ACIT'2017
The International Arab Conference on Information Technology
Yassmine Hammamet, Tunisia
December 22-24, 2017

Recognizing handwritten Arabic words using optimized character shape models and new features

[1]Anis Mezghani and Monji Kherallah

[1]National School of Engineers of Sfax, Tunisia

Faculty of Sciences of Sfax, Tunisia

***Abstract*—In this paper, we propose a segmentation-free word recognition system based on Hidden Markov Models (HMMs) where 48 features are computed on a fixed-length sliding window, some of them are never used in literature for the recognition task. The literature has proved the difficulty of Arabic text recognition systems to recognize all character shapes which are more than 170 shapes derived from 28 basic letters. To make training and recognition of characters more efficient, we used optimized character shape models to represent the different handwritten Arabic characters. Several experiments have been performed using the IFN/ENIT database of handwritten Arabic words. The results reveal the robustness and the effectiveness of our system.**

Keywords-Arabic text recognition; Hidden Markov Model; character shape models; profile based features; sliding window; IFN/ENIT database

I. INTRODUCTION

The last few score of years have witnessed a significant increase of research in off-line Arabic handwritten character recognition. Major challenges comprise the ligature and the interconnection of neighboring characters, the similarities between different character shapes and the variability in the shape of handwritten Arabic glyphs across different writers, as well as across different instances generated by the same writer [1].

In segmentation-free text recognition systems based on HMMs, the text images are usually modeled as the concatenation of different character shapes. These systems share similar theoretical foundations but hold opposing views on the model structures and parameters. The aim of this paper is to introduce a segmentation-free word recognition system using optimized character shape models representing the different handwritten Arabic characters. A set of new features applied on word images without their diacritical marks is proposed to represent better the selected character shapes. The performance of the proposed system has been tested on the IFN/ENIT database [2] containing handwritten Tunisian city names.

The remainder of this paper is organized as follows: Section 2 outlines the notable contributions in Arabic word recognition by using the IFN/ENIT database. Section 3 is devoted to the features extraction stage. Section 4 presents our choice for Arabic character shape modelling. Section 5 describes the training and the classification process of the proposed HMM based system. Experimental results using the IFN/ENIT database are reported in Section 6. Finally, conclusions are drawn in Section 7.

II. RELATED WORKS

Many handwriting Arabic recognition systems based on HMMs were proposed and achieved top results on the IFN/ENIT database. Al-Hajj et al. combined in [3] three HMM classifiers with the same topology. Three types of sliding window were used for feature extraction. This approach allows greater tolerance for the alignment of diacritical marks on the letter with which they are associated. The combination of the three classifiers thus achieves better performance than each of the three taken separately. The authors carried out their experiments using the IFN/ENIT database. They obtained the best recognition rate (90, 96% in top-1) when the fusion of the results was carried out with MLP.

In [4], the authors extracted different statistical and structural features using two strategies of segmentation in vertical bands. The authors proved that the non-uniform segmentation strategy produces more satisfactory results than the uniform one. They also proved that the use of Discrete HMM with explicit state duration modeling is more efficient than the use of the standard semi-continuous HMM. The best recognition rates obtained with 160 HMM models on the IFN/ENIT database (v1.0p2) was 89.08%.

AlKhateeb et al. presented in [5] an Arabic handwriting recognition system based on HMMs. Two types of features were used in this work: structural features (number of connected regions above and below the writing line) and intensity features (intensity of pixels in different local configurations). The intensity features were extracted from a sliding window moving on the mirror of the word image. The system was tested on the IFN/ENIT database and achieved a recognition rate of 83.55%.

Parvez et al. [6] presented a recognition system based on structural techniques. A segmentation algorithm based on the polygonal approximation was integrated into the recognition phase. Each text-line is thus segmented into words and pseudo-words. The characters are then modeled by "fuzzy" polygons that are later recognized using a fuzzy polygon matching algorithm. The dynamic programming is used to select the best assumptions of a sequence of recognized characters for each word/pseudo-word. Parvez et al. reported a recognition rate of 79.58% on the IFN/ENIT database.

Lawgali et al. presented in [7] a system for the recognition of Arabic handwritten words. In the training phase, the highest value of DCT coefficients of each Arabic handwritten character are extracted in zigzag mode and stored in a feature vector. In this phase, the HACDB database [8] containing 6600 Arabic handwritten character shapes, including ligatured characters, is used. In the test phase, the Arabic handwritten words are segmented. Then, the feature extraction is done by DCT while the classification is carried out by ANN. The classification is thus carried out in two stages: classification of the segmented characters and classification of the word. The system was tested on the IFN/ENIT database and an accuracy of 90.73% was reported.

The majority of the Arabic handwriting recognition systems cited in the literature are based on HMMs to build word, PAW, or letter models. Other techniques were also used and yielded satisfying results. The systems using PAW or word models for training suffer generally from limitations in recognizing words or PAWs that are not represented in the training set. In this paper, each character is modeled seperately and the word model is then built by character shape models concatenation. The system presented below gives a hight importance to the character shape models optimization by grouping similar shapes.

III. FEATURE EXTRACTION

The central idea of the proposed word recognition system is to identify and concatenate the different character shapes toward finding the word hypotheses. To attain that goal, a set of new profile based features is proposed and combined with other distribution features. All features are computed from a fixed width window of 10 pixels sliding from right to left and shifted by only one pixel. Thus, for each window, the feature extraction results in a vector of 48 coefficients including distribution features calculated on preprocessed word images and profile based features computed on the same word images after detection and removing of diacritical marks by the method described in [9]. Each word image is therefore transformed into a matrix of values where the number of rows corresponds to the number of analysis windows (n), and the number of columns is equal to the number of features. The feature extraction process is illustrated in Fig. 2.

A. *Distribution features*

Distribution features are generally based on foreground/background pixels densities and distributions in the analysed image. For each analysis window, the extracted features are the following:

- Density of foreground pixels
- 10 features representing the densities of black pixels for each column of pixels in the window;
- Number of black connected components;
- Number of white connected components;
- Number of background (white) pixels corresponding to P and representing the concavity configurations belonging to the five types of typological masks presented in Fig. 1;

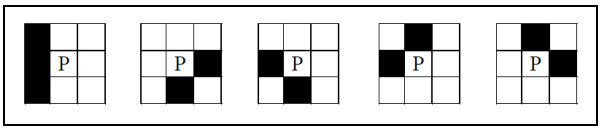

Figure 1. Typological masks around a background pixel P

- Six affine moment invariants proposed in [10], which are pure statistical measures of the pixels distribution around the center of gravity of the shape allowing the capture of the global shape information.

B. *Profile based features*

The profile based features represent the layout and the dispersion of profiles pixels on the X-axis and the Y-axis. Fig. 3 illustrates an example of drawing the raw top, bottom, left and right profiles of an analysis window from a word image. For each raw profile, six features could be calculated :

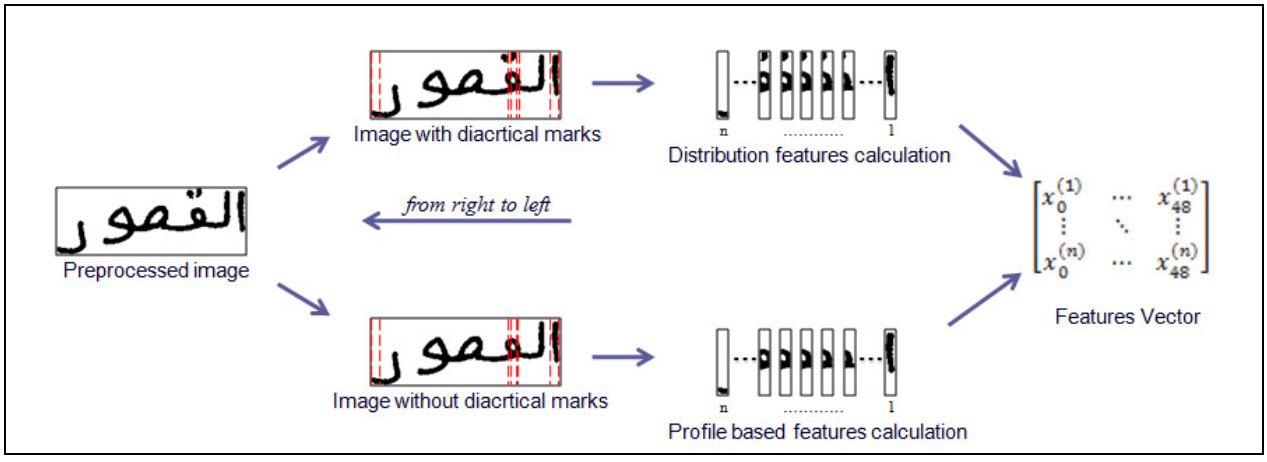

Figure 2. Feature extraction process

- The arithmetic mean of variances of the profile in the limits of the height/width of the shape, represented by the equation (1).

$$Ra = \frac{1}{H}\sum_{i=1}^{i=H} |y_i| \quad (1)$$

- The arithmetic mean of the absolute values of the differences between neighbors of the profile in the limits of the height/width of the shape, given by the equation (2).

$$DRa = \frac{1}{H-1}\sum_{i=1}^{i=H-1} |y_{i+1} - y_i| \quad (2)$$

- The quadratic mean of variances of the profile in the limits of the height/width of the shape, represented by the equation (3).

$$Rq = \sqrt{\frac{1}{H}\sum_{i=1}^{i=H} y_i^2} \quad (3)$$

- The quadratic mean of the absolute values of the differences between neighbors of the profile in the limits of the height/width of the shape, calculated by the equation (4).

$$DRq = \sqrt{\frac{1}{H-1}\sum_{i=1}^{i=H-1} |y_{i+1} - y_i|^2} \quad (4)$$

- The skewness, given by the equation (5) and describing the asymmetry of the values distribution of the profile.

$$Sk = \frac{1}{Rq^3}\frac{1}{H}\sum_{i=1}^{i=H} y_i^3 \quad (5)$$

- The kurtosis, given by the equation (6) and measuring the degree of the distribution of the profile deviations by report to the Gaussian distribution.

$$Ek = \frac{1}{Rq^4}\frac{1}{H}\sum_{i=1}^{i=H} y_i^4 \quad (6)$$

H represents the height of the shape in the case of right and left profiles, and the width in the case of top and bottom profiles. y_i represents the amplitude of the profile in position i by report to the width of the shape in the case of right and left profiles and by report to the height in the case of top and bottom profiles.

IV. CHARACTER MODELLING

Mezghani et al. [1] presented a study on character modelling for handwritten Arabic text recognition. By counting the number of character shapes in Arabic writing, they obtained the following observations:

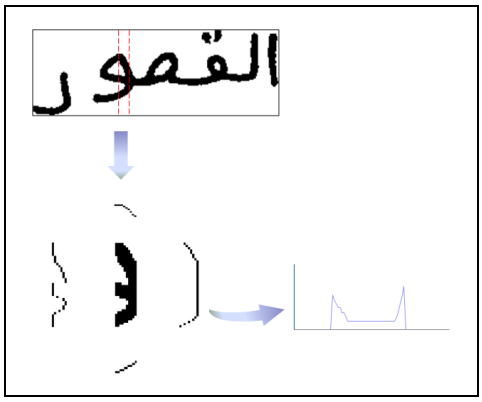

Figure 3. Raw top, bottom, left and right profiles of an analysis window

- The total number of shapes taking into account the positions of Arabic letters is equal to 100.
- Other shapes could be obtained by adding "Hamza" (ء) to some characters ("waaw" (و), "alif" (ا), or "alif maqsura" (ى)).
- Arabic letters could be combined with "Shadda" (ـّـ) which forms other character shapes.
- Vertical and horizontal ligatures could occur between two or three characters. The character jiim (ج), for example, could be ligatured with its neighboring characters in 67 ways [1].

Taken into account all these shapes, we obtain an enormous number of HMM models, which affect certainly the recognition rate of any system. The majority of Arabic recognition systems developed for the IFN/ENIT database used between 158 and 160 character shape models. Based on the experiments carried out on IFN/ENIT database by Mezghani et al. [1] to find the best character shape modeling, we used 61 shapes modeled as isolated HMMs. In these models, the different shapes are represented without ligatures and "Shadda" and the visually similar ones are grouped.

V. HMM-BASED CLASSIFICATION

HMMs, introduced by Baum et al. in 1970 [11], are statistical models based on strong theoretical foundations permitting the modeling of systems with a behavior partly predictable. HMM has been largely used for handwritten text recognition, especially in modeling connected character shapes of Arabic script ([12], [3], [13]). In this work, the bakis topology is employed with 8 states, each of which is associated with a probability distribution [14]. A set of state-transition probabilities is defined to allow a state to reach another one in a single step.

The training process is performed using the Expectation Maximization (EM) algorithm to estimate HMM parameters of a character shape model on the training set. The EM algorithm consists in iteratively estimating the component weights, the means and the variances of character shape models to increase the likelihood estimates of these HMM parameters [15]. During

training, we applied a simple binary splitting procedure after every 7 iterations to increase the likelihood value of the estimated model. A stabilization of the performance could be noticed after 256 Gaussian mixtures. The HMMs relating to the word recognition lexicon are built during the training process by linking their character shape models. The character shape sequence of the input image providing the best likelihood identifies the recognized word image using the Viterbi algorithm.

VI. RESULTS AND DISCUSSION

To demonstrate the feasibility and the validity of the developed system, several experiments were carried out on the publicly available IFN/ENIT database. In these experiments, five different sets (a, b, c, d and e) were used for training and testing. Fig. 4 summarizes the experimental results of the proposed recognition system using the different number of character shape models tested in [1] under the same system settings. These results prove the superiority of the word recognition rate when using 61 character shape models. Following this direction, we extracted the features described in section III to model each character shape with 8 states.

In order to compare the experimental results of the proposed recognition system to the works presented in the literature, we report on Table I the word recognition rate obtained on the IFN/ENIT database using two training–test configurations: "abc–d" and "abcd–e". In these experiments, we tried to highlight the importance of the new features related to the dispersion of profile pixels on the X-axis and the Y-axis. We investigated two features sets: set DF including the 24 distribution features presented in Section III.A, and set PBF including the 24 proposed profile based features presented in Section III.B. The use of set DF solely as a feature vector leads to a word recognition rate of 71.48% for set d and 59.03% for set e. After adding set PBF relative to the profile based features, the recognition rate increases to 91.22 % for set d and 82.78% for set e.

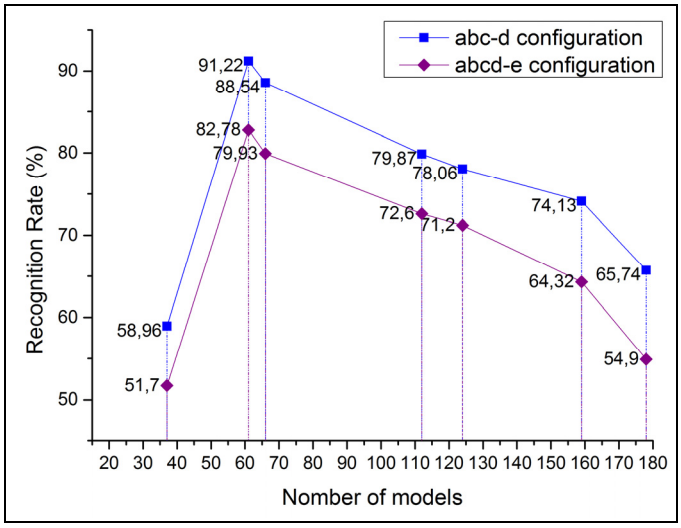

Figure 4. Recognition results using different number of character shape models

TABLE I. COMPARATIVE RECOGNITION RESULTS USING "ABC–D" AND "ABCD–E" CONFIGURATIONS OF IFN/ENIT DATABASE

Authors	Classifier	*Recognition rate(%)*	
		abc–d	*abcd–e*
Pechwitz and Margner [12]	HMM	89.74	
Benouareth et al. [4]	HMM	83.79	
Al-Hajj et al. [3]	HMM	87.6	
Hamdani et al. [16]	Multiple HMM		81.93
Elbaati et al. [17]	HMM		54.13
Xiang et al. [18]	HMM	84.09	
Kessentini et al. [19]	Multi-stream HMM		79.6
Alkhateeb et al. [5]	HMM	86.73	
Giménez et al. [20]	Bernoulli HMMs		**84**
Parvez et al. [6]	FATF with set- medians		79.58
Jayech et al. [21]	DBN	82	78.5
Proposed System (DF)	HMM	71.48	59.03
Proposed System (DF+ PBF)	HMM	**91.22**	82.78

The proposed system has accuracies comparable with the state of the art results on the same data. However, the developed system shows weakness due to the cursive nature of Arabic handwritten words and the high variability of shapes coming from the human writing styles and habits. In addition, the failure in detection and elimination of diacritical marks may causes inexact feature extraction, and consequently wrong classifications. These diacritical marks are often not at the exact position below or above the main part of the letter. Some other recognition errors can be attributed to the insufficiently representation of some character shapes in the database which implies the poor training of their models.

Finally, it must be stated that the main purpose of the carried experiments is to establish the contribution of reducing the number of HMMs by capturing the variations between the different character shape models, as well as the contribution of adding the profile based features to the feature vector. Regarding this investigation, the results show in both training–test configurations ("abc–d" and "abcd–e") the same important improvement of recognition performance when using only 61 character shape models and additional features based on top, bottom, left and right profiles.

VII. CONCLUSION

In this paper, a segmentation-free word recognition system based on HMMs, commonly known as powerful tools for sequence modelling, has been proposed. Led by the frequently observed errors, different character shape HMMs were used allowing the sharing of common models between different shapes of an Arabic character as well as between several characters. This led us to a robust and efficient recognition with only 61 models. For feature extraction, a fixed-length sliding window from right to left was used with no need of a priori segmentation into words or characters. In this context, a set of

new profile based features were proposed and proved their efficacy.

Several experiments were carried out on the IFN/ENIT database of Tunisian city names. The provided results showed significant improvement of the recognition rate coming from using optimized character shape models and profile based features. These results were compared with published works using the same database. The results reveal the robustness and the effectiveness of our system. In a future research, we intend to test the proposed system on Arabic word images with open vocabulary.

REFERENCES

[1] A. Mezghani, F.K. Jaiem, S. Kanoun and M. Kherallah, "Contribution on Character Modelling for Handwritten Arabic Text Recognition", International Afro-European Conference for Industrial Advancement, pp. 370–379, 2016.

[2] M. Pechwitz, S.S. Maddouri, V. Margner, N. Ellouze and H. Amiri, "IFN/ENIT - Database of Handwritten Arabic words", Colloque Internationale Francophone Sur l'Ecrit et le Document, pp. 129–136, 2002.

[3] R. Al-Hajj, L. Likforman-Sulem and C. Mokbel, "Combining Slanted-Frame Classifiers for Improved HMM-Based Arabic Handwriting Recognition", IEEE Transaction Pattern Analysis Machine Intelligence, vol. 31, no. 7, pp. 1165–1177, 2009.

[4] A. Benouareth, A. Ennaji and M. Sellami, "Semi-Continuous HMMs with Explicit State Duration for Unconstrained Arabic Word Modeling and Recognition", Pattern Recognition Letters, vol. 29, no. 12, pp. 1742–1752, 2008.

[5] J. AlKhateeb, J. Ren, J. Jiang, and H. Al-Muhtaseb, "Offline Handwritten Arabic Cursive Text Recognition using Hidden Markov Models and Re-Ranking," Pattern Recognition Letters, vol. 32, no. 8, pp. 1081–1088, 2011.

[6] M. Parvez and S. Mahmoud, "Arabic Handwriting Recognition using Structural and Syntactic Pattern Attributes", Pattern Recognition, vol. 46, pp. 141–154, 2013.

[7] A. Lawgali, M. Angelova, A. Bouridane, "A Framework for Arabic Handwritten Recognition Based on Segmentation", International Journal of Hybrid Information Technology, vol. 7, no. 5, pp. 413–428, 2014.

[8] A. Lawgali, M. Angelova, A. Bouridane, "HACDB: Handwritten Arabic Characters Database for Automatic Character Recognition", European Workshop on Visual Information Processing, pp. 255–259, 2013.

[9] A. Mezghani, S. Kanoun, S. Bouaziz, M. Khemakhem and H. El-Abed, "Baseline Estimation in Arabic Handwritten Text-line: Evaluation on AHTID/MW Database", International Conference on Pattern Recognition Applications and Methods, pp. 430–434, 2013.

[10] Y.C. Chim, A.A. Kassim and Y. Ibrahim, "Character recognition using statistical moments", Image and Vision Computing, vol. 17 (3/4), pp. 299–307, 1999.

[11] L.E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains", The Annals of Mathematical Statistics, vol. 41, no. 1, pp. 164–171, 1970.

[12] M. Pechwitz and V. Maergner, "HMM Based Approach for Handwritten Arabic Word Recognition Using The IFN/ENIT – Database", International Conference on Document Analysis and Recognition, pp.890–894, 2003.

[13] M. Khorsheed, "Recognizing Handwritten Arabic Manuscripts using a Single Hidden Markov Model", Pattern Recognition Letters, vol. 24, no. 14, pp. 2235–2242, 2003.

[14] L. Rabiner, and B.H. Juang, "Fundamentals of speech recognition", Prentice-Hall, Upper Saddle River, NJ, USA, 1993.

[15] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm", Royal Statistical Society Series B Methodological, vol. 39, no. 1, pp. 1–38, 1977.

[16] M. Hamdani, H. El-Abed, M. Kherallah and A. Alimi, "Combining Multiple HMMs using on-Line and Off-Line Features for Off-line Arabic Handwriting Recognition", International Conference on Document Analysis and Recognition, pp. 201–205, 2009.

[17] A. Elbaati, H. Boubaker, M. Kherallah, A. Alimi, A. Ennaji, and H. El-Abed, "Arabic Handwriting Recognition using Restored Stroke Chronology", International Conference on Document Analysis and Recognition, pp. 411–415, 2009.

[18] D. Xiang, H. Yan, X. Chen and Y. Cheng, "Offline Arabic Handwriting Recognition System Based on HMM", International Conference on Computer Science and Information Technology, vol. 1, pp. 526 –529, 2010.

[19] Y. Kessentini, T. Paquet and A. Ben Hamadou, "Off-line handwritten word recognition using multi-stream hidden Markov models", Pattern Recognition Letters, vol. 31, no. 1, pp. 60–70, 2010.

[20] A. Giménez, I. Khoury, J. Andrés-Ferrer, and A. Juan, "Handwriting Word Recognition using Windowed Bernoulli HMMs", Pattern Recognition letters, vol. 35, pp. 149–156, 2012.

[21] K. Jayech, M.A. Mahjoub, and N. Ben Amara, "Arabic Handwritten Word Recognition based on Dynamic Bayesian Network ", International Arab Journal of Information Technology, vol. 13, no. 3, pp. 276–283, 2016.