2017 International Conference on Smart, Monitored and
Controlled Cities (SM2C), Kerkennah, Tunisia, February, 17-19, 2017

Arabic Font Recognition System Applied to Different Text Entity Level Analysis

Faten Kallel Jaiem[1], Fouad Slimane[2], Monji Kherallah[3]
[1]ENIS, University of Sfax, Tunisia
[2]DHLAB, Ecole polytechnique Fédérale de Lausanne (EPFL), Switzerland
[3]Department of physics, Faculty of Sciences of Sfax,
University of Sfax, Tunisia
{kallelfaten,monji.kherallah}@gmail.com, fouad.slimane@epfl.ch

Abstract—**Arabic text recognition systems encounter many difficulties in processing multi-font and multi-size text images. Some Arabic font families introduce complex variability such as overlaps and ligatures. In this case, developing a font recognition phase for text recognition has become a necessity. In this paper, we presented a new version of the steerable pyramids font recognition system. We also compared and evaluated this system with two font recognition systems based on different methods. We adapted all the systems to a common measurement protocols, and evaluated them with text block, text line and word images. The results are very interesting in the different text entity levels and show the superiority of the steerable pyramid method over the text blocks image with the high resolution and comparable recognition rates with other text entities.**

Keywords—font recognition; steerbale pyramids; GMMs;O-LBP;APTID/MF; APTI.

I. INTRODUCTION

Our major focus in this study was on texts Arabic Font recognition. The Arabic script has a cursive nature for printed and handwritten text. The Arabic letters change forms according to their position in the word. Most of them have four shapes: isolated, initial, medial and final[1]. In fact, an Arabic recognition phase is a complex task. For this reason, most researchers have resorted to a font recognition phase, as a pre-recognition step to select the font dependent text recognition system, for printed Arabic documents[1].The analysis of the literature shows that, font recognition systems differ in the technique of feature extraction, the classifier and the used text entity level. Thus, font recognition systems can be stored in four classes based on the text entity level analysis: text block, text line, word and character.

For the first class, we cited many related, Zhuet al. [2] proposed a font recognition system using Gabor filter for the feature extraction phase. A set of 16 filters with 4 orientations and 4 frequencies has been used. To evaluated the proposed method 14,000 are used, the image data set was divided into128*128 uniform blocks of text. This database presents24 Chinese fonts and 32 English fonts. The recognition rate achieved is about99.1%. Ma et al.[3], also described a font recognition system based on Gabor filter. In [3], 24 filters have been used and the input images are transformed into 256*256 uniform blocks. This method was experimented with 3 scripts: Latin, Greek and Cyrillic, with 5 fonts. The best recognition rates are about 98.28% for the Latin script, 98.48% for the Greek script and 98.35% for Cyrillic script. Furthermore, researches use other texture analysis techniques than Gabor Filters such as Gray Level Co-occurrence Matrix (GLCM) and wavelet. In [4], Bataineh et al. used Edge Direction Matrix based on GLCM as a technique for the feature extraction phase to identify 7 Arabic fonts. The dataset contains 700 text-block image samples, which means of 100 samples from each font. The final images used for evaluation are in a 512*512 pixel image size. As for the use of wavelet, Imani et al. presented a technique for the feature extraction in [5]. Their experimental studies relied 10 Persian fonts, and their database contained 5000 images. For the feature extraction phase, they used uniform blocks with 128*128 pixel image size.

The second class consists of a font recognition system using text line. In [6], Zarmdini et al. proposed a statistical approach based on global typographical features. The used database includes a text line images written with 280 font models, 7 sizes and 4 styles. In [7], Ben Moussa et al. used Fractal multi-dimensions as a feature. They used 1000 text lines of an Arabic database including 10 fonts. Khosarvi et al. [8], also presented a Farsi font recognition system based on the text line level. In these works the authors presented a texture analysis technique for feature extraction, using Sobel and Robert filters. In [9], Luqman et al. presented a new Arabic Font database KAFD and a font recognition system. The database includes text blocks and text line datasets. The Arabic font recognition system uses the Log-Gabor Filters for feature extraction. Using 20 fonts, 10 sizes, and 4 styles of KAFD, the recognition rate obtained is 98.1%.

The third class includes font recognition systems based on the word image. Among the works proposed in the literature. Ben Amara [10] suggested a multi-resolution font recognition system based on sub-word database and wavelet, to identify 9Arabic fonts. The font recognition rate reached about 99%. Also Slimane et al. [1] proposed a font-size recognition system using Gaussian Mixture Models (GMMs) as a classifier and the sliding window technique to extract features. This system was evaluated using word images from the APTI database [11].

On the other hand, some researchers propose a character font identification system. Fu et al. in [12] and Ding et al. in [13] proposed a font identification system based on texture techniques using Chinese characters. Whereas Chaker et al. presented Arabic character font recognition [14]. The suggested recognition method is based on a shape index that allows the characterization of each character.

To our knowledge, the steerable pyramid which is applied for various applications of images is not used by the current works [15]. In [15] we suggested an Arabic Font Recognition System based on steerable pyramids (AFR/SP system), compared to other texture analysis technique. This paper presented a novel version of our proposed system. The novelties in this paper are:

To present the advantage of the new Font recognition system compared to the previous one.

To evaluate the performance and the behavior of the system with text block, text line and word images.

To compare our system with the leading Gaussian Mixture Model system (AFR/GMMs system) and the Oriented Local Binary Patterns System (AFR/O-LBP System).

The remaining of this paper was organized as follows: the second section described the font recognition system. The third section presented the experimental results and a comparison between the AFR/SP and the AFR/GMMs methods. In the final section, we concluded and suggested some perspectives for our future works.

II. SYSTEM DESCRIPTION

In this section, our new version of the AFR/PS system was presented, and the AFR/GMMs [1] and the AFR/O-LBP[16] systems were briefly described.

A. AFR/PSsystem

The AFR/PS system is an Arabic font recognition system based on steerable pyramids, using the Back propagation Artificial Neural Network (BpANN) as a classifier. At the feature extraction phase, we used steerable pyramids, which applied different steerable filters in multi-orientations to decompose the image in many sub-bands. To implement the steerable pyramids, a "k" band-pass filters as steerable filters in multi-orientations (FPBd) and a low-pass filter (FPBl) consists in generating many images of input text with a different resolution representing the recursive part of the algorithm. The structure of the steerable pyramid decomposition was presented by Figure1.

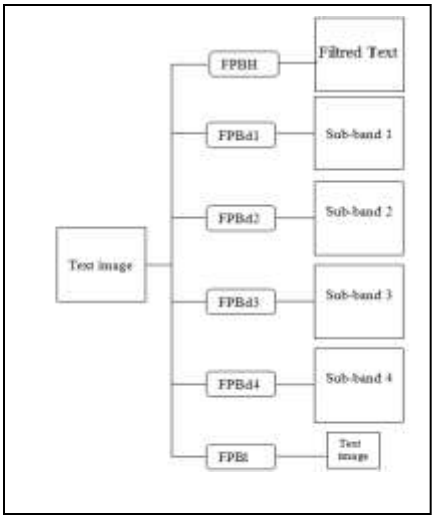


Fig. 1. The structure steerable pyramid decomposition

For the feature vector, for each level, we calculate statistical variables of the sub bands. For more details about feature extraction phases, we refer to [15].

In this section a new version of AFR/PS system is presented. For the recognition step, a tree-layer feed-forward neural networks classifier was applied, using a back-propagation method.

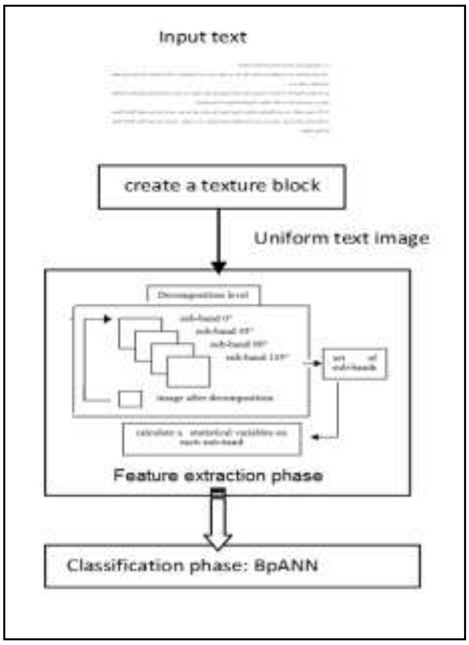


Fig. 2. The AFR/SP system architecture.

A deep understanding of in the current works, presented in our first section, reveals that many parameters are usually used for texture analysis techniques. To calculate the feature vector with a multi-orientation texture analysis technique two statistical variables have been used namely mean and standard deviations and 4 orientations(0°, 45°, 90° and 135°). The figure 2 presents the architecture of our system and shows the structure of steerable pyramid decomposition. For the input image, different sizes of the uniform image are used, for the texts blocks level, which are 128*128, 256*256 or 512*512. To be under the same condition, we chose to modify our system.

The new version presents the usually used parameters. The table below presents the difference between the two versions of AFR/SP system.

TABLE I. OLD AND NEW VERSIONS OF THE AFR/PS SYSTEM

AFR/PS system	Image size	Statistical variables used for feature extraction	Used orientations
Old version	2048*2048	variance	0°, 30°, 60°, 90°, 120° and 150°
New version	512*512	mean and standard deviations	0°, 45°, 90° and 135°

B. *AFR/GMMs system*

The AFR/GMMs system is based on Gaussian Mixture Models to estimate the likelihoods of font classes and the sliding window technique for local feature extraction.

The feature extraction method consist on transform the line image into a sequence of feature vectors based on an analysis window moved from right-to-left. The analysis window has a width of 8 pixels and a shift of 1 pixel. 110 typographical and statistical features and delta coefficients are extracted from each window. For the text blocks, we segment an image into text lines and we concatenate the feature vector sequences of all the extracted text lines of one text block into a single sequence.

In the training phase, the Expectation-Maximization (EM) algorithm is used to iteratively refine the component weights, means and variances to monotonically increase the likelihood of the training feature vectors. In our experiments we used the EM algorithm to build the models by applying a simple binary splitting procedure after each 10iterations to increase the number of Gaussian mixtures through the training procedure up to 1024 mixtures. We experimentally determined the number of Gaussians to have an optimal system without any prior class distribution knowledge in the decision process.

At recognition, the GMMs are fed in parallel with the features extracted from the image. The font recognition is performed through a simple class score comparison. The model with the highest score is selected and determines the font and size hypothesis. Performances are evaluated in terms of classification rates using an unseen set of word/text line or text block images. For more details about this system, we refer to [1].

C. *AFR/O-LBP system*

This system is based on Local Binary Patterns (LBP), which is a texture feature. The system is made up of two steps: the definition of LBP feature-set and classification using the Neural Networks (NN). In first step, 8 LBP operators are defined based on the radius and the number of neighborhood pixels. The LBP operator is a spatial structure analysis technique of the local texture in the text image. For each operator, an LBP histogram is defined. LBP features are usually derived from the occurring pattern histogram [16]. In the second step, the performance of AFR/O-LBP system was evaluated with a multi-class classifier. For the classification, the Neural Network classifier was used. For more details about this system, we refer to [16].

III. EVALUATION

Different text entity level analysis was used for the evaluation of font recognition systems: text blocks level, text line level and word level. In the first part of our experimental study, we compared our system with AFR/GMMs system. To evaluate the performance of our system with a low-resolution image and compare the results with low-resolution and high-resolution we used the APTI database. In the second part, we presented the performance evaluation of AFR/SP with different text entity levels.

In our study, we used two databases with 10 Arabic fonts (presented in Figure 3):APTID/MF database [17] and APTI database[11].In the APTID/MF database, all the text images contain incomplete text lines, different spaces between lines and words. The font recognition method based on texture analysis technique needs a normalized and uniform block of text. Firstly, for each input image we normalize the line space and word space as a first step. Then, to create a texture block, for each incomplete text line, we fill the blank space by a line block. The same process is applied for the text line images and word images.

Fig. 3. 10 Arabic fonts used in APTID/MF and APTID[17].

The first text image dataset was extracted from APTID/MF, which was created in high-resolution of 300dpi. All the text images have a big size.2048*2048 texture blocks was generated and divided into 16: 512*512 texture block consisting of different texts. For each font we save 334: 512*512 uniform text block images as a training set and 166: 512*512 uniform text block images as a test set. An example of texture blocks with the Andalus font is presented in Figure 4.

Fig. 4. A texture blocks forAndalus font: (a) a high resolution text image, (b) a low resolution text image, (c) a low resolution line image, (d) a low resolution word image.

For the second text image datasets, we used some part of APTI database, an ultra-low resolution(72 dpi) and synthetic images. For the text block images and line text images, we used the text image datasets proposed in [16]. The set of text blocks and text line were created with the first 1000 words from the set 1 for training and the first 1000 words from the set 5 for testing, the text block contains 250 words, 7 words for the text line. For each font and size we have 8 text images, 4 samples for the training and 4 samples for the test. The text line set contains 286 images for each font and size, divided into a training set and a test set. 800 uniform text block images and 28,600 uniform texture blocks for text line images with 512*512 size of pixels were used as an input image set. The third set of images, word images, is the same set used in [1], 100,000 words in set 1 for training and the first 100,000 words in set 5 for testing of APTI database. At this level we create a 256*256 uniform texture block from the word images.

A. AFR/SP vs. AFR/GMMs on high resolution images

Many researchers have compared their system with AFR/GMMs proposed by Slimane et al. [1] such as Luqman et al. [8] and Nicolaou et al. [16]. In tables II and III we displayed a comparison of our system and AFR/GMMs system applied to different text entity levels and different resolutions.

For the high resolution(300 dpi), we evaluated the systems with APTID/MF database. The examination of the fonts shows a strong similarity between 'Simplified Arabic' font and 'Arabic transparent' font. In the second part of this experiment we deleted the 'Simplified Arabic' font from the training and test steps.

TABLE II. COMPARISON OF AFR/SP AND AFR/GMMS WITH APTID/MF

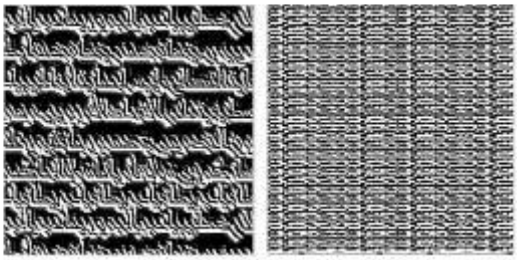
	10classes	9 classes
AFR/GMMs	99.58%	99.97%
AFR/SP	99.32%	99.87%

The analysis of the results shows a slight superiority of AFR/GMMs compared to AFR/SP. The difference between the recognition rates using APTID/MF is about 0.26% for 10 classes and 0.1% for 9 classes. We note that AFR/GMMs has a complex process that made it very low compared to AFR/SP. However, the simplicity of AFR/SP gives an advantage to our proposed method to be used for font recognition on text blocks with high resolution.

B. The AFR/SP vs. AFR/GMMs and AFR/O-LBP applied to different text entity level on low resolution.

In this section, we compared three systems relying on text block, text line and word levels with APTI database which is created in a low resolution of 72 dpi on 10 fonts.

The experimental results, presented in Table III., using the APTI database showed that the use of AFR/GMMs and AFR/O-LBP give high recognition rates compared to AFR/SP. This superiority is due to multi-resolution property of the steerable pyramid. This technique consists in the decomposing the image in level$_i$ with n*n size into 4 sub image n/2*n/2 pixels in level$_{i-1}$filtered with 4orientations (0°, 45°, 90° and 135°). Figure5 shows a sub-band in 45° of the text blocks images of Figure 4 after three decompositions. In Figure 5(a), which is high-resolution image (Figure 4(a)), the text lines and the characters are clear. However, the decomposition in three levels of the low-resolution image of Figure 4(b), gives a very ambiguous text presented in Figure 5(b). It is clear, then that the text image in low resolution after multi decompositions became confusing. This property cause a disadvantage for font recognition based on a multi-resolutions method such as steerable pyramids applied on low-resolution images.



(a) (b)

Fig. 5. A sub-band in 45° of text blocks after 3 decompositions: (a) a high-resolution text, (b) a low-resolution text.

No one denies that, there are many works based on texture analysis techniques for feature extraction. In this paper, however we used a steerable pyramid as a texture analysis technique. In this method, the text image is regarded as an image containing some specific textures. The font recognition is based on the identification of character shape combination in the text. In our font recognition scheme, we use steerable filters to characterize and analyze local orientation in writing Arabic patterns[15]. It is clear, that the number of words and characters in a text is a significant factor, which acts directly on the recognition rates. The results in TABLE III prove this idea.

TABLE III. RECOGNITION RATESWITHDIFFERENT TEXT ENTITY LEVEL ON 10 FONT CLASSES

	Text block level (250 words)	**Text line level (7 words)**	**Word level (1word)**
AFR/SP	93.25%	90.67%	78.98%
AFR/GMMs	*	95.85%	94.5%
AFR/O-LBP	99.75%	94.27%	*

Comparing the results, the AFR/GMMs system is noticed to perform better results compared to the two others. However, is still slower in its execution time compared to the others.

IV. CONCLUSION

In this paper, we presented a comparative study of a font recognition system applied on different image resolutions and text entities. The proposed system based on steerable pyramids was compared with AFR/GMMs and AFR/O-LBP, and evaluated on high and low resolution images. In the second part of this work, we introduced an analysis of texture technique performance for different text font recognition, text blocks, text lines and words. The experimental results showed that the use of the steerable pyramid yields high recognition rates of about 99% for high-resolution text block images.

In our future studies, we will test our proposed strategy for script identification and writer identification and we will develop an optimization method to determine automatically the best selected features.

In addition, as a perspective of our studies, we will evaluate the performance of our system for various applications for the image processing, such as, biometric and medical image analysis...

REFERENCES

[1] F. Slimane, S. Kanoun, J. Hennebert, A.M. Alimi and R. Ingold, “A study on font-family and font-size recognition applied to arabic word images at ultra-low resolution”, PRL, 34 (2) 2013, pp. 209–218.

[2] Y. Zhu, T. Tan, and Y. Wang. “Font Recognition Based on Global Texture Analysis” PAMI 2001, Vol.23, N0.10.

[3] H. Ma and D. Doermann, “Font Identification Using the Grating Cell Texture Operator”, DRR 2005, pp.148-156.

[4] B. Bataineh, S.N.H.S. Abdullah, and K. Omar, , “A novel statistical feature extraction method for textual images:Optical font recognition”, ESWA 2012, pp. 5470–5477.

[5] M. B. Imani, M. R. Keyvanpour and R. Azmi, “Semi-supervised Persian font recognition”, PCS2011, pp. 336–342.

[6] A. Zramdini and R. Ingold, “Optical font recognition using typographical features”, IEEE PAMI 1998, pp. 877–882.

[7] S. Ben Moussa, A. Zahour, A. Ben abdelhafid, A. M. Alimi, “New feature susing fractal multi-dimensions for generalized arabic font recognition”, PRL,2010, pp. 361–371

[8] H. Khosaravi and E. Kabir, “Farsi font recognition based on Sobel-Roberts features”, PRL,2010, pp. 75–82.

[9] H. Luqman, S.A. Mahmoud and S. Awaida “ KAFD Arabic Font Database”, PR, 2014, pp. 2231–2240.

[10] N.E. Ben Amara and S.Gazzah, “Une approche d’identification des fontes arabes”, CIFED 2004.

[11] F. Slimane, R. Ingold, S. Kanoun, A.M. Alimi and J.Hennebert, “A New Arabic Printed Text Image Database and Evaluation Protocols”, ICDAR 2009, pp. 946–950.

[12] J.Fu, M. Guo and X. Tan, “Font recognition of chinese character based on multi-scale wavelet”, Internat. Conf. on Circuits, Systems, Electronics, Control & Signal Processing, 2006, pp. 90–93.

[13] X. Ding, L. Chen and T. Wu, “Character independent font recognition on a single chinese character”, IEEE PAMI 200, pp. 195–204.

[14] I.Chaker, M.Harti, H. Qjidah and R. Benslimane, “Recognition of Arabic characters and fonts”, IJEST, 2010, pp. 5959–5960.

[15] F.K. Jaiem, S. Kanoun and V. Eglin, “Arabic font recognition based on a texture analysis”, ICFHR, 2014, pp. 673–677.

[16] A. Nicolaou, F. Slimane, V. Margner and M. Lwicki, “Local binary patterns for Arabic optical font recognition”, DAS 2014, pp. 76-80.

[17] F.K. Jaiem, S. Kanoun, M. Khemakhem, H. El Abed, and J. Kardoun, “Database for Arabic Printed Text Recognition Research,” ICIAP 2013, Part I, LNCS 8156, pp. 251–259.