A novel Arabic handwriting recognition system based on image matching technique

1st Maamar Kef
Department of Computer Sciences
Universit Mostefa Benboulaid - Batna 2
Batna, Algeria
lm_kef@yahoo.fr

2nd Leila Chergui
Department of Computer Sciences
Universit Mostefa Benboulaid - Batna 2
Batna, Algeria
pgleila@yahoo.fr

***Abstract*—This paper presents a new off-line recognition system for Arabic handwritten words. The proposed system uses scale-invariant descriptor namely SIFT, and based on an image matching technique for achieving classification. The recognition process was done through a Keypoints matching procedure, using a nearest-neighbor distance-ratio. The paper presents also a new large Arabic handwritten word database. This database provides a new framework for benchmarking and gives a new freely available Arabic handwritten word dataset. Several tests have been performed using our new database and the well known IFN/ENIT database for comparison purposes. A high correct recognition rate was reported.**

***Index Terms*—Arabic handwriting recognition, Features extraction, SIFT descriptors, Keypoints matching, New Arabic database.**

Automatic recognition of handwritten scripts is an area of pattern recognition that is extremely useful in numerous fields, including documentation analysis, mailing address interpretation, bank check processing and more recently the reconstruction and recognition of historical manuscripts.

Recognition of Arabic handwriting remains one of the most challenging problems in the pattern recognition domain. Arabic is written by more than 240 million people, in over 20 different countries. The standard Arabic script contains 28 letters. Each letter has either two or four different shapes, depending on it position within a word.

One of the most challenging aspects of off-line handwriting recognition is finding a good database that well represents the variety of handwriting styles. Comparing with the great number of existing databases for English script, IFN/ENIT database [1] was the only freely accessible Arabic database; this incited us to develop a new large database which will be freely available for research and academic use.

In this research we present a new fast and robust Arabic handwriting recognition system based on SIFT descriptor and a recognizing procedure that use keypoints matching. Contrary to the majority of handwritten characters recognition systems, the proposed method operates without any preprocessing steps, since the used features are invariant regarding images' transformations and are highly distinctive in a large database. We also introduce a new large database of Arabic handwritten words which provides a comparison tool for research works in characters recognition domain.

The remainder of this paper is divided into six sections. The next section resumes several works done in handwritten Arabic recognition field. Section 3 detail the feature extraction method and section 4 describes our new Arabic handwritten words database. Experimental results including keypoints detection and matching are reported in section 5, where a comparative analysis of the experimental results is also discussed. Finally, some concluding remarks end the paper.

I. Related works

The main idea of scale invariant feature descriptor (SIFT) [6] is resumed on detecting distinctive invariant features from images that can be later used to perform reliable matching between different views of an object or scene. Because of the proved efficiency of the SIFT keypoint detector, a large number of researcher are attracted further for expanding or using these descriptors in many applications. In handwritten recognition domain, SIFT was addressed in a few published papers.

Diem and Sablatnig [5] tried to solve the problem of degraded handwritten characters recognition using SIFT descriptors. In order to recognize a character, the local descriptors are initially classified with a Support Vector Machine (SVM) and then identified by a voting scheme of neighboring local descriptors.

De Campos [4] presented a solution to the problem of recognizing characters in images of natural scenes. Such situations could not be well handled by traditional OCR (Optical Character Recognition) techniques. The problem is addressed in an object categorization framework based on a bag-of-visual-words representation. For feature extraction, authors used SIFT and other descriptors.

Zhang et al. [13] proposed a novel SIFT based feature for off-line handwritten Chinese character recognition. The presented feature is a modification of SIFT descriptor taking into account of the characteristics of handwritten Chinese samples. MQDF classifier was used in classification phase and showed that the proposed method outperforms original SIFT feature and two traditional features, Gabor feature and gradient feature.

In [11] a new method for the off-line recognition of Tamil handwriting characters based on local feature extraction was investigated. Authors represented each character by a set of local SIFT feature vectors.

Character type classification on a document image problem was addressed in [12]. In that work, authors proposed a method based on a probabilistic topic model and SIFT descriptor. The character' types are: mathematical formula, printed Japanese, printed and handwritten English.

Ramana et al. [9] examined the issues in recognizing the Devanagari characters in the wild like sign boards, advertisements, logos, shop names, notices, and address posts. They used a variation of SIFT, namely Dense SIFT features. These are derived by densely sampling keypoints from the character and extracting SIFT descriptors around them.

Mao et al. [8] incorporated SIFT descriptors in Chinese calligraphy word style recognition domain (seal script, clerical script, standard script, semi-cursive script and cursive script). In this study, authors proposed a method based on K-Nearest Neighbors (KNN) and feature vector filtering. Experiments show that SIFT feature has better recognition result than that of Gabor feature and GIST feature.

For Arabic handwriting recognition, we found only one work which uses SIFT as descriptor introduced by Rothacker et al. [10]. They applied the Harris detector to extract coins and for each coin, they detect keypoints using SIFT descriptors; they also used a segmentation phase with a set of Hidden Markov models.

Aouadi and Kacem Echi [14] presented a new method for Arabic handwritten word recognition. The authors extracted some structural features from words image and trained a classic right-left Hidden Markov Model. Experiments were carried on a set of ancient Arabic manuscripts and the IFN-ENIT standard database. An average recognition rate of 87% was reported.

Rabi et al. [15] presents a recognition system of Arabic cursive handwriting using embedded training based on hidden Markov models. The extracted features were based on the densities of foreground pixels, concavity and derivative features using sliding window, some of these features depends on baselines estimation. the system achieved 87.93% of correct recognition.

II. SIFT DESCRIPTOR

SIFT was developed by David Lowe in 2004 [7] as a continuation of his previous work on invariant feature detection [6], and it presents a method for detecting distinctive invariant features from images that can be later used to perform reliable matching between different views of an object or scene. This approach consists of four major computational stages (figure 1).

Each of these stages are executed in a descending order (cascade approach) and on every stage a filtering process is applied so that only the keypoints that are robust enough are allowed to pass to the next stage. According to Lowe, this will reduce significantly the cost of detecting the features. The descriptor is formed from a vector containing the values of all the orientation histogram entries.

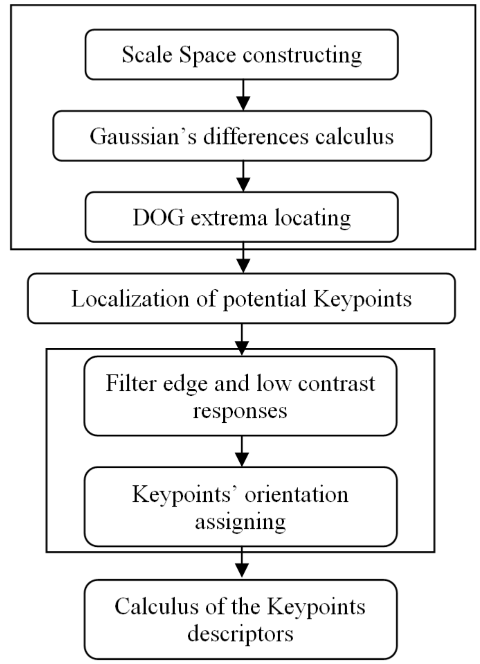

Fig. 1. SIFT features detection algorithm.

For image matching and recognition, SIFT features are first extracted from a set of learning images and stored in a database. A new image is matched by individually comparing features extracted from it to those previously stocked in the database and finding candidate matching based on Euclidean distances calculated from their feature vectors. The Euclidean distance between the SIFT feature descriptors is considered as a cost measure.

The experiments conducted in this paper use a 4x4x8 = 128 elements in each feature vector of a keypoint. Regarding the image matching procedure, the local descriptors from several images are matched. A complete comparison is performed by computing the Euclidean distance between all potential matching pairs. A nearest-neighbor distance-ratio matching criterion is then used to reduce mismatches.

III. THE NEW ARABIC HANDWRITING WORD DATABASE

In order to make the databases as much representative as possible, we have focused on most aspects responsible of variations of handwriting styles like the age, the sex, the educational level, the profession, the residence town, etc.

Data collection was conducted using 2100 forms. Each writer was asked to fill one form comprising 11 Algerian village names, each word is written twice. Also, there is a field for writer's personal informations including; his name, his age, his residence town, and his profession. Each form possesses 15 exemplars. An example of a filled form represented in a grayscale level is shown in figure 2.

All the extracted images have been archived in two different formats: grayscale and binary formats in TIFF file format at 300 dpi resolutions. The Arabic handwritten data were sorted and saved into four sets. Figure 3 shows some statistics

Fig. 2. Example of a filled form.

concerning the number of words, sub-words and characters in each set.

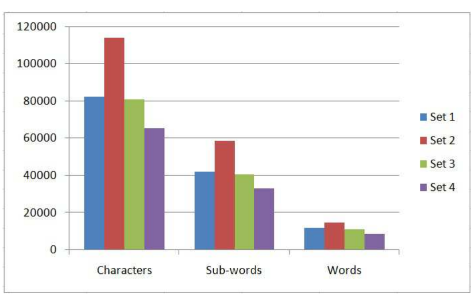

Fig. 3. Character' number, sub-words' number and words' number of our new database.

IV. A NOVEL RECOGNITION SYSTEM

In order to show the efficiency of the proposed system, experimental tests were achieved on both databases; the IFN/ENIT and our new database. IFN/ENIT was produced by the Institute for Communications Technology at Technical University of Braunschweig (Institut für Nachrichtentechnik, IFN) and the l'Ecole Nationale d'Inégnieurs de Tunis. This database was used as a comparison tool to evaluate researchers' works during the three competitions of the ICDAR (International Conference on Document Analysis and Recognition) organized in 2005, 2007 and 2009 [1].

A. Keypoints detection

In our study, we are not interested by the matching of two distinct images representing the same scene (or parts of the same scene) taken from two different views; our aim is to compare two images of two handwritten words whose similar contents will be in the same area, for all images representing a given word class.

The suggested method divides vertically the word images to be recognized into five frames of equal size. The objective here is to compare the detected keypoints in a given frame with its corresponding in another image representing the same word class. Figure 4 shows an example.

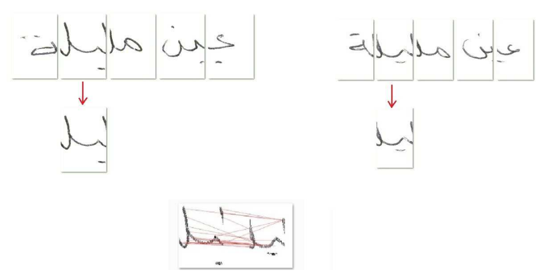

Fig. 4. Keypoints matching of two corresponding frames in two images.

The number of frames was selected through different tests of several scenarios and their impact on the recorded recognition rate (table 1).

For each word class, we build a model of keypoints using 25 images as training samples. Each class model contains a given number of keypoints divided into five subsets, representing the different frames composing the word images. The construction process of each class model is detailed in the flow chart presented in figure 5. This process allows us to filter and improve the robustness of keypoints extracted from the training images of a given class.

The number of training images used to build each class model was also fixed through several experiments. We noticed that using more than 25 images during the learning process will increase the number of detected keypoints without bringing a significant improvement to the recognition rate (figure 6).

A set of 128 features are extracted for each keypoint, since a keypoint descriptor consists of eight 4x4 orientation histograms. Figure 7 presents the keypoints detection process using SIFT descriptors for the five frames representing a word image taken from our database.

Several tests were conducted in order to determine the matching ratio; this parameter fixes the matched keypoints' number which affects the recognition rate. Tests show that the number of keypoints and the matching ratio are rising at the same time (figure 8), but the discriminating capacity of these

TABLE I
IMPACT OF THE FRAMES' NUMBER ON THE RECOGNITION RATE

Frames' number	1	2	3	4	5	6	7	8
Recognition rate (%)	57.94	63.38	76.77	87.61	**93,72**	90.61	86.88	81.16

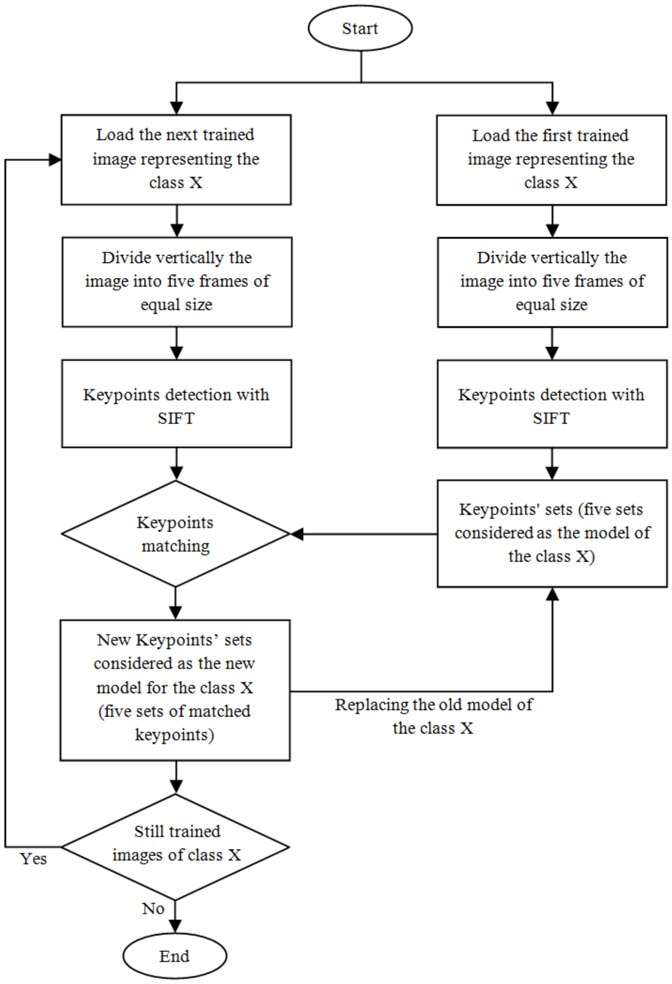

Fig. 5. Construction process of classes' models.

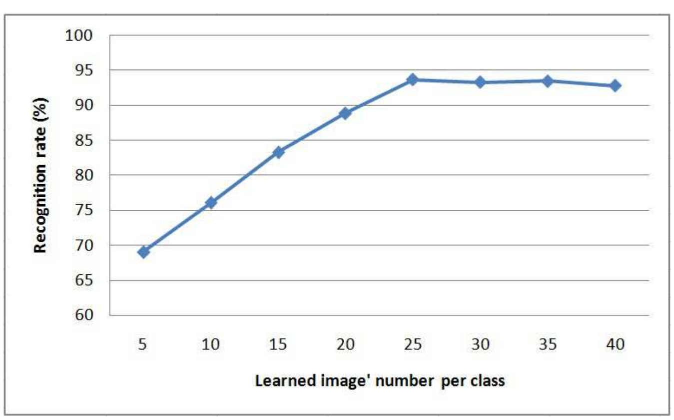

Fig. 6. Effect of the training images' number per class on the recognition rate.

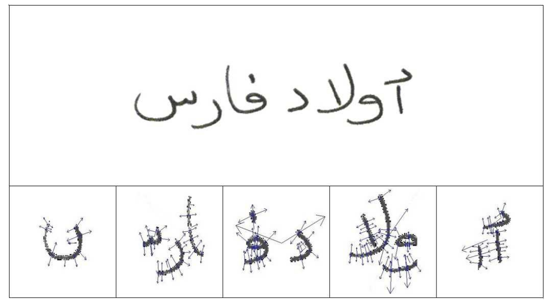

Fig. 7. Keypoints detection using SIFT descriptors for a handwritten Arabic word.

keypoints decreased. Figure 9 shows that Keypoints matching becomes more efficient when the matching ratio is fixed to 0.9 even if the number of keypoints is reduced. Worse still, the recognition rate tends to decrease when the ratio gets higher values.

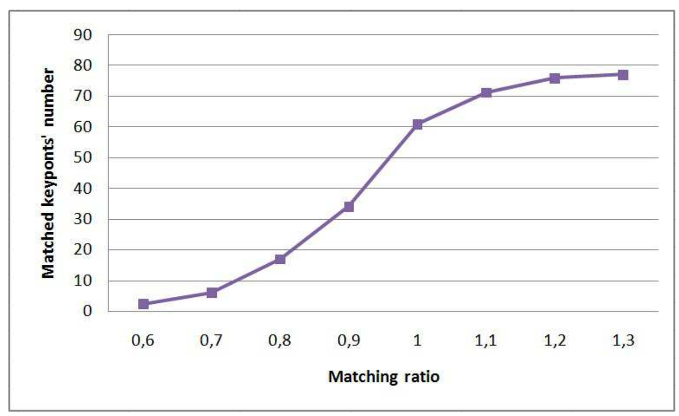

Fig. 8. Effect of the matching ratio on the matched keypoints' number.

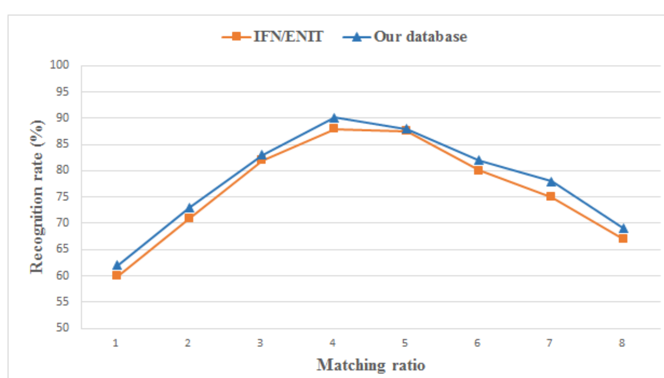

Fig. 9. Impact of the matching ratio on the recognition rate.

The number of keypoints representing each model of the 200 used classes, with which the system registered the highest recognition rate, is given in figure 10.

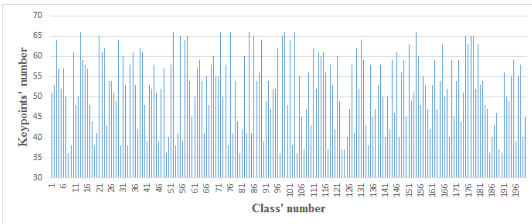

Fig. 10. Keypoints' number in each class's model.

B. Keypoints matching

Once the keypoints were detected in two images, they should be paired. The best candidate match for each keypoint in the first image is found by identifying its nearest neighbor in the second one. In this work, matching keypoints are calculated from features vectors by comparing the Euclidean distance of the closest neighbor to that of the second closest neighbor. Keypoints matching of the five frames representing an image pair is illustrated in figure 11.

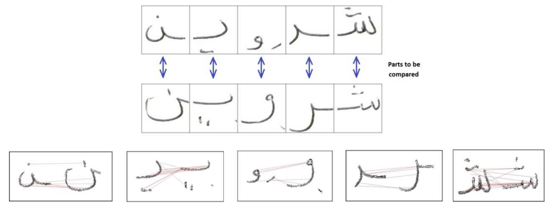

Fig. 11. Keypoints matching of an image pair.

In the recognition process, each image of test must be firstly divided into five frames, then the keypoints are calculated for each frame. The matching process is then performed as follows:

Repeat the following steps for each class model and each test image:

1) Each frame representing a part of a test image is compared with its correspondent part of a class model.
2) The matched keypoints rate (MKR) is then calculated for each frame as follows:

$$MKR = \frac{\text{matched keypoints' number}}{\text{detected keypoints' number from a test image + model keypoints' number}} \quad (1)$$

3) An average matching rate (AMR) is then established:

$$AMR = \frac{MKR(frame1) + MKR(frame2) + MKR(frame3) + MKR(frame4) + MKR(frame5)}{5} \quad (2)$$

Finally, the model recording the highest average matching rate will be considered as the target class. Figure 12 shows an example summarizing these stages.

The keypoint descriptors are highly distinctive, which allows a single feature to find its correct match with good probability in a large database of features.

Tests conducted on both databases (IFN/ENIT and our new database) are listed in table 2, where we can observe that the system registered high performances with scalability, since a slight loss of approximatively 8% of the accuracy was registered when the number of classes that have to be recognized increased from 40 to 200. We also noticed that a small improvement of the recognition rate was reported during tests done on our new database compared to the IFN/ENIT database.

TABLE II
REGISTERED PERFORMANCES USING IFN/ENIT AND OUR NEW DATABASES

Classes number	Recognition rate (%)	
	IFN/ENIT database	*Our database*
40	97.33	98.83
60	96.77	98.11
80	94.58	96.41
100	93.46	95.13
120	90.61	93.72
160	88.90	91.74
200	88	90.10

C. Results comparison

In order to prove the efficiency of the proposed method, we compare the obtained results with some pertinent works done on handwritten Arabic words recognition. However, only the systems tested on IFN/ENIT database have been mentioned. The reported results (table 3) show that our proposed system outperforms the other systems which proves the effectiveness of our approach.

V. CONCLUSION

The contribution of this paper is twofold. Firstly, a new large and free database for Arabic handwriting words is presented. Secondly, an effective and robust off-line handwritten Arabic words recognition system is presented and evaluated on this new database.

The developed sytem use a new type of features, namely SIFT descriptors and an efficient recognition method based on

TABLE III
COMPARISON RESULTS

Systems	Used classifier	Features extraction method	Recognition rate (%) (IFN/ENIT database)
Azizi [2]	MLP	Structural features	87.12
		Statistical features	87.46
		Selected features	87.05
Burrow [3]	KNN	Zernike moments	80
Aouadi and Kacem Echi [14]	HMM	Structural features	87
Rabi et al. [15]	HMM	densities of foreground pixels concavity and derivative features	87.93
Our system	Matching based on Euclidean distance	SIFT descriptor	88

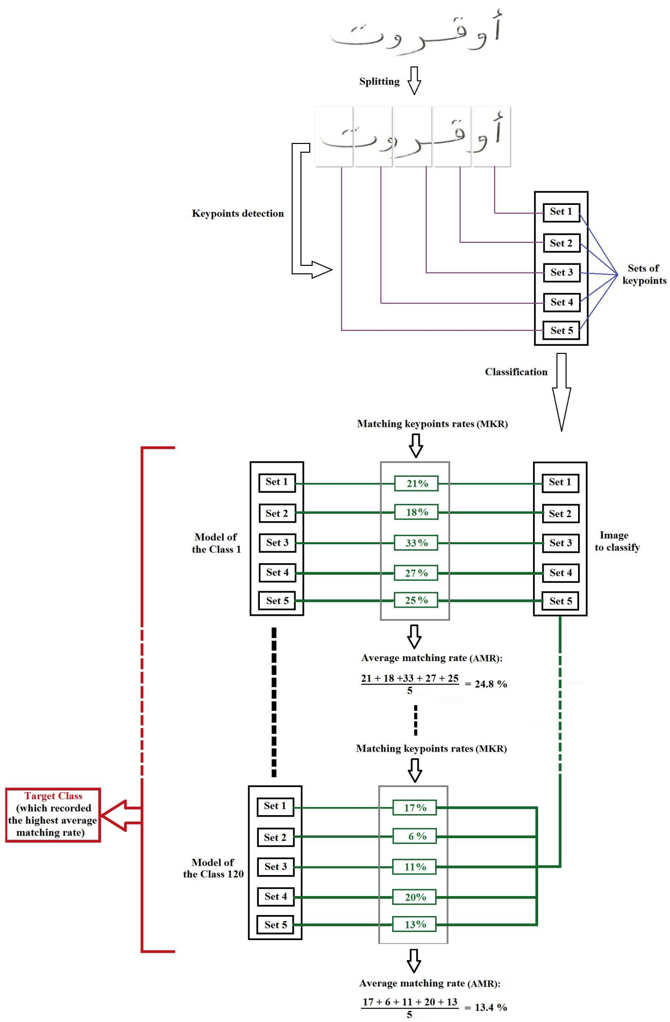

Fig. 12. Classification procedure.

an image matching procedure. A heigh recognition rate was recorded through several experiments conducted on IFN/ENIT and our new database.

REFERENCES

[1] H. Al Abed, V. Margner, ”ICDAR 2009 - Arabic handwriting recognition competition,” International Journal on Document Analysis and Recognition, Springer, vol. 14, no. 1, pp.3–13, 2011.

[2] N. Azizi, N. Farah, M. T. Khadir, M. Sellami, ”Arabic handwritten word recognition using classifiers selection and feature extraction/selection,” Proc. The 17th IEEE Conference in Intelligent Information System, Proceedings of Recent Advances in Intelligent Information Systems, Academic Publishing House, Warsaw, pp.735–742, 2009.

[3] P. Burrow, ”Arabic handwriting recognition,” Thesis, School of Informatics, University of Edinburgh, 2004, England.

[4] T.E. De Campos, B.R. Babu, M. Varma, ”Character recognition in natural images,” Proc. The International Conference on Computer Vision Theory and Applications, Lisbon, Portugal, vol. 2, pp.273–280, 2009.

[5] M. Diem, R. Sablatnig, ”Recognition of degraded handwritten characters using local features,” Proc. The 10th International Conference on Document Analysis and Recognition, Barcelona, Spain, pp.221–225, 2009.

[6] D. G. Lowe, ”Object recognition from local scale-invariant features,” Proc. of the International Conference on Computer Vision, Corfu, Greece, pp.1150–1157, 1999.

[7] D. G. Lowe, ”Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp.91–110, 2004.

[8] T. Mao, J. Wu, P. Gao, Y. Xia, Y. Lin, ”Calligraphy word style recognition by KNN based feature library filtering,” Proc. The 3rd International Conference on Multimedia Technology, uangzhou, China, pp.934–941, 2013.

[9] O. V. Ramana, S. Roy, V. Narang, M. Hanmandlu, ”Devanagari character recognition in the wild,” International Journal of Computer Applications, vol. 38, no. 4, pp.38–45, 2012.

[10] L. Rothacker, S. Vajda, G. A. Fink, ”Bag-of-features representations for offline handwriting recognition applied to Arabic script,” Proc. The 3rd International Conference on Frontiers in Handwriting Recognition, Bari, Italy, pp.149–154, 2012.

[11] A. N. Subashini, D. Kodikara, ”Novel SIFT-based codebook generation for handwritten tamil character recognition,” Proc. The 6th International Conference on Industrial and Information Systems, Sri Lanka, pp.261–264, 2011.

[12] T. Yamaguchi, M. Maruyama, ”Character type classification via probabilistic topic model,” International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 5, no. 2, pp.123–140, 2012.

[13] Z. Zhang, L. Jin, K. Ding, X. Gao, ”A novel feature for offline handwritten Chinese character recognition,” Proc. The 6th International Conference on Industrial and Information Systems, Sri Lanka, pp.763–767, 2009.

[14] N. Aouadi and A. Kacem Echi, ”Word Extraction and Recognition in Arabic Handwritten Text,” International Journal of Computing & Information Sciences, vol. 12, no. 1, pp.17–23, 2016.

[15] M. Rabi, M. Amrouch, Z. Mahani, ”Recognition of cursive Arabic handwritten text using embeddedtraining based on HMMs,” Journal of Electrical Systems and Information Technology, 2017 (article in press).