_n(v_n)=p(y_1,\ldots,y_n,v_n)$ and backward function $\beta_n(v_n)=p(y_{n+1},\ldots,y_N|y_n,v_n)$ that can be computed iteratively using (17) and (18).

$$\alpha_1(v_1)=p(y_1,v_1);$$

$$\alpha_{n+1}(v_{n+1})=\sum_{v_{n-1}}\alpha_{n-1}(v_{n-1})p(u_n|u_{n-1})p(z_n|z_{n-1},u_n) \quad (17)$$

$$\beta_N(v_N)=1;$$

$$\beta_n(v_n)=\sum_{v_{n+1}}\beta_{n+1}(v_{n+1})p(u_{n+1}|u_n)p(z_{n+1}|z_n,u_{n+1}) \quad (18)$$

The posterior distributions of V, X and U may then be derived:

$$p(v_n|y)\propto\alpha_n(v_n)\beta_n(v_n) \quad (19)$$

$$p(x_n|y)=\sum_{u_n}p(v_n|y) \quad (20)$$

$$p(u_n|y)=\sum_{x_n}p(v_n|y) \quad (21)$$

D. Parameters Estimation

In this section, we show how to estimate the switching Gaussian *PMC* parameters when these latter are unknown. Let $A_{\lambda,\lambda'}$, $(P_m)_{m=1}^{M}$, $\mu_1^{i,j}$, $\mu_2^{i,j}$, $\sigma_1^{i,j}$, $\sigma_2^{i,j}$ and $\rho^{i,j}$ be such parameters. In this paper, we propose to use the Expectation-Maximization algorithm (*EM*). Accordingly, the parameters are estimated iteratively in the following manner

1) Choose an initial set of parameters $\Theta^0=\left(A_{\lambda,\lambda'},(P_m)_{m=1}^{M},\mu_1^{i,j},\mu_2^{i,j},\sigma_1^{i,j},\sigma_2^{i,j},\rho^{i,j}\right)^0$.
2) For each iteration, we compute $\psi_n(v_n,v_{n+1})=p(v_n,v_{n+1}|y)$ and $\xi_n(v_n)=p(v_n|y)$ according to Θ^q thanks to:

$$\psi_n(v_n,v_{n+1})\propto\alpha_n(v_n)a_{u_n,u_{n+1}}p(z_{n+1}|z_n,u_{n+1})\beta_{n+1}(v_{n+1}) \quad (22)$$

$$\xi_n(v_n)=\sum_{v_{n+1}}\psi_n(v_n,v_{n+1}) \quad (23)$$

Then we derive Θ^{q+1} as follows

$$\left(\mu_1^{i,j}\right)^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})y_n 1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}}{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}} \quad (24)$$

$$\left(\mu_2^{i,j}\right)^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})y_{n+1} 1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}}{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}} \quad (25)$$

$$\left(\sigma_1^{i,j}\right)^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})\left(y_n-\mu_1^{i,j}\right)^2 1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}}{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}} \quad (26)$$

$$\left(\sigma_2^{i,j}\right)^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})\left(y_{n+1}-\mu_2^{i,j}\right)^2 1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}}{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}} \quad (27)$$

$$(\rho^{i,j})^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})\left(y_n-\mu_1^{i,j}\right)\left(y_{n+1}-\mu_2^{i,j}\right)1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}}{\sum_{n=1}^{N-1}\sum_{v_n,v_{n+1}}\psi_n(v_n,v_{n+1})1_{[(x_n,x_{n+1})=(\omega_i,\omega_j)]}} \quad (29)$$

$$a_{\lambda,\lambda'}^{q+1}=\frac{\sum_{n=1}^{N-1}\sum_{v_n}\psi_n(v_n,v_{n+1})1_{[(u_n,u_{n+1})=(\lambda,\lambda')]}}{\sum_{n=1}^{N-1}\sum_{v_n}\xi_n(v_n)1_{[u_n=\lambda]}} \quad (30)$$

$$P^{q+1}_{m,\omega,\omega'} = \frac{\sum_{n=1}^{N-1}\sum_{v_n}\psi_n(v_n,v_{n+1})1_{[(u_{n+1},x_n,x_{n+1})=(\lambda_m,\omega,\omega')]}}{\sum_{n=1}^{N-1}\sum_{v_n}\psi_n(v_n,v_{n+1})1_{[(u_{n+1},x_n)=(\lambda_m,\omega)]}} \quad (31)$$

3) We repeat the previous step until an end criterion is reached.

IV. EXPERIMENTS

In this section, we apply our proposed model to segment two types of data: synthetic ones that we generate according to the switching Gaussian *PMC* formalism and real ones that we noise in different ways. More explicitly, we consider image segmentation as illustrative application field. For this purpose, one-dimensional signals are converted from and to images using the Hibert Peano scan [2]. Experiments are achieved on 256×256 images. The results of *MPM* segmentation using the *EM* estimation algorithm are provided.

A. Experimental Images

Let us consider the case where $\Omega = \{\omega_1, \omega_2\}$. Experiments are achieved on the following images sets:

- Images set 1: We generate 256×256 synthetic images using the following parameters:

$$A_{\lambda,\lambda'} = \frac{1}{10^3}\begin{bmatrix} 998 & 1 & 1 \\ 1 & 998 & 1 \\ 1 & 1 & 998 \end{bmatrix}$$

$$P_1 = \begin{bmatrix} 0.49 & 0.01 \\ 0.01 & 0.49 \end{bmatrix}, P_2 = \begin{bmatrix} 0.39 & 0.11 \\ 0.11 & 0.39 \end{bmatrix},$$
$$P_3 = \begin{bmatrix} 0.29 & 0.21 \\ 0.21 & 0.29 \end{bmatrix}.$$

$$\mu_1 = \begin{bmatrix} -1.25 & -0.75 \\ 0.75 & 1.25 \end{bmatrix}, \mu_2 = \begin{bmatrix} -1.25 & 0.75 \\ -0.75 & 1.25 \end{bmatrix}.$$

$$\sigma_1 = \sigma_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Three different values of ρ are considered:

Experiment 1.1: $\rho = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$

Experiment 1.2: $\rho = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$

Experiment 1.3: $\rho = \begin{bmatrix} 0.9 & 0.9 \\ 0.9 & 0.9 \end{bmatrix}$

- Images set 2: We generate 256×256 synthetic images using the same parameters except the noise ones that are replaced as follows:

Experiment 2.1:

$$\mu_1 = \begin{bmatrix} -5 & -3 \\ 3 & 5 \end{bmatrix}, \mu_2 = \begin{bmatrix} -5 & 3 \\ -3 & 5 \end{bmatrix}$$

$$\sigma_1 = \begin{bmatrix} 4 & 4 \\ 8 & 8 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 4 & 8 \\ 4 & 8 \end{bmatrix}, \rho = \begin{bmatrix} 0.99 & 0.99 \\ 0.99 & 0.99 \end{bmatrix}$$

Experiment 2.2:

$$\mu_1 = \begin{bmatrix} -5 & -3 \\ 3 & 5 \end{bmatrix}, \mu_2 = \begin{bmatrix} -5 & 3 \\ -3 & 5 \end{bmatrix} \sigma_1 = \begin{bmatrix} 14 & 7 \\ 9 & 20 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 14 & 9 \\ 7 & 20 \end{bmatrix}, \rho = \begin{bmatrix} 0.99 & 0.99 \\ 0.99 & 0.99 \end{bmatrix}$$

- Images set 3: We consider the zebra non stationary image (Fig. 2) that we noise in three different ways:

Experiment 3.1: the image is noised according to a classical Gaussian *HMC* with $\mu = [1\ 3]$ and $\sigma = [1\ 1]$.

Experiment 3.2: a correlated noise is introduced using the following parameters:

$$\mu_1 = \begin{bmatrix} -2 & -1 \\ 1 & 2 \end{bmatrix}, \mu_2 = \begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix}, \sigma_1 = \sigma_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$
$$\rho = \begin{bmatrix} 0.99 & 0.99 \\ 0.99 & 0.99 \end{bmatrix}.$$

Experiment 3.3: the same as Experiment 3.2 with

$$\mu_1 = \begin{bmatrix} -2 & -1 \\ 1 & 2 \end{bmatrix}, \mu_2 = \begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix}, \sigma_1 = \sigma_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$
$$\rho = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

The aim of the first experiment is to check whether the new model permits to segment the image when this latter is noised according to a classical switching hidden Markov chain. In the next experiments, a correlated noise is introduced to the image in order to assess our new model (S-PMC) against the classical S-HMC.

For image sets 1 and 2, restoration is performed using real and estimated parameters. For the last image set, it is achieved according to *S-HMC* and *S-PMC* formalisms. For parameters estimation, *EM* algorithm is initialized via K-means algorithm.

B. Experimental Results

In this section, we show the results of the experiments described above. Segmentation results are given in TAB. I (images sets 1 and 2) and TAB. II (images set 3). Some of them are illustrated in Fig. 1, Fig. 2 and Fig.3.

TABLE I. MISCLASSIFICATION RATIOS OF SYNTHETIC IMAGES RESTORATION

Experiment	Images set 1			Images set 2	
	Exp. 1	*Exp. 2*	*Exp. 3*	*Exp. 1*	*Exp. 2*
$\tau_{\hat{x}_\theta}$ (%)	9.72	9.07	2.69	14.38	6.67
$\tau_{\hat{x}_{\hat{\theta}}}$ (%)	9.73	9.11	2.73	14.42	6.70
$\tau_{\hat{u}_\theta}$ (%)	5.19	6.05	1.77	7.30	5.84
$\tau_{\hat{u}_{\hat{\theta}}}$ (%)	5.45	8.40	1.72	8.54	6.05

The experimental results of experimental sets 1 and 2 show that *EM* algorithm provides comparable results with those given by real parameters (TAB. I). Notice that, other algorithms (Stochastic EM and ICE [2]) that provide comparable results can be used for the parameters estimation.

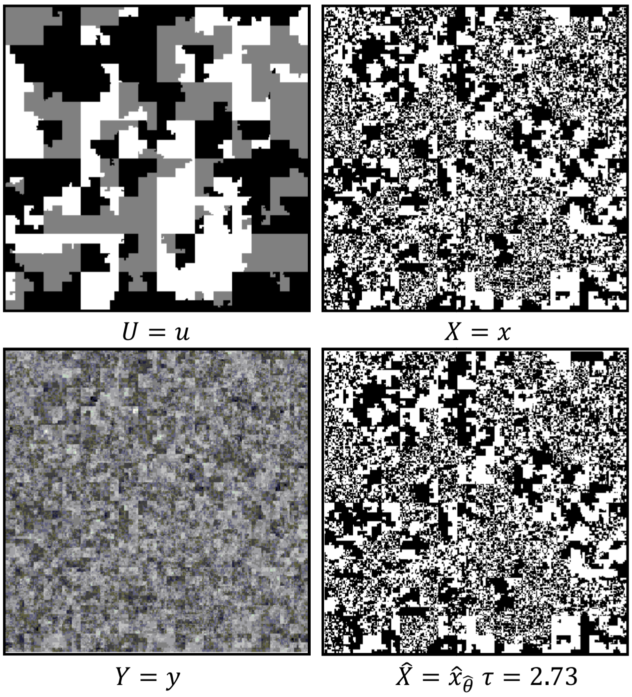

Figure 1. MPM restoration of a synthetic non stationary image (exp. 1.3)

On the other hand, the segmentation results of image set 3 demonstrate the dominance of S-PMC over S-HMC when applied to images presenting different homogeneous classes (stationarities). In fact, as shown in TAB. II, S- PMC allows to better recover the hidden process realization $X = x$ than S-HMC.

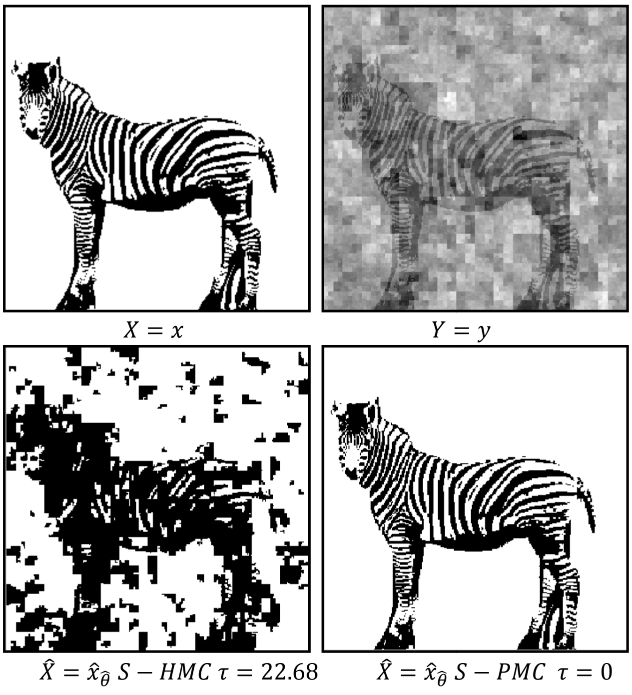

Figure 2. MPM restoration of Zebra image (exp. 3.2)

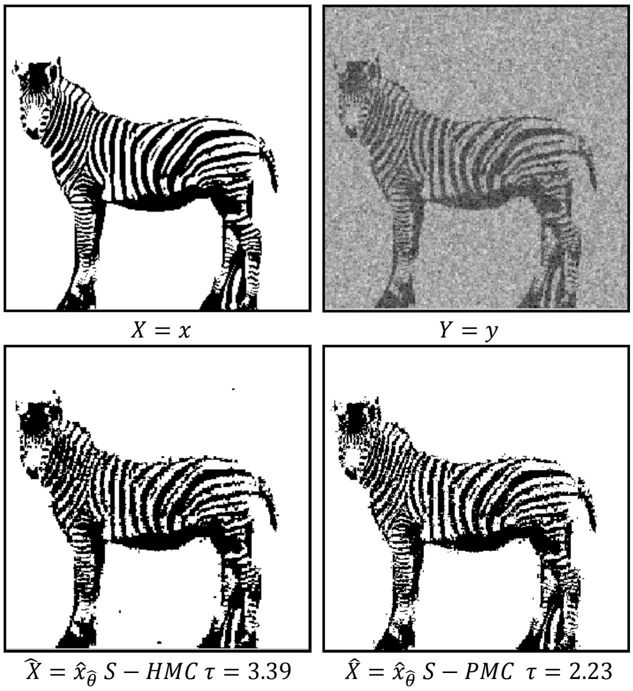

Figure 3. MPM restoration of Zebra image (exp. 3.3)

TABLE II. MISCLASSIFICATION RATIOS OF ZEBRA IMAGE RESTORATION

Experiment	**Images set 3**		
	Exp. 1	*Exp. 2*	*Exp. 3*
$\tau_{\hat{x}_{\hat{\theta}}}\ S - HMC$ (%)	3.42	22.68	3.39
$\tau_{\hat{x}_{\hat{\theta}}}\ S - PMC$ (%)	3.43	0	2.23

V. CONCLUSION

In this paper, we showed how switching hidden Markov chains can be generalized to switching pairwise Markov chains to model non stationary data. To validate our model and its corresponding estimation and restoration processing, we have chosen unsupervised images segmentation as application field. The obtained results demonstrated the superiority of *S-PMC* over *S-HMC*.

As perspective for future work, we propose to extend the present model to consider the case where the noise densities depend on the underlying process *U* realization, which may be of interest to model multi-textured images.

REFERENCES

[1] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition”, *Proc. of IEEE*, vol. 77, no. 2, pp. 257-286, 1989.

[2] 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.

[3] P. Lanchantin, J. Lapuyade-Lahorgue, and W. Pieczynski, “Unsupervised segmentation of randomly switching data hidden with non-Gaussian correlated noise”, *Signal Processing*, Vol. 91, No. 2, pp. 163-175, Feb. 2011.

[4] W. Pieczynski, "Multisensor triplet Markov chains and theory of evidence", *International Journal of Approximate Reasoning*, Vol. 45, No. 1, pp. 1-16, 2007.

[5] B. Sin and J. H. Kim, "Nonstationary hidden Markov model," Signal Processing, vol. 46, pp. 31–46, 1995.

[6] S.-Z. You and H. Kobayashi, "A hidden semi-Markov model with missing data and multiple observation sequences for mobility tracking", Signal Processing, vol. 83, pp. 235–250, 2003.

[7] P. Lanchantin and W. Pieczynski, "Unsupervised restoration of hidden nonstationary Markov chains using evidential priors", *IEEE Trans. on Signal Processing*, Vol. 53, No. 8, pp. 3091-3098, 2005.

[8] W. Pieczynski, "Triplet Markov chains and image segmentation". In Inverse problems in Vision and 3D Tomography, Chapter 4, A. Mohammed-Djafari ed., Wiley.

[9] M. E. Y. Boudaren, W. Pieczynski, E. Monfrini, "Unsupervised segmentation of non stationary data hidden with non stationary noise", in *Proc. of IEEE WoSSPA*, Tipaza 2011.