Unsupervised Multispectral Image Classification By Fuzzy Hidden Markov Chains Model For SPOTHRV Images

Faiza DAKKA dakka_f@yahoo.fr
Faculty of Sciences, Department of Physics,
Mohamed V-Agdal University
Rabat B.P. 1014, Morocco

Ahmed HAMMOUCH
GIT-LGE Laboratory, ENSET,
Mohamed V-Souissi University
Rabat, B.P. 6207, Morocco

Driss ABOUTAJDINE
Faculty of Sciences, Department of Physics,
Mohamed V-Agdal University
Rabat B.P. 1014, Morocco

Abstract

This paper deals with unsupervised classification of multi-spectral images, we propose to use a new vectorial fuzzy version of Hidden Markov Chains (HMC).

The main characteristic of the proposed model is to allow the coexistence of crisp pixels (obtained with the uncertainty measure of the model) and fuzzy pixels (obtained with the fuzzy measure of the model) in the same image. Crisp and fuzzy multi-dimensional densities can then be estimated in the classification process, according to the assumption considered to model the statistical links between the layers of the multi-band image. The efficiency of the proposed method is illustrated with a Synthetic and real SPOTHRV images in the region of Rabat.

The comparisons of two methods: fuzzy HMC and HMC are also provided. The classification results show the interest of the fuzzy HMC method.

Keywords: Bayesian Iimage Classification, Markov Chains, Fuzzy Hidden Markov, Unsupervised Classification.

1 INTRODUCTION

Hidden Markov models are used in many forms and with different spatial structures like strings, fields or trees [3]. The popularity of these models is illustrated by the multitude and diversity of applications that have been proposed, in signal processing in speech recognition and also in image processing. In particular, the model of hidden Markov chains (CMC) has been successfully used for image classification [3].

In this case, 2D images are first transformed into a 1D vector path through Hilbert-Peano [5] to fit the one-dimensional structure of a chain Fig. 1 (a).The interest for this model comes from the fact that when the hidden process $X = (X, ..., X_N)$ can be represented by a finite Markov chain and when the structure of the noise is not too complex, then X can be reconstructed from the only observed process $Y = (Y_1, ..., Y_N)$, Y_n□ R,, using different criteria for classification like MAP "Maximum A Posteriori" or MPM Maximal Posterior Mode. It is sometimes interesting to take into account not only the uncertainty in the observations (often due to noise), but also their vagueness (fuzzy). This imprecision may come from a motion sensor at the time of the acquisition of a photo, or the effect of surface / volume part found in medical imaging and satellite. The latter effect reflects the fact that a pixel of the image is the result of integration over a surface not necessarily homogeneous, where several components are mixed in different proportions and unknown. In terms of classification, these pixels have an ambiguous and sometimes it is better not to get a "harsh response" from classifier (eg, Class 0 or Class 1) but a "vague response" measured by an interval value]0,1[. This response reflects the ambiguity of the pixels can be observed and interesting information. Indeed, the level of degrees or can be interpreted as the mixing ratio of class 0 versus class 1 in the pixel, or as the degree of belonging to class 0.

This paper is organized as follows: The CMC model is briefly presented in Section 2. In Section 3 the CMC model for fuzzy image classification is detailed, and also The CMC Parameter estimation, carried out with an extension of the algorithm Iterative Conditional Estimation (ICE) [1, 2]. Sections 4 illustrate the comparative results of Classification Unsupervised. Finally, Section 5 reviews the theoretical and experimental work and offers some perspectives.

2 UNSUPERVISED CLASSIFICATION USING HMC

This section is intended to give some recalls about the HMC model [6,7] and its use for unsupervised image classification [8]. The HMC model can be adapted to a 2D analysis through a Hilbert-Peano scan of the image, see Fig. 1. Hence all estimation and classification processings are applied on the 1D sequence, and the segmented 2D image is reconstructed by using a reverse Hilbert-Peano scan from the 1D classified sequences.

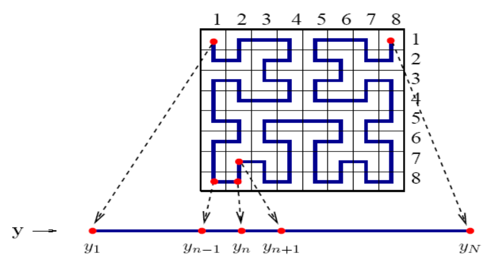

FIGURE 1: Hilbert-Peano scan construction for an 8*8 image. This scan is used to transform a 2D image into a 1D signal (y), and conversely.

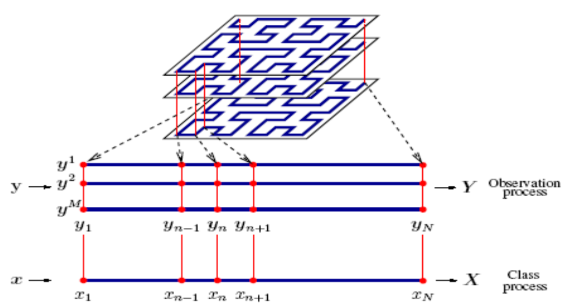

FIGURE 2: Hilbert-Peano scan for an 8*8 multi-component image (M = number of layers in the image, N = number of pixels in each layer).

The Hilbert-Peano scan presents the ability to take into account the neighborhood of the pixel of interest [6]. Fig. 2 shows the scan for a 8 * 8 multi-component image with M layers.

A. The HMC model

An image $y = \{y_1, \ldots, y_N\}$, N being the total number of pixels, is considered as a realization of the 1D observed process $Y = \{Y_1, \ldots, Y_N\}$, each Y_n is a real-valued random variable. The segmented image $x = \{x_1, \ldots, x_N\}$ is considered as a realization of a hidden process $X = \{X_1, \ldots, X_N\}$, each $X_n \in \Omega = \{1, \ldots, K\}$ is a discrete random variable. In classical HMC modeling, X is assumed to be a Markov chain, i.e. $p(x_{n+1} \mid x_n, \ldots, x_1) = p(x_{n+1} \mid x_n)$.

The distribution of X is consequently determined by the distribution of X_1, denoted by $\pi_k = p(X_1 = k)$, and the set of transition matrices $(A_n)1 \le n \le N$ whose entries are $a^n_{ij} = p(X_{n+1} = j \mid X_n = i)$.
We further assume that X is a stationary Markov chain, i.e. entry $a^n_{ij} = a_{ij}$ does not depend on n. With the following additional properties:
(i) Y_n are independent conditionally to X, i.e. $p(y \mid x) = \pi^N_{n=1} p(y_n \mid x)$, and
(ii) $p(y_n \mid x) = p(y_n \mid x_n)$, the distribution of the pairwise process (X,Y) can be written as

$$p(x, y) = \pi_{x1} f_{x1}(y1) \prod_{n=2}^{N} a_{x_{n-1}}, x_n f x_n(y$$

with $fx_n(y_n) = p(y_n \mid x_n)$ the data-driven densities, which are assumed to be Gaussian in this work [8].

B. Classification:

The estimation of X from Y can be done by applying the MPM Bayesian criterion:

$$\forall n \in [1, \ldots N], x_n^{MPM}(y) = \arg\max \xi_n(k)$$

(1)
with $\xi_n(k) = p(X_n = k \mid y)$ the marginal a posteriori probabilities.
The HMC model allows explicit computation of the MPM solution using the well-known Baum's "forward" $\alpha_n(k)$ and "backward" $\beta_n(k)$ probabilities [10], modified by Devijver [9] for computational reasons:

$\alpha_n(k) \approx P(X_n = k, Y_1 = y_1, \ldots, Y_n = y_n)$ (2)

$$\beta_n(k) \approx \frac{p(Y_{n+1} = y_{n+1}, \ldots, Y_N = y_N \mid X_n = k)}{p(Y_{n+1} = y_{n+1}, \ldots, Y_N = y_N \mid Y_1 = y_1, \ldots, Y_n = y}$$

(3)
In the following, we use the numerically stable forward–backward recursions resulting from these approximations:
Forward initialization:
for $1 \le i \le K$.

$$\alpha_n(i) = \frac{\pi_i f_i(y_1)}{\sum_{1 \le j \le k} \pi_j f_j(y_1)}$$

(4)

Forward induction:
for n =2,…,N and $1 \le i \le K$.

$$\alpha_n(i) = \frac{f_i(Y_n) \sum_{1 \le j \le k} \alpha_{n-1}(j) a_{ij}}{\sum_{1 \le l \le k} f_l(Y_n) \sum_{1 \le j \le k} \alpha_{n-1}(j)}$$

(5)
for $1 \le i \le K$

$$\beta_N(i) = 1$$

Backward induction:
for n=N-1,…,1 et $1 \le i \le K$

$$\beta_n(i) = \frac{\sum_{1 \le j \le k} a_{ij} f_j(Y_{n+1}) \beta_{n+1}(j)}{\sum_{1 \le l \le k} f_l(Y_{n+1}) \sum_{1 \le j \le k} \alpha_{n-1}(j)}$$

(6)
It can be shown that marginal a posteriori probabilities involved in MPM classification can be written

$$\xi_n(i) = \alpha_n(i) \beta_n(i),$$

(7)
and joint a posteriori probabilities $\Psi_n(i,j) = P(X_n = i, X_{n+1} = j \mid Y = y)$ as:

$$\Psi_n(i,j) = \frac{\alpha_n(i) a_{ij} f_j(Y_{n+1}) \beta_{n+1}(j)}{\sum_{1 \le m \le k} f_m(Y_{n+1}) \sum_{1 \le l \le k} \alpha_n(l) a}$$

(8)

C. Estimation:

Before classification, all parameters involved in the CMC model have to be estimated.

$$\theta = \{\pi_k, a_{kl}, f_k\} \quad k,l \square \Omega$$

(9)

One well-known solution is to use the EM iterative procedure [11] aiming at optimizing the log-likelihood of data, according to the steps described in Algorithm 1, with following update equations

$$\forall \mathrm{k} \in \Omega, \Pi_k^{[q]} = \frac{1}{\mathrm{N}} \sum_{n=1}^{N} \xi_n^q(k) \tag{10}$$

$$\forall \mathrm{k}, \mathrm{l} \in \Omega, \hat{a}_{kl}^{[q]} = \frac{\sum_{n=1}^{N-1} \Psi_n^q(k,l)}{\mathrm{N} \hat{\Pi}_k^q} \tag{11}$$

$$\forall \mathrm{k} \in \Omega, \hat{\mu}_k^{[q]} = \frac{\sum_{n=1}^{N} \xi_n^q(k) y_n}{\mathrm{N} \hat{\Pi}_k^q} \tag{12}$$

$$\forall \mathrm{k} \in \Omega, \hat{\sigma}^{2}{}_k^{[q]} = \frac{\sum_{n=1}^{N} \xi_n^q(k)(y_n - \hat{\mu}_k^{[q]})^2}{\mathrm{N} \hat{\Pi}_k^q} \tag{13}$$

Algorithm 1 : HMC parameters estimation using EM.
q ← 0: Initialize parameters $\theta^{[0]}$
repeat
q ←q + 1
- Compute “Forward” $\alpha_n^{[q]}$ and “Backward” $\beta_n^{[q]}$ probabilities using equations (5) and (6).
- Compute a posteriori probabilities $\xi_n^{[q]}(k)$ using eq. (6) and $\psi_n^{[q]}(k,l)$ using eq. (7) and (8).
- Estimate CMC parameters θ[q] using equations (10) and (11) for Markov parameters and equations (12) and (13) for Gaussian data-driven parameters.
until |θ[q] - θ[q-1] | <Threshold.
Iterative estimation of parameters is stopped when parameters do not vary much.

3 UNSUPERVISED CLASSIFICATION USING FUSSY HMC

A. Fuzzy Markov chains model

We consider a multi-component image of M layers. According to the Hilbert-Peano scan, we get N series of M data, denoted by y = { y_1 ,… y_N }, where y_n={ $y_n^{[1]}$,…, $y_n^{[M]}$ }^{t}, 1≤n ≤ N.

In classical HMC approach, the aim is to classify each yn□RM into a set of K classes, the state space Ω ={w_1,…,w_K}, in order to obtain the segmented chain x ={x_1,…, x_N} (Fig. 2). The segmented image is then reconstructed from x using an inverse Hilbert-Peano scan.
For the sake of simplicity, we confine our study to the K = 2 case, i.e. Ω= {0,1}.
In fuzzy HMC context, the range of x_n is now the interval Ω= [0, 1]. In the following, εn will denote a realization of random variable X_n and we will adopt the notation:
ε_n = 0 if the pixel is from class 0,
ε_n □]0, 1[if the pixel is a fuzzy one,
ε_n = 1 if the pixel is from class 1.

B. Probabilities in fuzzy Markov chains context

As stated previously, each x_n takes its value in two types of sets: a hard one {0, 1}, and a fuzzy one defined over the range]0, 1[.
Let δ_0 and δ_1 be Dirac weights on 0 and 1, and □$_1$ the Lebesgue measure on]0, 1[.
By taking v = δ_0 + δ_1 + □$_1$ as a measure on Ω, the distribution of X_n can be defined by a density h on Ω with respect to v.
If we assume that X is homogeneous and the distribution of each X_n is uniform on the fuzzy class, P (X_n = ε_n) =h(ε_n) = $\pi_{\varepsilon n}$ can be written:
h(ε_n = 0) = π_0;

$h(\varepsilon_n = 1) = \pi_1$;
$h(\varepsilon_n) = \pi_{]0,1[}$; ε_n □]0, 1[;
with $\pi_0 + \pi_1 + \pi_{]0,1[} = 1$.
We can now detail the new expression for the transition probabilities of the Markov chain $t_{\varepsilon_{n-1},\varepsilon_n}$:

$$t_{\varepsilon_{n-1},\varepsilon_n} = P(X_n = \varepsilon_n / X_{n-1} = \varepsilon_{n-1}) =$$
$$P(X_n = 0 / X_{n-1} = \varepsilon_{n-1})\delta_0(\varepsilon_n)$$
$$+ P(X_n = \varepsilon_n / X_{n-1} = \varepsilon_{n-1})\iota_{]0,1[}(\varepsilon_n)$$
$$+ P(X_n = 1 / X_{n-1} = \varepsilon_{n-1})\delta_1(\varepsilon_n)$$
$$\forall \varepsilon_{n-1}, \varepsilon_n \in \Omega \text{ and } \forall n \in \{2,...,N\}$$

With $t_{\varepsilon_{n-1},\varepsilon_n} \geq 0$ and $\int_\Omega t_{\varepsilon_{n-1},\varepsilon_n} d_{\varepsilon_{n-1}} = 1$

$$P(X = x) = \Pi_{\varepsilon_1} \prod_{n=2}^{N} t_{\varepsilon_{n-1},\varepsilon_n}$$

C. Multi-component fuzzy HMC implementation.
ε_n is a realization of a random variable X_n, and each y_n is a realization of a random vector $y_n = \{ y_n^{[1]},\dots, y_n^{[M]} \}^t$, Thus the problem is to estimate the unobserved realization x of a random process X from the observed realization y of a random process y = { y_1 ,… y_N };
Similarly to classical HMC, multi-component fuzzy HMC based image classification methods consider the two following assumptions:
H1: the random variables $Y_1,\dots,Y_N$ are independent conditionally on X;
H2: the distribution of each Y_n conditionally on X is equal to its distribution conditionally on X_n.
It is important to note that the random variables ($y_n^{[m]}$)1≤m≤M are not assumed to be mutually independent conditionally on X_n.

Assuming that distributions of $(X_n,Y_n,X_{n+1},Y_{n+1})$ are independent of n, each state ε_n of the state space (i.e. hard classes {0,1}, as well as the fuzzy class]0,1[) is associated to a distribution characterizing the M-dimensional observations y_n:

$$f\varepsilon_n(y_n) = P(Y_n = y_n \mid X_n = \varepsilon_n), \quad (14)$$

Given an observed sequence y = {y_1,…, y_N}, the joint state-observation probability is given by:

$$P(X = x, Y = y) = \Pi_{\varepsilon_1} f_{\varepsilon_1}(y_1) \prod_{n=2}^{N} t_{\varepsilon_{n-1},\varepsilon_n} f_{\varepsilon_n}(y_n) \quad (15)$$

In unsupervised classification, the distribution P (X = x,Y = y) is unknown and must first be estimated in order to apply a Bayesian classification technique (MAP or MPM). Therefore the following sets of parameters need to be estimated:
1) The set Γ characterizing the fuzzy Markov chain parameters, i.e. the initial probability vector π = {π_ε} ε□Ω and the transition probabilities $t_{\varepsilon_{n-1},\varepsilon_n} \forall \varepsilon_{n\_1}, \varepsilon_n \in \Omega$
2) The set Δ regrouping the parameters of the M- dimensional distributions presented in (14), i.e. the distributions associated with the hard classes and the fuzzy one.

D. Fuzzy Markov Chain parameters estimation
For the estimation of the parameters in ¡, we propose to use an adaptation of the general ICE algorithm [12] which is an alternative to the well-known Estimation-Maximization (EM) algorithm. In fact, ICE does not refer to the likelihood, but it is based on the conditional expectation of some estimators from the complete data (x, y). It is an iterative method which produces a sequence of estimations θq of parameter θ as follows:
1) initialization θ_0, obtained with an initial classification algorithm (k-means algorithm).

2) computation of $\theta^{q+1} = [\hat{\theta}(x,y)/Y = y]$ where $\hat{\theta}(x,y)$ is an estimator of θ.

3) Stop the algorithm when θ_{Q-1} □ θ_Q.

This section is not intended to give a complete description of the ICE algorithm in the HMC context, interested readers may consult [5]. Similarly to the classical case, parameters in Γcan be calculated analytically by using the Baum-Welch algorithm [13]:

for the hard classes, the classical normalized Baum-Welch probabilities [14] can be used directly.

for the fuzzy class, the forward and backward probabilities are defined by:

$$\alpha_{n+1}(\xi) \propto f_{\xi}(y_{n+1}) \int_{[0,1]} \alpha_n(\varsigma) t_{\varsigma,\xi} d\varsigma \quad (16)$$

$$\beta_n(\xi) \propto \int_{[0,1]} \beta_{n+1}(\varsigma) t_{\varsigma,\xi} f_{\xi}(y_{n+1}) d\varsigma \quad (17)$$

The integrals above can not be calculated analytically and numerical integration must be performed. Hence, the continuous interval]0,1[is partitioned into a given number F of subintervals. We thus reduce the domain of fuzzy membership degree]0,1[to F attributed values, corresponding to the medium value of the sub-interval of interest (see Fig. 3), thus resulting in a quantization of the fuzzy measure]0; 1[.

We obtained F fuzzy classes "discrete, whose value corresponds to the fuzzy median subinterval considered.

For example, F = 2 implies ε □ {0.25, 0.75 }, F = 3 implies x □{0. 165,0:5, 0.825}. More F is large, over the parameter estimation is accurate.

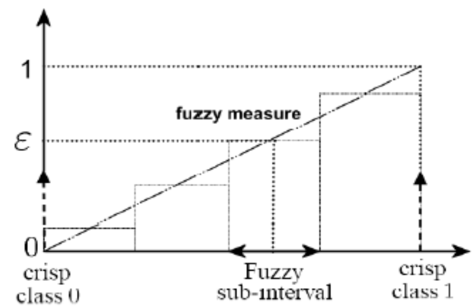

FIGURE 3: Partition of the continuous interval]0,1[into F = 4 sub-intervals. The attributed values ε correspond to the medium value of the sub-interval of interest.

E. Multi-dimensional density estimation

At each ICE iteration, we need to estimate the multidimensional densities $f\varepsilon_n(y_n)$. Several strategies from multivariate data analysis are available, depending on the assumptions made on the statistical links between the layers, and on the choice of the shape of the one-dimensional densities. If independence between the layers is assumed, $f\varepsilon_n(y_n)$ is the product of M densities $g^1_{\varepsilon_n},...,g^M_{\varepsilon_n}$ defined on R:

$$f_{\varepsilon_n}(y_n) = \prod_{m=1}^{M} g^m_{\varepsilon_n}(y^m_n)$$

For example, if we consider that fεn(yn) are multidimensional

Gaussian densities, parameters estimation can be easily achieved from the first and second moments of a M-dimensional sample. Denoting by N(m, σ^2) the normal distribution with mean m and variance σ^2, the pdf of the hard classes are then expressed according to:

$\varepsilon_n = 0$: N(m0; σ20),

$\varepsilon_n = 1$: N(m1; σ21);

The parameters $\Delta=\{m_0,m_1,\sigma_0, \sigma_1\}$ can be estimated by computing the empirical mean of

several estimates according to $\theta^{q+1} = \frac{1}{L}\sum_{l=1}^{L} \hat{\theta}(x^l, y)$, where xl is an a posteriori realization of X conditionally on Y . It can be shown that X | Y is a non homogeneous Markov chain whose parameters can be computed from the forward and backward probabilities: (16) and (17).
The definition of the fuzzy measure Ā: "the pixel belongs to class 1" corresponding to the fuzzy class]0,1[, and its fuzzy membership function μ_A, allows us to estimate the fuzzy parameters of the set Δ in this new context. The proposed fuzzy membership function μ_A is defined by:

$$\mu_A(m) = \begin{cases} 1 - \frac{m_1 - m_1}{m_1 - m_0} \forall m \in [m_0, m_1] \\ 0 \quad \text{elsewhere} \end{cases}$$

Accordingly, the parameters of the Gaussian pdf for the fuzzy class can then be estimated by:

$$m_{\varepsilon_n} = (1-\varepsilon_n)m_0 + \varepsilon_{n,m}$$

$$\sigma_{\varepsilon_n} = (1-\varepsilon_n)^2\sigma_0 + \varepsilon_n{}^2\sigma_1^2$$

4 MULTI-SPECTRALES IMAGES CLASSIFICATION

This section is intended to evaluate the two local classification approaches, and the benefit of relaxing the HMCF model.

A. Validation

To validate the proposed approach we consider the synthetic image in Figure 4, we have considered a hard class(the opaque background and the umbrella), and the gradient between this class as the fuzzy zone (representing the border).Fig.4 present the results of two methods of classification based on HMC model and fuzzy HMC model .
We can be seeing that the image of Figure 4 appears to be classified much more precise with fuzzy CMC. The Taking into account a higher number of classes in the classical model of CMC did not have much interest because the additional classes tend to specialize on individual pixels.

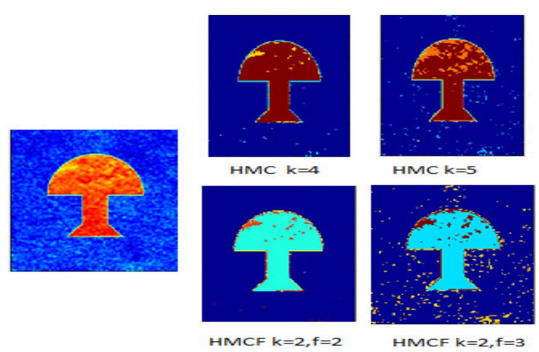

FIGURE 4: The synthetic image and their results of classification with CMC model (4 class and 5 class) and fussy CMC model (F=2,and F=3).

B. SPOTHRV Images Classification

1. The studied images

The image chosen for the study is that of the forest region of Rabat Morocco. This forest is an ecological unit consisting of a typical natural vegetation and reforestation.
Figure 5 shows the satellite images studied, and its multi-spectral images (green, red and near infrared) of the forest Rabat.

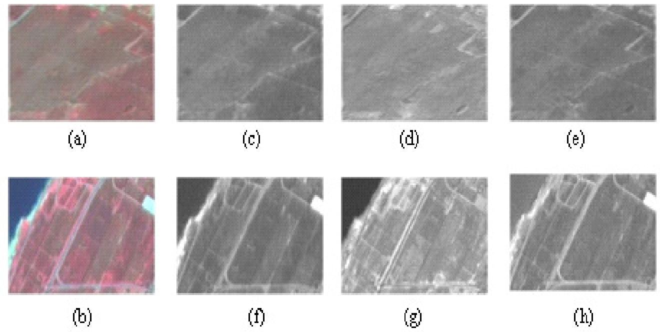

FIGURE 5: The two images studied (a and b) and their multi-spectral images (c, d and e) for the image (a) and (f, g and h) for the image (b).

HRV image - SPOT XS of this region, was acquired in December 1999 with a spatial resolution of 20m. in this image we extracted two test areas of size 140x90 pixels occupied mainly by a natural forest of orkoaks and eucalyptus planting, pine and acacia.
The reality of the ground was obtained from the official plans of the forest, followed by field verification. These areas are characterized by heterogeneity in the distribution offorest Stands. Discrimination between different textures is not obvious to human observer.

2. Experimental protocol
In all experiments, parameters initialization was done with a k-means classifier. The ICE algorithm was stopped after 10 iterations, assuming it has converged.

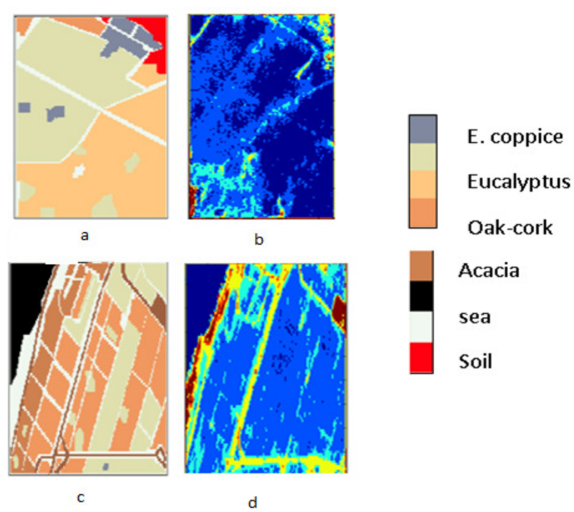

FIGURE 6: (a) and (c) Maps of the reality of the ground. (b) and (d) Results of classification with CMC (6 class) .

Figure 6 shows images of the realities ground (a and c) and the results of classification with CMC on both multi-spectral images of the forest in Rabat.(b and d). Computation times are 339.246seconds.
In this application, the “Sea” or the “Acacia” were considered as crisp class 0 and “Eucalyptus” or the “Oak-cork” crisp class 1. The fuzzy measure A thus corresponds to: “the pixel belongs to the free sea class”. The classification results depend on the partition of]0,1[, the choice of the F sub-intervals implies different values of fuzzy measure ε, e.g. F = 2 implies $\varepsilon \in \{0.25, 0.75\}$, F = 3 implies $\varepsilon \in \{0.165, 0.5, 0.825\}$, F = 4 implies $\varepsilon \in \{0.125, 0.375, 0.625, 0.875\}$
Figure 7 shows images of the realities ground (a and c) and a classification results obtained with fussy vectorial CMC models for F=2 sub-intervals(b and e) and F=3 sub-intervals (c and f). Computation times are 545.636 seconds.

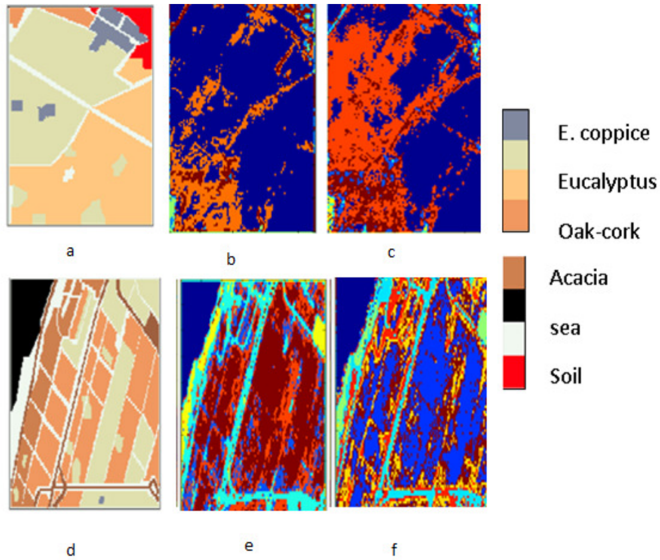

FIGURE 7: (a) and (d) Maps of the reality of the ground. Results of classification with fussy CMC for 4 crisp class and 2 fuzzy class (K=4,F=2) (b) and (e) and for 3 crisp class and 3 fuzzy class (k=3,F=3)) (c) and (f).

The computational complexity involved in the fuzzy vectorial HMC model is quite higher than the one involved in the classical one. One can observe that a big number of fuzzy sub-intervals provide a fine characterization of the observed scene. Furthermore, the classification results produced by the fuzzy vectorial HMC model are more homogeneous. The global shape and frontiers of the class seem to be better defined. Nevertheless, this model clearly characterizes the frontiers between class and produces a reliable classification of class, and then illustrates the interest of the vectorial fuzzy HMC model.

5 CONCLUSION

In this work, we propose to extend the fuzzy HMC model to unsupervised multi-spectral image classification. In multi-component case, it is interesting not only to take into account the uncertainty measure of the noisy observation (characteristic of probabilistic approach in classical HMC), but also the imprecision measure of this observation (characteristic of fuzzy approaches [6]) By adding a fuzzy measure in a statistical model, we obtain an original model, different from classical and models. Indeed, it preserves the robustness of the statistical classification (based on measures of uncertainty), and enriches it with the fuzzy characteristic (measure of imprecision).The Experimental results confirm the validity of the proposed approach. The vectorial fuzzy HMC model seems to be p promising in the field of multi-component image classification, due to the imprecision measure ability to take into account the multivariate information. We believe extending the fuzzy model to another model strictly more general than the CMC, called Markov chain couple [15] could also benefit from the fuzzy extension proposed in this paper.

REFERENCES

[1] W. Pieczynski. Statistical image segmentation. Mach. Graph. and Vis., 1 : 261-268, 1992.

[2] H. Caillol, W. Pieczynski, A. Hillion. Estimation of fuzzy Gaussian mixture and unsupervised statistical image segmentation. IEEE Trans. on Im. Proc., 6(3) : 425-440, 1997.

[3] W. Pieczynski. Modµeles de Markov en traitement d'images. Traitement du Signal, 20(3) :255{278, 2003.

[4] B. Benmiloud and W. Pieczynski. Estimation des paramètres dans les CMC et segmentation d'images. Traitement du Signal, 12(5) :433{454, 1995.

[5] W. Skarbek. Generalized Hilbert scan in image printing. In R. Klette and W. G. Kropatsch, editors, Theoretical Foundations of Computer Vision. Akademik Verlag, Berlin, 1992.

[6] S. Derrode, G. Mercier, and W. Pieczynski, "Unsupervised change detection in SAR images using a multicomponent HMC model," in MultiTemp'03, (Ispra, Italy) (2003).

[7] L. R. Rabiner, "A tutorial on HMMs and selected applications in speech recognition," Proc. IEEE 77(2), 257–286 (1989). [doi:10.1109/5.18626].

[8] N. Giordana and W. Pieczynski, "Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation," IEEE Trans. Pattern Anal. Mach. Intell.19(5), 465–475 (1997). [doi:10.1109/34.589206].

[9] P.-A. Devijver, "Baum's Forward-Backward algorithm revisited," Pattern Recogn. Lett. 3,369–373 (1985).

[10]L. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," Ann. Math. Stat. 41(1), 164–171 (1970).

[11]A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Roy. Stat. Soc. 39(1), 1–38 (1977).

[12]W. Pieczynski, "Statistical image segmentation," Mach. Graph. and Vis.,vol. 1, pp. 261–268, 1992.

[13]L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in Proc. IEEE, vol. 77, no. 2, Feb. 1989, pp. 257–286.

[14]P. A. Devijver, "Baum's forward-backward algorithm revisited," Pattern Recognition Letters, vol. 39, no. 3, pp. 369–373, december 1985.

[15]S. Derrode, C. Carincotte, and J.M. Boucher. Unsupervised image segmentation based on high-order hidden Markov chains. In IEEE ICASSP, Montreal (Ca), 17-21 mai 2004.