# Features Modelling in Discrete and Continuous Hidden Markov Models for Handwritten Arabic Words Recognition

Amine Benzenache[1], Hamid Seridi[1], and Herman Akdag[2]
[1]LabSTIC, University of 8 Mai 1945 of Guelma, Algeria
[2]LIASD, University Paris 8, France

**Abstract**: *The arab writing is originally cursive, difficult to segment and has a great variability. To overcome these problems, we propose two holistic approaches for the recognition of the handwritten arabic words in a limited vocabulary based on the Hidden Markov Models (HMMs): discrete with wk-means and continuous. In the suggested approach, each word of the lexicon is modelled by a discrete or continuous HMM. After a series of pre-processing, the word image is segmented from right to left in succession frames of fixed or variable size in order to generate a sequence vector of statistical and structural parameters which will be submitted to two classifiers to identify the word. To illustrate the efficiency of the proposed systems, significant experiments are carried out on IFN/ENIT benchmark database.*

**Keywords**: *Recognition of the handwritten arabic words, holistic approach, DHMMs, CHMMs, k-means, wk-means, algorithm of Viterbi, modified EM algorithm.*

*Received September 22, 2014; accepted April 23, 2015*



## 1. Introduction

The nature of Arabic script poses some problems in the conventional system of automatic recognition such as words that are not separated along a line, short connection between two successive characters in a word, and the presence of ligatures on the horizontal and vertical strings, which make very difficult the recognition in approaches based on analytical segmentation of characters [27]. A survey and bibliography on recognition of Arabic script covering all the research publications highlighting Arabic text characteristics is presented [2, 8, 27].

Certain research tasks were devoted to the isolated characters [7, 31], others were oriented towards the texts and invested in the segmentation of the words in characters taking into account the aspect related to variation of forms according to their position in the word [37]. Other works were also devoted to the recognition of the isolated handwritten words [4, 6].

According to the manner of perceiving a word, two main methods have been used in the literature, the holistic and analytical approaches. In the first approach, the word is seen as an indivisible entity which preserves the character in its context of vicinity but reduces the recognition vocabulary [6]. The second approach is based on modeling the alphabet of the language and word segmentation entities representing a character or pseudo-character; word recognition is to identify these entities and to propose words hypotheses; this approach is mainly related to the results of segmentation [28] and has the advantage of being able to handle an extended vocabulary.

Most of the works done are focused on combining structural, textural and local features [1, 14, 26], or on introducing a new feature [14]. Recently the problem of writer identification in a multi-script environment [25] and the utility of detecting and removing ruling lines, from handwriting documents, for writer identification [15].

The progress made in current years in handwritten Arabic script recognition is due in large part to statistical [23, 34, 36, 37, 47], and structural approaches [3, 5, 45], approaches using neural network [3, 12, 22, 23, 35, 43, 47] approaches based on hidden Markov models [10, 11, 18, 19, 20, 32, 41, 42, 44, 46] and approaches based classifiers combination [9, 23, 41, 47, 48]. Another approach based on Artificial Immune Systems (AIS) and inspired from computer science technique embodies the principles of biological immune-systems for tackling complex real-world problems such as handwritten Arabic script recognition [17].

In order to address the problems associated with the Arabic script processing, some researchers have focused on using a markovian holistic approach for handwriting recognition. This success can be attributed to the probabilistic nature of Hidden Markov Models (HMM) models, which can perform a robust modelling of the handwriting signal with huge variability and sometimes corrupted by noise [21, 30]. Compared to other recognition approaches, HMMs are distinguished by their ability to effectively model various sources of knowledge. Indeed, on one hand they offer a coherent integration of different levels of modeling and on the other hand, the existence of efficient algorithms for determining the optimum parameters providing the best models, so why HMMs have been used in several works such as those cited in [10, 16, 21, 29].

Motivated by these advantages, our research has been focused on developing recognition systems using Discrete Hidden Markov Models (DHMMs) and Continuous Hidden Markov Models (CHMMs) with global and embedded training, in which we choose an adequate set of statistical and structural features extracted from two types of segmentation, uniform and non uniform. To evaluate the performance of the suggested approaches tests are carried out on database of Tunisian cities names.

## 2. Related Works

The very successful use of HMMs in speech recognition has led many researchers to apply them to various problems in the field of handwriting recognition such as character recognition, offline word recognition.

Chen *et al.* [16] used HMMs with explicit state duration named continuous density variable duration HMM. After explicit segmentation of the word into subcharacters, the observations used are based on geometrical and topological features. Each letter is identified with a state which can account for up to four segments per letter. The parameters of the HMM are estimated using the lexicon and the manually labeled data. A modified Viterbi algorithm is applied to provide the N best paths. Khorsheed [29] has presented a method for offline handwritten script recognition, using a single HMM with structural features. The single HMM is composed by concatenation of multiple character models. After preprocessing, the skeleton graph of the word is decomposed into a sequence of links. The line segment sequence is transformed into discrete symbols by vector quantization. A modified Viterbi algorithm is applied to provide the N best paths. Benouareth *et al.* [10] describe an extended version of an offline unconstrained Arabic handwritten word recognition system based on segmentation-free approach and discrete HMMs with explicit state duration. After preprocessing intended to simplify the later stages of the recognition process, the word image is first divided according to two different schemes from right to left into frames using a sliding window. Then each frame is represented by a vector having 42 features. This latter sequence is submitted to an HMM classifier to carry out word discrimination by a modified version of the Viterbi algorithm. Several experiments have been performed using the IFN/ENIT benchmark database.

## 3. Hidden Markov Models

A process of Markov is a system at discrete time being at every moment in a state taken among N distinct states. The transitions between the states occur between two consecutive discrete moments, according to a certain law of probability. An HMM is defined by $\lambda = (A, B, \pi)$ where A denotes the transition matrix, B the matrix of discrete output probabilities corresponding to each possible discrete observation, and $\pi$ the initial state distribution [40]. The modeling of real processes by hidden Markov models requires resolution of three problems [9] such as: The assessment of the likelihood of observations sequence $O=\{o_1, o_2,\dots,o_T\}$ issued by a model $\lambda$; the search for the sequence of states $Q=\{q_1, q_2,\dots,q_T\}$ of $\lambda$ that produced the observations; the learning parameters from a model $\lambda$. In response to the three constraints, most of the researches focus on using simultaneously the Forward-Backward algorithm, the Viterbi algorithm and Baum-Welch algorithm. Discrete HMMs representing a continuous signal into a sequence via discrete vector quantization degrades the performance of the model. It would be advantageous to use hidden Markov models with continuous probability densities.

In the case of continuous HMMs, in addition to the matrix of probabilities of transition and the vectors of initial probabilities, the parameters of the emission probability density of the observations for each state were re-estimated [40].

## 4. Proposed Systems

The architecture of both developed systems based on DHMM and CHMM are shown respectively in Figures 1 and 2.

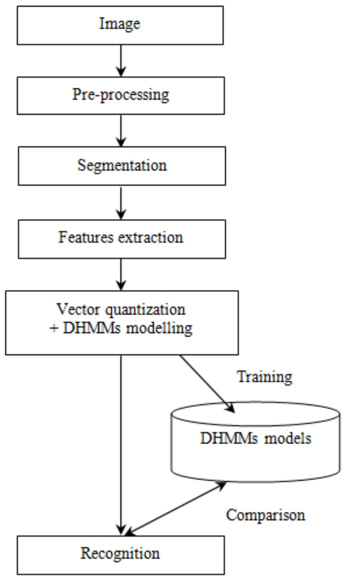


Figure 1. Architecture of discrete hidden markov models.

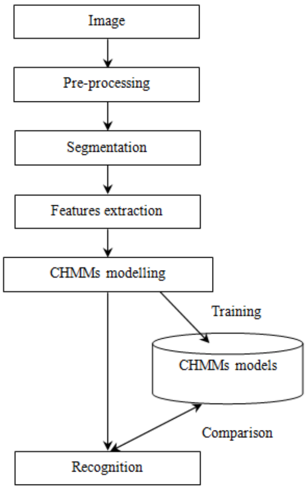


Figure 2. Architecture of continuous hidden markov models.

### 4.1. Pre-Processing

Pre-processing operations prepare the image of the word for the later stages of the recognition process. Practically these pre-processing include: binarization, smoothing, estimation of the baseline based on the analysis of the horizontal histogram projection after transformation of the binary image of the word into Hough space [39]. Skeleton methods using algorithm described in [38].

### 4.2. Segmentation

The word image is segmented from right to left into sequence frames of fixed or variable size [10, 20]. Uniform segmentation divides the frame into a series of windows having the same width; the optimum value of windows width is the one which gives the best recognition rate as shown in Figure 3. However in the non uniform segmentation, the windows widths of frames are not necessarily equal as shown in Figure 4. In both cases, the beginning and the end of each frame are determined by the vertical and horizontal histograms of the projected word.

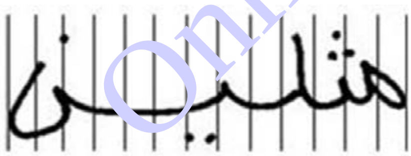

Figure 3. Uniform segmentation.

Figure 4. Non uniform segmentation.

### 4.3. Features Extraction

The step of feature extraction consists in extracting the sequence of vectors with structural and statistical features. These features are extracted from right to left, from the picture or the skeleton of the word using the technique of sliding windows with a uniform or non uniform segmentation. In our approach, frame is analyzed and characterized by a vector having 68 features of which 49 are statistical features related to pixels density and 19 structural features that take into account the local configuration, projection points, inflection points, reflection points, diacritical points and loops.

#### 4.3.1. Statistical Features

a) Histograms of projections and transitions features:

The features of a frame represented by F(n, m) are computed from histograms of projections and transition. In what follows, m and n are the length and width of the frame representing the word.

In the case of histograms of projections and transitions from white to black, for a given frame F(n, m) four directions have been considered in both cases and they are defined as:

- Horizontal projection

$$H_{ph}(i) = \sum_{j=1}^{m} F(i,j), \ 1 \le i \le n \quad (1)$$

- Vertical projection

$$H_{pv}(j) = \sum_{i=1}^{n} F(i,j), \ 1 \le j \le m \quad (2)$$

- 45° Diagonal projection

$$H_{pd+}(k) = \sum_{\substack{j=\min(k,m),\dots,1 \\ i=k-j+1,\, i \le n}} F(i,j), \ 1 \le k \le m+n-1 \quad (3)$$

- 135° Diagonal projection

$$H_{pd-}(k) = \sum_{\substack{j=\min(k,m),\dots,1 \\ i=k-j+1,\, i \le n}} F(i,j), \ 1 \le k \le m+n-1 \quad (4)$$

- Horizontal transition

$$HT_{NBH}(i) = NBTNB(F,i,H) \quad 1 \le i \le n \quad (5)$$

- Vertical transition

$$HT_{NBV}(j) = NBTNB(F,j,V) \quad 1 \le j \le m \quad (6)$$

- 45° Diagonal transition

$$HT_{NBD+}(k) = NBTNB\ (F,k,D+) \quad 1 \le k \le n+m-1 \quad (7)$$

- 135° Diagonal transition

$$HT_{NBD-}(k) = NBTNB\ \ (F,k,D-) \qquad 1 \le k \le n+m-1 \tag{8}$$

Where: NBTNB (F, L, D) is a function that evaluates the number of transition from white to black along a straight line L in the direction D from the frame F.

Based on histograms of projections and transitions, 33 statistical features have been computed from the following equations:

$$f_1 = \frac{\sum_{i=1}^{n} H_{ph}(i)}{n} \tag{9}$$

$$f_2 = \frac{\sum_{i=1}^{n} (H_{ph}(i) - f_1)^2}{n} \tag{10}$$

$$f_3 = mod\,e(H_{ph}) \tag{11}$$

$$f_4 = \frac{\sum_{j=1}^{m} H_{pv}(j)}{m} \tag{12}$$

$$f_5 = \frac{\sum_{j=1}^{m} (H_{pv}(j) - f_4)^2}{m} \tag{13}$$

$$f_6 = mod\,e(H_{pv}) \tag{14}$$

$$f_7 = \frac{\sum_{k=1}^{n+m-1} H_{pd+}(k) \times L(k)}{n \times m} \tag{15}$$

$$f_8 = \frac{\sum_{k=1}^{n+m-1} (H_{pd+}(k) \times L(k) - f_7)^2}{n \times m} \tag{16}$$

$$f_9 = mod\,e(H_{pd+}) \tag{17}$$

$$f_{10} = \frac{\sum_{k=1}^{n+m-1} H_{pd-}(k) \times L(k)}{n \times m} \tag{18}$$

$$f_{11} = \frac{\sum_{k=1}^{n+m-1} (H_{pd-}(k) \times L(k) - f_{10})^2}{n \times m} \tag{19}$$

$$f_{12} = mod\,e(H_{pd-}) \tag{20}$$

$$f_{13} = \frac{\sum_{i=1}^{n} HT_{NBH}(k)}{n} \tag{21}$$

$$f_{14} = \frac{\sum_{i=1}^{n} (HT_{NBH}(i) - f_{13})^2}{n} \tag{22}$$

$$f_{15} = mod\,e(HT_{NBH}) \tag{23}$$

$$f_{16} = min(HT_{NBH}) \tag{24}$$

$$f_{17} = max(HT_{NBH}) \tag{25}$$

$$f_{18} = \frac{\sum_{j=1}^{m} HT_{NBV}(j)}{m} \tag{26}$$

$$f_{19} = \frac{\sum_{j=1}^{m} (HT_{NBV}(j) - f_{18})^2}{m} \tag{27}$$

$$f_{20} = mod\,e(HT_{NBV}) \tag{28}$$

$$f_{21} = min(HT_{NBV}) \tag{29}$$

$$f_{22} = max(HT_{NBV}) \tag{30}$$

$$f_{23} = \frac{\sum_{k=1}^{n+m-1} HT_{NBD+}(k) \times L(k)}{n \times m} \tag{31}$$

$$f_{24} = \frac{\sum_{k=1}^{n+m-1} (HT_{NBD+}(k) \times L(k) - f_{23})^2}{n \times m} \tag{32}$$

$$f_{25} = mod\,e(HT_{NBD+}) \tag{33}$$

$$f_{26} = min(HT_{NBD+}) \tag{34}$$

$$f_{27} = max(HT_{NBD+}) \tag{35}$$

$$f_{28} = \frac{\sum_{k=1}^{n+m-1} HT_{NBD-}(k) \times L(k)}{n \times m} \tag{36}$$

$$f_{29} = \frac{\sum_{k=1}^{n+m-1} (HT_{NBD-}(k) \times L(k) - f_{28})^2}{n \times m} \tag{37}$$

$$f_{30} = mod\,e(HT_{NBD-}) \tag{38}$$

$$f_{31} = min(HT_{NBD-}) \tag{39}$$

$$f_{32} = max(HT_{NBD-}) \tag{40}$$

$$f_{33} = \frac{m}{n} \tag{41}$$

b) Densities features:

Let be *U* and *L* respectively the upper and lower values of the baseline. H represents the height in pixels of the sliding window (height of the bounding box of the word processed). The window is divided vertically into $n_C$ cells, that is $n(i)$ the number of pixels of writing (black pixels) in the cell i.

Let $b(i)$ the intensity of the cell $i$: $b(i)=0$ if $n(i)=0$, $b(i)=1$ otherwise. The characteristics of density are:

$f_{34}$: density of black pixels in the window.

$f_{35}$: number of transitions from black to white between cells.

$$f_{35} = \sum_{i=2}^{n} |b(i) - b(i-1)| \tag{42}$$

$f_{36}$:difference in position between the centers of gravity *g* of writing pixels in two consecutive windows:

$$f_{36} = g(t) - g(t-1) \tag{43}$$

$f_{37}$ to $f_{44}$: are the densities of pixels in each writing column of the window.

The following features depend on the estimated position of the baseline.

$f_{45}$: normalized vertical gravity center of the writing pixels compared to the low baseline.

$$f_{45} = \frac{g - L}{H} \tag{44}$$

$f_{46}$-$f_{47}$: writing pixel densities above and below the low baseline.

$f_{48}$: number of transitions from black to white between cells located above the low baseline.

$f_{49}$:zone belongs to the gravity center of writing in the window, it could be upper ($f_{49}$=1), middle ($f_{49}$=2) or lower ($f_{49}$=3).

### 4.3.2. Structural Features

a) Local configurations features:

Concavity features are features that provide local concavity information and stroke direction within each frame. Each concavity feature $f_{50}$ to $f_{54}$ represents the (normalized) number of white pixels (background) that belong to five types of concavity configurations.

They are explored by using a 3 x 3 window, as shown in Figure 5.

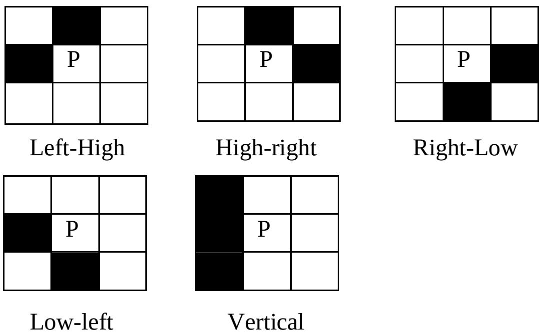


Figure 5. The five local configurations around a background pixel P

The following characteristics $f_{55}$ to $f_{59}$ depend on the positions of the baselines and the number of pixels located in the middle of writing between the two baselines.

b) Projecting points:

They correspond to points in the word skeleton with a number of neighbors different to 0 and 2. There are two types: the extreme points and junction points. An extreme point correspond the start / end of a line segment. A junction point connects three branches or more. For the latter, there are two types: the branch points and crossing points. $f_{60}$=E, $f_{61}$=B and $f_{62}$=C.

Where E, B and C are respectively the number of extremes points, branch points and crossing points in the frame F.

c) Inflection points

They correspond to the sign changes of the curvature. $f_{63}$=I.

Where I is the number of inflection points in the frame F.

d) Reflection points

They correspond to the skeleton points $P_i$ for which the sum of the smoothed curves of the previous sequence of points $P_i$ is positive (respectively negative). The overall smoothed curvature is defined by:

$$\delta_{is} = \theta_{(i+1,s)} - \theta_{(i-1,s)} \tag{45}$$

With

$$\theta_{ik} = Arctg\left[ (y_i - y_{(i-k)}) \Big/ (x_i - x_{(i-k)}) \right] \tag{46}$$

Where ($x_b$, $y_b$) are the coordinates of point P on the curve analysis (i.e., sequence of points in the skeleton), and S is a smoothing factor, in other words, the optimal interval for which the noise quantization is reduced and significant details are preserved in each point of the curve. For a smooth curve, the value of S must experimentally be in the range (5 to 15). The optimal value for S was empirically fixed at 7. $f_{64}$=R.

Where R is the number of reflection points in the frame F.

e) Diacritical points

These are the black pixels in the skeleton with a number of neighbors equal to 0. This type of points characterizes the formed characters of the secondary parts. These are distinguished by their positions relative to baseline (below or above). $f_{65}$=DB, $f_{66}$=DH.

Where DB and DH are the number of diacritic points being below or above the baseline in the frame F.

f) Loops

They correspond to the internal contours of the word image, and can be easily calculated from the skeleton of the word. The loops are distinguished by their degree of inclusion (integral or partial) in the frame in question. $f_{67}$=BP, $f_{68}$=BI.

Where BP and BI are the number of loops partially and completely included in the frame F.

## 4.4. Learning

The topology adopted for the global modeling of handwritten Arabic words is that of right-left with break inter-state and intra-states [40]. This type of model has the advantage of roughly keeping the notion of time in the modeling, thus approaching the nature of writing. However, we must define the different parameters in the model structure (“Number of states, N”, “transition class”, “Number of Gaussian, G”). The choices about the exact structure of the model can be identified fairly empirical and are still not easily justifiable. In our case, the values of numbers of states and Gaussians were empirically determined, those which give the best recognition rate. For learning, the algorithm of Baum-Welch based on a criterion of maximum likelihood is used for the optimization of HMMs parameters.

### 4.4.1. Vectors Quantization and Discrete HMMs Modeling

In the modeling of discrete HMMs, the words and their probability densities observations are discrete and require the use of a vector quantizer to match continuous vector to a discrete index of a dictionary of reference (CodeBook). Once the dictionary reference is obtained, the correspondence between the characteristic vectors of the frames and the indices of the dictionary is evaluated by the method of the nearest neighbor.

This procedure is achieved through three steps: the partition of the characteristic vectors obtained from all learning images in disjoined units; the representation of each set by a single vector ($v_k$=1≤k≤M), which is usually the centroïde vector of characteristics of the training set assigned to the same class and the optimization of the partition of the dictionary (the centroïde of each region). In our case, hybridizes Invasive Weed Optimization with K-means have been used; the solutions generated by the Invasive Weed Optimization Algorithm are used as initial solutions for the K-means algorithm [13].

Two types of modeling were implemented in the case of DHMM.

- The holistic learning: The concept of state is associated to the number of character in the word. Consequently, the number of HMM states varies from one model to another. For example, the model associated with the word "مارث" contains 4 states, whereas the number of states in the model of "القيروان" is 8 states.
- Embedded learning: Each character is modeled by an HMM model, thus the modeling of the word is carried out by the concatenation of the models characters, the number of states of characters is evaluated experimentally.

### 4.4.2. Continuous Hidden Markov Model

The choice of the number of Gaussian mixtures and the number of states is crucial in generating an appropriate model. In practice, the number of states and Gaussian model in learning process are fixed, where each frame may belong to any state with some probability.

### 4.4.3. Recognition

The classification of the word is based on a modified EM algorithm for mixture models based on the Bregman divergence [24]. For each model $\lambda_v$ the word v of lexicon L, the classifier computes the probability $p(O/\lambda_v)$ which corresponds to the probability of obtaining the sequence O of model $\lambda_v$.

## 5. Results and Discussions

To validate the suggested approach, significant experiments have been performed on the IFN/ENIT benchmark database composed of 26459 Arab words corresponding to a lexicon of 946 of Tunisian cities names written by 411 scriptwriters. This database is composed of four distinct units (a, b, c, d) among which three are used for training and one for testing. Several tests were carried out to evaluate the rate of recognition of the two systems, according to modeling types, number of states (N), size of quantization codebook (M), procedure of segmentation word image and the number of Gaussian (G).

Experimentally we define the values of parameters which give the best rate recognition. Tables 1 and 2 shows respectively the values of the parameters and the configuration of the database for learning and testing.

Table 1.DHMMs/CHMMs parameters and configurations.

| | Holistic learning | | Embedded learning | |
|---|---|---|---|---|
| | DHMMs | CHMMs | DHMMs | CHMMs |
| Number of gaussiennes (G) | X | 25 | X | 25 |
| Number of states (N) | Define | Define | 4 | 4 |
| Size of codebook (M) | 80 | X | 80 | X |

Table 2. Configurations used for learning and test

| Number of test | Learning | | Test | |
|---|---|---|---|---|
| | Data | Size | Data | Size |
| 1 | a, b, c | 20209 | D | 6250 |
| 2 | a, c, d | 20010 | B | 6449 |
| 3 | b, c, d | 19069 | A | 7390 |
| 4 | a, b, d | 20089 | C | 6370 |

The non uniform segmentation of handwritten Arabic words in a series of frames seems better suited to modeling based on CHMMs. Tables 3 and 4 shows that the average gain rate obtained in top 10 is 1,27% for embedded learning and 1,81 % for holistic learning. Indeed, this segmentation method has the advantage in giving in most cases, the frame that has the form of complete characters or parts of characters. On the other hand, the uniform segmentation process often generates frames whose form is a combination of characters which will lead to the recognition phase, the risk of confusion at character level allowed the choice to emit this frame.

Table 3. Recognition rates obtained with discrete HMMs modelling and embedded learning.

| Test | Segmentation | Top1 | Top2 | Top6 | Top10 |
|---|---|---|---|---|---|
| 1 | Uniform | 81,14 | 85,45 | 90,75 | 91,95 |
| | Non uniform | 82,79 | 86,66 | 91,83 | 92,89 |
| 2 | Uniform | 80,54 | 83,79 | 88,93 | 90,50 |
| | Non uniform | 81,62 | 85,15 | 90,33 | 91,80 |
| 3 | Uniform | 79,42 | 81,25 | 86,15 | 88,61 |
| | Non uniform | 80,74 | 83,75 | 88,46 | 90,63 |
| 4 | Uniform | 82,73 | 87,88 | 90,96 | 92,94 |
| | Non uniform | 84,46 | 88,62 | 91,66 | 93,75 |

Table 4. Recognition rates obtained with discrete HMMs modelling and holistic learning.

| Test | Segmentation | Top1 | Top2 | Top6 | Top10 |
|---|---|---|---|---|---|
| 1 | Uniform | 47,69 | 49,20 | 51,79 | 52,92 |
| | Non uniform | 50,95 | 52,06 | 53,59 | 54,97 |
| 2 | Uniform | 46,55 | 48,13 | 50,91 | 51,68 |
| | Non uniform | 48,75 | 50,24 | 51,35 | 53,11 |
| 3 | Uniform | 46,33 | 47,56 | 49,72 | 50,92 |
| | Non uniform | 48,82 | 49,63 | 51,66 | 52,76 |
| 4 | Uniform | 49,98 | 51,18 | 53,39 | 54,73 |
| | Non uniform | 51,84 | 53,19 | 54,24 | 56,66 |

The non uniform segmentation of handwritten Arabic words in a series of frames seems better suited to modeling based on continuous HMMs. Average gain rate is obtained in top 10 is 1,01% for embedded learning and 1,88 % for holistic learning as shown in Tables 5 and 6. In all performed tests, the best recognition rate is obtained experimentally by continuous Markovian modeling for 4 states with 25 Gaussian mixtures and non uniform segmentation.

Table 5. Recognition rates obtained with continuous HMMs modelling and embedded learning.

| Test | Segmentation | Top1 | Top2 | Top6 | Top10 |
|---|---|---|---|---|---|
| 1 | Uniform | 84,66 | 87,88 | 92,90 | 95,58 |
| | Non uniform | 85,37 | 89,18 | 94,47 | 96,80 |
| 2 | Uniform | 85,21 | 89,33 | 94,66 | 96,98 |
| | Non uniform | 86,88 | 90,75 | 95,93 | 97,44 |
| 3 | Uniform | 83,25 | 86,55 | 90,19 | 93,66 |
| | Non uniform | 84,42 | 87,88 | 92,75 | 94,32 |
| 4 | Uniform | 86,65 | 91,94 | 94,56 | 98,96 |
| | Non uniform | 89,29 | 92,77 | 95,86 | 99,66 |

Table 6. Recognition rates obtained with continuous HMMs modelling and holistic learning.

| Test | Segmentation | Top1 | Top2 | Top6 | Top10 |
|---|---|---|---|---|---|
| 1 | Uniform | 48,75 | 50,25 | 52,78 | 53,66 |
| | Non uniform | 51,55 | 53,44 | 54,70 | 55,19 |
| 2 | Uniform | 49,72 | 51,14 | 53,96 | 54,85 |
| | Non uniform | 52,58 | 54,36 | 55,79 | 56,65 |
| 3 | Uniform | 46,85 | 48,68 | 50,55 | 51,79 |
| | Non uniform | 49,42 | 51,21 | 52,44 | 53,88 |
| 4 | Uniform | 51,35 | 53,18 | 55,89 | 56,76 |
| | Non uniform | 53,24 | 55,41 | 56,65 | 58,87 |

Also note that in the non uniform segmentation, the average gain rate obtained with continuous HMM in top 10 is 4,79% for embedded learning and 1,78% for holistic learning compared to discrete HMM.

A comparison made between the proposed system (CHMMs) and others systems such as UOB [19], SCHMMs with explicit state duration [10], TH-OCR, REAM and ARAB-IFN shows that continuous HMMs provides better results with proposed set of features as shown in Table 7.

Table 7. Comparison with others systems which are represented in [10, 19, 33] : with the dataset d (67[illegible]5 images)

| Systems | Features | Top 1 | Top 10 |
|---|---|---|---|
| UOB | 24 statistical | 85,00% | 93,56% |
| SCHMMs with explicit state duration | [illegible]3 statistical + 9 structural | 89,08% | 95,98% |
| TH-OCR | structural | 30,13% | 46,59% |
| REAM* | Structural | 89,06% | 99,62% |
| ARAB-IFN | Statistical | 87,94% | 95,62% |
| Proposed DHMMs | 49 statistical + 19 structural | 84,46% | 93,75% |
| **Proposed CHMMs** | 49 statistical + 19 structural | 89,29% | **99,66%** |

(*)The system is trained on a reduced set with 1000 names

## 6. Conclusions

In this paper, we have proposed a modeling approach based on hidden Markov models, discrete and continuous. Many improvements have been planned such as: the learning strategy, the algorithm display, the selection and extraction of features from the word image. We have adopted two methods for segmenting the word into a sequence of frames namely uniform and non uniform segmentation. On the other hand, the nonuniform segmentation scheme is more appropriate than the uniform one.

We have studied the influence of different parameters that are involved in the construction of a Markov model such as the number of states and the number of Gaussians. Frame is analyzed and characterized by a vector having 68 components and combining a new set of relevant statistical related to pixels density and structural features that take into account the local configuration, projection points, inflection points, reflection points, diacritical points and loops. The performances of recognition obtained by our approach are acceptable: they have shown that modeling word recognition using hidden Markov models with continuous probability density improves the discriminating power and speeds up the computation time. The rates of goodness recognition using continuous hidden Markov model with non uniform segmentation and embedded learning are higher than those obtained by discrete HMMs; this improvement varies from 2,52% to 5,91%.

In addition, the recognition rates obtained by our system based CHMMs are greater than those obtained by UOB, SCHMMs with explicit state duration, DHMMs, TH-OCR, REAM and ARAB-IFN. The experiments show that significant improvements can be achieved by using a strategy that reinforces the classifier performances by using an adequate mechanism of selecting the parameters or attributes which are most relevant to an appropriate approach.

## Acknowledgments

The authors sincerely wish to thank Samir Beloucif and Amar Bendjazia for their useful comments.

## References

[1] Abdi M. N. Khemakhem M. and Ben-Abdallah H. "Off-Line Text Independent Arabic Writer Identification Using Contour Based Features," *In International Journal of Signal and Image Processing*, vol. 1, pp.4-11, 2010.

[2] Al Badr B. and Mahmoud S. A., "Survey and Bibliography of Arabic Optical Text Recognition," *Signal Processing*, vol. 41, pp. 49-77, 1995.

[3] Al-Ma'adeed S., "Recognition of off-ligne handwritten Arabic words using neural network," *Proceeding of GMAI'06*, International Conference on Geometric Modeling and Imaging, pp.141-114, London, England, July 2006.

[4] Almuallim H. and Yamaguchi S., "A method of recognition of Arabic cursive handwriting,"

*IEEE Transactions on PAMI*, vol. 9, no. 5, pp.715-722, 1987.

[5] Alsallah B. and Safadi H., "An Arabic online handwriting recognition system," *Proceeding of ICTTA'06*, 2nd IEEE International conference on Information et Communication Technologies : from Theory to applications, vol.1, 1844-1849, Damascus, Syria, April 2006.

[6] Ameur A. Romeo-Pakker K. Miled H.and Cheriet M., "Approche Globale pour la Reconnaissance de Mots Manuscrits Arabes," *Actes CNED'94*, 3ème Colloque National sur l'Ecrit et le Document, pp.151-156, 1994.

[7] Amin A., "Off-Line Arabic Character Recognition: the State of the Art. *Pattern Recognition*," vol. 31, no. 5, pp.517–530, 1998.

[8] Amin A. Kaced A. Haton J. P. and Mohr R., "Handwritten Arabic Character Recognition by the IRAC System," *Proceedings of ICPR'80, 5th International Conference on Pattern Recognition*, pp.729-731, 1980.

[9] Azizi N. Farah N. and Sellami M., "Off-line handwritten word recognition using ensemble of classifier selection and features fusion," *Journal of Theoretical and Applied Information Technology JATIT*, vol. 14. 2, pp.141-150, 2010.

[10] Benouareth A. Ennaji A. and Sellami M., "Arabic Handwritten Word Recognition using HMMs with Explicit State Duration," *EURASIP Journal on Advances in Signal Processing*, vol. 2008.

[11] Biadsy F. El-Sana S. and Habash N., "Online Arabic handwriting recognition using Hidden Markov Models," *Proceeding of IWFHR'06*, 10th International Work shop on Frontiers in Handwriting Recognition, pp.85-90, La Baule, France, 2006.

[12] Blumenstein M. and Verma B., "A Segmentation Algorithm used in Conjunction with Artificial Neural Networks for the Recognition of Real-World Postal Addresses," *ICCIMA*, Australia, 1997.

[13] Boobord F., Othman Z. and Abu Bakar A.," A WK-means Approach for Clustering," *IAJIT*, August 2013.

[14] Bulacu M. Schomaker L. and Brink A., "Text-Independent Writer Identification and Verification on Off-Line Arabic Handwriting," *In 9th ICDAR*, vol. 2, pp.769-773, 2007.

[15] Chen J. Lopresti D. and Kavallieratou E., "The Impact of Ruling Lines on Writer Identification," *In the Proceedings of the 12th International Conference on Frontiers in Handwriting Recognition*, Kolkata, India, pp.439-444, November 2010.

[16] Chen M Y. Kundu A. and Srihari S. N. "Variable duration hidden Markov model and morphological segmentation for handwritten word recognition," *IEEE Transactions on Image Processing*, vol. 4, no. 12, pp. 1675–1688, 1995.

[17] Djeddi C. and Souici-Meslati L., "Artificial Immune Recognition System for Arabic Writer Identification," *Fourth International Symposium on Innovation in Information and Communication Technology*, pp.159-165, 2011.

[18] El-Hadj R., "Reconnaissance hors ligne de textes manuscrits cursifs par l'utilisation de systèmes hybrides et de techniques d'apprentissage automatique," *Thèse de doctorat*, Ecole Nationale Supérieure de Télécommunications, Paris, Juillet 2007.

[19] El-Hadj R. Likeforman-Seulem L. and Mokbel C., "Arabic Handwriting Recognition Using Baseline Dependant Features and Hidden Markov Models," *Proceedings of ICDAR'05*, International Conference on Document Analysis and Recognition, Seoul, Korea, vol. 2, pp.893-897, 2005.

[20] El-Hadj R. Mokbel C. and Likeforman-Seulem L., "Reconnaissance de l'écriture arabe cursive: combinaison de classifieurs MMCs à fenètres orientées," *Actes CIFED'06*, 9èmeColloque International Francophone sur l'Ecrit et le Document, pp.271-276, Fribourg, Suisse, 2006.

[21] El-Yacoubi M. A. Gilloux M. Sabourin R. and Suen C. Y., "An HMM-based approach for off-line unconstrained handwritten word modeling and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp.752–760, 1999.

[22] Farah N. Ennaji A. Kadir T. and Sellami M., "Benefit of multiclassifier systems for Arabic handwritten words recognition," *ICDAR'05*, 8th International Conference on Document Analysis and recognition, vol.1, pp.222-226, Seoul, August-September 2005.

[23] Farah N. Souici L. and Sellami M., "Classifiers combination and syntax analysis for Arabic literal amount recognition," *Engineering Applications of Artificial Intelligence*, vol. 19, no. 1, pp.9-39, February 2006.

[24] Fujimoto Y. and Murata N. "A modified EM algorithm for mixture models based on Bregman divergence," *Springer*, 2007.

[25] Garian U. and Paquet T., "Off-Line Multi-Script Writer Identification Using AR Coefficients," *ICDAR 2009*, Spain, pp.991-995, July 2009.

[26] Gazzah S. and Ben Amara N. E., "Arabic Handwriting Texture Analysis for Writer Identification Using the DWT-Lifting Scheme," *In 9th ICDAR*, vol. 2, pp.1133-1137, 2007.

[27] Kammon W. and Ennaji A., "Reconnaissance de textes arabes à vocabulaire ouvert. Laboratoire Perception," *Système Information (PSI) FRE-CNRS 2645*, Université de Rouen 76821 Mont Saint Aignan Cedex, France, 2001.

[28] Khorsheed M. S., "Off-line arabic character recognition– a review," *Pattern Anal*, Appl. 5, pp.31-45, 2002.

[29] Khorsheed M. S., "Recognising handwritten Arabic manuscripts using a single hidden Markov model," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2235–2242, 2003.

[30] Koerich A. L. Sabourin R. and Suen C. Y., "Recognition and verification of unconstrained handwritten words," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp.1509–1522, 2005.

[31] Mahmoud S. A., "Arabic Character Recognition Using Fourier Descriptors and Character Contour Encoding," *Pattern Recognition*, vol. 27, no. 6, pp.815-824, 1994.

[32] Margner V. El Abed H. and Pechwitz M., "Off-Line Handwritten Word Recognition Using HMM- a Character Approach Without Explicit Segmentation," *Actes CIFED 06*, 9ème Colloque International Francophone sur l'Ecrit et le Document, Fribourg, Suisse, pp.259-264, 2006.

[33] Margner V. Pechwitz M. and El Abed H., "ICDAR 2005 Arabic handwriting recognition competition," *In Proceedings of the International Conference on Document Analysis and Recognition*, (ICDAR '05), vol.2005, pp. 70–74, 2005.

[34] Mezghani N., "Densités de probabilité d'entropie maximale et mémoires associatives pour la reconnaissance en ligne de caractères Arabes," *Thèse de PHD*, institut national de la recherche scientifique (INRS), centre énergie, Matériaux et télécommunication, Canada, Juin 2005.

[35] Mezghani N. Mitiche A. and Cheriet M., "A new representation of shape and its use for high performance in online Arabic character recognition by an associative memory," *International Journal on Document Analysis and recognition*, IJDAR, vol.7, no.4, pp.201-210, September 2005.

[36] Mezghani N. Mitiche A. and Cheriet M., "Estimation de densités de probabilité par maximum d'entropie et reconnaissance bayesienne de caractères Arabes en ligne," *RFIA' 06*, 15ème congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle, Actes sur CD-ROM, Tours, Janvier 2006.

[37] Olivier C. Miled H. Remeo K. and Lecourtier Y., "Segmentation and Coding of Arabic Handwritten Words," *Proceedings 13th International Conference on Pattern Recognition (ICPR 96)*, Vienne, Autriche, pp.264-268, 1996.

[38] Pavlidis T., "Algorithms for Graphic and Image Processing. Rockville," *MD: Computer science press*, pp.195-200, 1982.

[39] Pechwitz M. and Maergner V., "HMM Based Approach for Handwritten Arabic Word Recognition Using the IFN/ENIT- Database," *ICDAR (Proceedings of the Seventh International Conference on Document Analysis and Recognition)*, pp.890-894, 2003.

[40] Rabiner L. R., "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp.257-286, 1989.

[41] Rahman A. F. R. and Fairhurst M.C., "Multiple classifier decision combination strategies for character recognition," *International Journal on Document Analysis and Recognition*, IJDAR, vol.5, no.4, pp.166-194, July 2003.

[42] SaonG., "Modèles Markoviens uni-bidimensionnels pour la reconnaissance de l'écriture manuscrite hors-ligne," *Thèse de doctorat*, Université Henri Poincaré Nancy 1, 1998.

[43] Sari T. and Sellami M., "Cursive Arabic script segmentation and recognition system," *International Journal of Computers and Applications*, vol. 27, no.3, 2005.

[44] Slimane F. Kanoun S. Hennebert J. Alimi A. M. and R. Ingold, "Modèles de Markov Cachés et Modèle de Longueur pour la Reconnaissance de l'Ecriture Arabe à Basse Résolution," *MajecSTIC*, Avignon, France, November 2009.

[45] Souici L. Aoun A. and Sellami M., "Versune architecture modulaire de reconnaissance de montants de chèques arabes," *Proceedings de la Conférence Internationale en Informatique*, CII'99, Annaba, Algérie, Numéro spécial de la revue Synthèse (publication de l'université d'Annaba, Algérie), no.6, pp.260-270, November 1999.

[46] Vinciarelli A. Bengio S. and Bunke H., "Off-Line Recognition of Unconstrained Handwritten Texts Using {HMM}s and Statistical Language Models," *IEEE Transactions on PAMI*, vol. 26, no. 6, pp.709-720, 2004.

[47] Zermi N. Ramdani M. and Bedda M., "Arabic handwriting word recognition based on hybride HMM/ANN approach," *International Journal of Soft Computing*, vol.2, no.1, pp.5-10, 2007.

[48] Zouari H. Heutte L. and Lecourtier Y., "Controlling the diversity in classifier ensembles trough a measure of agreement," *Pattern Recognition*, vol.38, no.1, pp.2195-2199, November 2005.

**Amine Benzenache** received his Engineer degree in Electronics from University of 08Mai 1945, Guelma, Algeria. in 2004 and the Magister degree in Intelligence artificial University of Guelma in 2007. Currently, he is a researcher in the laboratory LabSTIC, University of 8 Mai 1945 of Guelma and member in GADM team at LabSTIC laboratory, Guelma, Algeria. Her current research interests include Handwriting recognition, HMM, pattern recognition and document processing.

**Hamid Seridi** received his Bachelor's degree with honours in 1981, from the University of Annaba, Algeria, and the Master's degree from the Polytechnic Institute of New-York, USA in 1984, both in Electrical Engineering. He received his PhD in Computer Science with distinction in 2001 from the University of Reims, France. He was Vice Dean of the Post-Graduation, Scientific Research and External Relations in the University of Guelma. Currently he is Professor and Director of Laboratory of Science and Information Technologies and Communication "LabSTIC". He is also Chairman of the Scientific Council of the Faculty of Mathematics and Computing and Material Sciences. He is an expert member at the national committee for evaluation and accreditation national projects research. His research interests include approximate knowledge management, pattern recognition and artificial intelligence, data mining, video mining, machine learning and cryptography.

**Herman Akdag** received his PhD and HDR degree at Paris 6 University in 1980 and 1992, respectively. Assistant Professor since 1980, he obtained a Full Professor position at Reims University, France in 1995. He is was a Senior Researcher at LIP6, CNRS, Paris and the head of the research group MODECO in University of Reims. Currently, he is a Full Professor at the University of Paris 8 and member at LIASD, France. His research interests include Fuzzy Set Theory and Machine Learning approaches to decision-making, image classification and image retrieval. He also works on approximate reasoning, fuzzy abduction, data mining, and user modelling.