Feature Extraction Techniques of Online Handwriting Arabic Text Recognition

Mustafa Ali Abuzaraida[1], Akram M. Zeki[2]
Kulliyyah of Information and Communication Technology
International Islamic University Malaysia
Kuala Lumpur, Malaysia
[1]abuzaraida@yahoo.com, [2]akramzeki@iium.edu.my

Ahmed M. Zeki
Department of Information Systems
College of Information Technology
University of Bahrain
Manama, Kingdom of Bahrain
amzeki@uob.edu.bh

***Abstract*— Online recognition of Arabic handwritten text has been an ongoing research problem for many years. Generally, online text recognition field has been gaining more interest lately due to the increasing popularity of hand-held computers, digital notebooks and advanced cellular phones. Most of the online text recognition systems consist of three main phases which are preprocessing, feature extraction, and recognition phase. This paper compares between different techniques that have been used to extract the features of Arabic handwriting scripts in online recognition systems. Those techniques attempt to extract the feature vector of Arabic handwritten words, characters, numbers or strokes. This vector then will be fed into the recognition engine to recognize the pattern using the feature vector. The structure and strategy of those reviewed techniques are explained in this article. The strengths and weaknesses of using these techniques will also be discussed.**

Keywords—Text Recognition; Features Extraction; Online Arabic Recognition System.

I. INTRODUCTION

Word and characters recognition methods have been improved since many years. These methods used for printed or handwritten scripts and used two different approaches of processing which are online and offline. It has been gaining more interest lately due to the increasing popularity of hand-held computers, digital notebooks, and advanced cellular phones. These devices nowadays are commonly used worldwide that encouraged companies to improve their products to support multi languages. However, these devices can deal with many languages spoken by billions of people around the world such as Latin, Chinese, Japanese, Indian, Korean, Arabic, and many others from textual or speech manner.

Arabic language is the main language of all Arabic countries with more than 280 million people are speaking this language as a first language and by 250 million as a second language. Arabic language comes as the fifth rank of most commonly used languages in the world. However, there are some other languages related to Arabic language. These languages have some similarities with Arabic language whence from the characters shapes or from the pronunciation. These languages are Jawi, Persian, Urdu, Pashto, Bengali, and others. These languages are spoken by millions of people in many Islamic countries such as Iran, Afghanistan, Pakistan, parts of India, Bangladesh, Sri Lanka, Malaysia, Indonesia, and other countries [1].

In text recognition field, it is noticeable from the literature that most of the research works were dedicated to Latin characters and other languages such Chinese. A few researches and studies have been conducted to develop new methods and algorithms in this area for Arabic language. Besides that, online Arabic handwriting recognition researches are still too few in number compared to those done in offline Arabic recognition.

The typical structure of online text recognition systems basically consists of three main phases [2] which are preprocessing, feature extraction, and recognition phase as shown in Figure 1. However, some researches provide extra phases or ignored some phases of the systems [3]. In this paper, a brief overview of the features extraction techniques proposed in the past works in the area of online Arabic handwriting recognition will be described and compared.

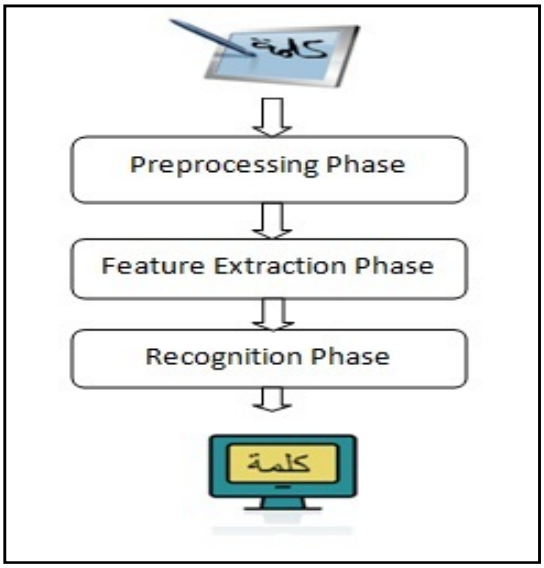

Fig. 1. Typical phases of online Text Recognition System

II. ARABIC LANGUAGE ALPHABET CHARACTERISTICS

Several reasons make Arabic language different from other languages from the shape and the writing style [4]. Here are some of these reasons:

- First, Arabic letters come in different shapes depending on their position in the word. These positions are beginning, middle, terminate, and single shape.

978-1-4799-0136-4/13/$31.00 ©2013 IEEE

Although, some letters come in two shapes which are single and terminate shape only as shown in Table I.

- Second, several Arabic characters have the same main body and it can be distinguished from each other by special extra strokes that may be dots or Hamza as shown in Table I.

TABLE I. ARABIC CHARACTERS CHAPES

Letter	Single	Beginning	Middle	Terminate
Alif	أ	أ	ـا	ـا
Baa	ب	بـ	ـبـ	ـب
Taa	ت	تـ	ـتـ	ـت
Thaa	ث	ثـ	ـثـ	ـث
Jeem	ج	جـ	ـجـ	ـج
Haa	ح	حـ	ـحـ	ـح
Khaa	خ	خـ	ـخـ	ـخ
Dal	د	د	ـد	ـد
Dhal	ذ	ذ	ـذ	ـذ
Raa	ر	ر	ـر	ـر
Zai	ز	ز	ـز	ـز
Seen	س	سـ	ـسـ	ـس
Sheen	ش	شـ	ـشـ	ـش
Sad	ص	صـ	ـصـ	ـص
Dad	ض	ضـ	ـضـ	ـض
Taa	ط	طـ	ـطـ	ـط
Dhad	ظ	ظـ	ـظـ	ـظ
Ain	ع	عـ	ـعـ	ـع
Ghain	غ	غـ	ـغـ	ـغ
Faa	ف	فـ	ـفـ	ـف
Qaf	ق	قـ	ـقـ	ـق
Kaf	ك	كـ	ـكـ	ـك
Lam	ل	لـ	ـلـ	ـل
Meem	م	مـ	ـمـ	ـم
Noon	ن	نـ	ـنـ	ـن
Haa	ه	هـ	ـهـ	ـه
Waw	و	و	ـو	ـو
Yaa	ي	يـ	ـيـ	ـي

- Third, Arabic words must be written in a cursive way and the letters must be connected unless one of the characters (ا, د, ذ, و, ر, ز) exist in the middle of the word to make sub-words as shown in Figure 2.

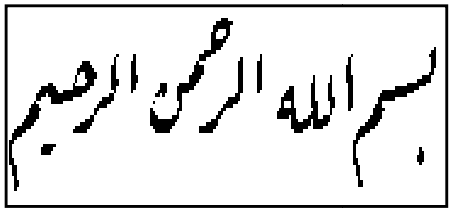

Fig. 2. The connectivity of the Arabic Characters

- Fourth, in such a writing way some Arabic words may have overlap case especially in Arabic handwritten. Here, some characters come in vertical position above their next character. Overlapping may cause an ambiguity to the system to recognize the proper word as illustrated in Figure 3.

Fig. 3. Arabic Words Overlapping

- Fifth, Arabic words can be written using different writing styles that make the letters and words having different outlook which make ambiguity to any recognition system. Figure 4 shows different styles of writing the word *Valley* (وادي).

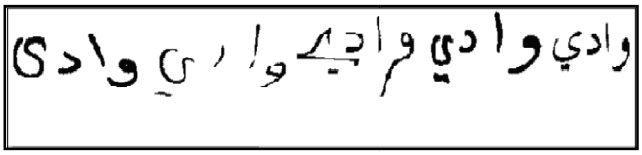

Fig. 4. Different writing styles of the word Valley (وادي)

III. FEATURE EXTRACTION METHODS

Text features are measurements applied to the word or a character and listed together to produce a measurement vector "feature vector". Any character, word, digit, or stroke has its own specific feature vector to identify it from any other characters, words or strokes [5]. For pattern recognition in general and text recognition approaches in particular, extracting an appropriate set of features and an efficient extraction method have been stated as the most important factors in achieving high recognition performance [6].

In online recognition systems, the information about how the character has been written is available. Although, complex preprocessing steps cannot be performed in practical online systems such as tablets and (personal digital assistant) PDAs since data is collected as the text is being written. Hence, taking advantage of the dynamic characteristics of the data is crucial which consists of speed, angular velocity, and other features of this kind. These features would be available for processing as the character is being written on the tablet [7].

Choosing the proper type of features depends on the nature of the text, the type of the system processing which may be online or offline, and the scripts types that can be handwritten or printed. However, feature types of recognizing any text can be categorized into three main types which are structural features, statistical features, and global transformation [7]. The description of each type is explained as follows.

Structural Features: This type of features illustrates and describes the geometrical and topological characteristics of a text which may be word, character, digit, or stroke by describing its global and local properties. Going deeply, the structural features depend on the nature of text that wants to be classified. For Arabic text, for example, these features include such as zigzags, dots, loops, cross points, branch points and strokes in different directions.

Also, the dots number and their positions, the length of contour segment, and the distance between the start and end points of the contour projection of the *x–axis* and *y–axis* can be considers as structural features. In general, structural features are difficult to extract from the Arabic text image

and many errors occur because of the similarity of Arabic characters shapes. However, extracting these features might be easier and more effective for online system due to the way of recording the writing data. For that, this type of features is commonly used in online recognition systems which have been noticed in the literatures.

Statistical Features: In this type of features, a set of characteristic measurements used to describe the text. For Arabic text recognition, the statistical features used for character representation include such as zoning, characteristic loci, projections and profiles, and crossing and distances. An example of these features is zoning of pixels which used by dividing the text image into zones and using the density of pixels in those zones as a feature. This type of features is used in offline systems and used in few online systems by combining them with structural type.

Global Transformations: This type is used in image processing in general and hence for offline recognition systems. Here, the Fourier transform of the contour of the image is calculated in order to reconstruct the image contour by calculating *n* coefficients of the Fourier transform, and then these *n* coefficients are considered to be a *n*-dimensional feature vector that represents the pattern.

Another classification of features types can be identified according to [8] , in which text features can be classified into two main categories which are quantitative and qualitative features. Quantitative features of any text include the number of dots, text measurements, text area, and text weight above and below the baseline. On the other hand, qualitative features include text loops, branches, topological descriptions, topological relations, dot position, connection points, and junction points. The combination of features can give more general description of the character, word, digit, or stroke. The resulting features can be either directly measured on the input character or obtained as a combination of the local features [9].

IV. Past works on online Arabic handwriting

Early research on online Arabic handwriting recognition was done by Almuallim and Yamaguchi [10]. They proposed a structural recognition method for cursive Arabic handwritten words. In this method, words were segmented into strokes and each stroke was defined as a continuous curve represented by a string of coordinates (X_i, Y_i), where the order of these coordinates represents an approximation of the pen movement during writing. The first and last points of the stroke are called the start and end points respectively. The extraction of a stroke is made by finding out its start point and then following the curve until a point which is inferred to be the stroke end point is reached. Experimental results on handwritten texts by two people showed high recognition accuracy with a rate of 81.25%.

A few years later, an online Arabic handwriting recognition system [11] based on decision-tree techniques was proposed to recognize Arabic handwritten words. For extraction of features phase, each character is specified by a number of direction codes from 1 to 4 and separation codes (0). Every character has its own collection of codes named the code number. A decision tree then used to store the information on the characters in the learning process, and for the search to find a character in the recognition process. The nodes of the tree each have five branches, corresponding to the above five numbers from 0 to 4 in the code number. The results were promising based on the early date of this research.

In [9] a new hybrid approach for recognition of on-line handwritten Arabic characters was described. The algorithm was based on features extracted by structural techniques and modeled by fuzzy sets. With the use of fuzzy linguistic features, it was shown that vagueness and uncertainties originated from the various writing styles and variations of characters can be broadly represented by fuzzy sets. The combination of certain fuzzy features was a natural way of describing characters. Simple fuzzy if-then rules were used to classify characters.

It is noticeable that the number of research works has been increased starting from the year of 2005. In [12] an efficient structural approach for recognizing on-line Arabic handwritten digits was proposed based on the changing signs of the slope values those represented by grammar string. In this system, features extraction phase has two steps; the first step is for extracting primary primitives of the digits using the primary break points; while the second step is for secondary primitives when there are no primary break points. The proposed system has been tested on an on-line dataset representing the digits 0-9 collected from 100 users. The average of the recognition accuracy was about 95%.

In the same year [13], an Arabic handwritten recognition system based on matching template algorithm and genetic algorithm was developed. In the features extraction phase, they used a heuristic method by using 13 visual codes to represent the handwritten trajectory that was divided into two zones which are median and striking zone and each zone is composed of its own visual. The system recognized correctly 92% for testers how have used system before and 70% for new users.

In the next year [14], a system based on Hidden Markov Models (HMMs) to resolve most of the difficulties in recognizing Arabic script was developed. Here, three features from the point sequence were extracted for each segment which are local-angle, super-segment, and loop-presence were used as features vectors for this system. The results were an average of 92% of both independent and dependent test. In the same year also [15], a trainable system for recognizing Arabic online handwriting and the system was called AraPen to recognize Arabic characters was

developed. Four features sets were used which are coordinates series to represent the *x*-*y* coordinates of the points that constitute the stroke, tangent angles series to represent the angles of the line segments that constitute the stroke, the winding value to represent the algebraic sum of direction changes from 0 for straight or s-shaped strokes to almost 360 for o-shaped ones and the aspect ratio to represent the height to the width ratio of the letter. The system results showed high recognition rate even by using a small collected corpus. The recognition rate in the non-cursive mode was about 91% and it increased to 98% after training the system with the user's handwriting. For cursive mode, the system gives a low accuracy with less than 50%.

In 2007 [16], a system to recognize online Arabic cursive handwriting using rule-based method has been proposed. For features extraction, three freeman chain codes were used. The first code is for representing eight long strokes and the second is for representing eight short strokes. The third is used to represent eight pen-up movements directions numbered from (1-8) and (0) for pen-up-down. The experimental results showed 92% success rate for recognizing the characters, with the same accuracy for the strokes, and 77% for the words. Another work in this year has been presented [17], here a structured model for recognizing online Arabic handwriting written in continuous formatting based on Hidden Markov Models (HMMs) to recognize Arabic strokes has been developed. For the feature extraction, they used a feature vector that was also adopted by[18] by using a digitizer that captures the successive pen-tip positions as a time-series of (x,y) coordinates and at each point, local position, angles, and curvature information are captured to make the feature vector. Then a technique called sliding windows with length of 10 samples was used to cover the whole word's feature vector sequence. The length of each window obtained from the shortest simple length of the strokes average sample length. The system was evaluated and gave an accuracy average of 75% for both strokes and letters.

In 2008, [19] a recognition system for recognizing handwritten Arabic characters was developed based on support victor machine (SVM) and using new feature called relative context (RC). RC was obtained from the relative pairwise distances and the angles of the writing trajectory. The system gave a high accuracy rate, comparing with previous systems, with 97.8%.

Due to the increase number of electronic devices which support online writing process, the research on text recognition has become a very popular area. Here, in the year of 2009 many research and studies were done. In [20], for instant, a new template matching scheme and applied to the recognition of cursive Arabic script has been presented. After the segmentation phase, each segment whether from the template database or the input sample is identified by the features angle (φ), arc type (T), length ratio (λ), connection angle (θ) and the relative positions R_x,R_y adding up to a feature point *f*. The results showed that the system gave the rate of 94.8% for isolated character recognition and word recognition rates of about 92%. In the same year [21], a handwriting recognition system based on visual coding and fitness evaluation function for Arabic script called (REGIM-CV) was developed. The feature extraction of the system was based on the Beta–elliptic model. This model considered a simple movement as the response to the neuromuscular system, which is described by an elliptic trajectory and a Beta velocity profile.

Another research done by Saabni and El-Sana [22]. The researchers presented a multi-level recognizer for online Arabic handwriting. Features extraction was done using two types of features from the body part which are global and local features. The global features include loops, ascenders, and descenders. The local features characterize local relation between adjacent or nearby points on the polyline. The system has been tested and several tests have being performed on the various datasets in which encouraging results were received. Later in the same year, an on-line Arabic handwritten recognition system based on a new stroke segmentation algorithm has been developed [23]. For extracting the features of the characters, Hu's seven moments [24] was used to extract the features for every letter which is considered as offline features extraction method. Hu's seven moment invariants are region-based image features which take all of the pixels of the image into account. The proposed system gave a recognition rate of up to 97% for words and 92% for letters recognition.

In [25] a decision tree and matching algorithm to recognize online Arabic character handwriting were presented. The Decision Tree makes the classifier faster to recognize characters and reduce the feature execution time, as the author claimed. Every stroke is represented by one Freeman direction to extract its features. The character recognition system was based on matching and similarity directional stroke string. The average accuracy rate of this system was 97.6%.

Another study was done by Al-Taani and Al-Haj [26] here, a recognition system based on Decision Trees to recognize online isolated Arabic characters was introduced. Furthermore, three main features in their system were used which are; number of segments, left-right density ratio, and bottom-up density ratio of each character. The system gave an accuracy rate in average of 75.3%. Next year, [27] an online Arabic handwriting system to recognize handwritten Arabic names was presented. In this system, HHM was used as a classifier and two main types of feature called local and vicinity features were used. Each of the main type got three sub features which are delta *x* and *y*, writing direction, and the range between each two points of the writing trajectory for local feature type. For the other type, there are the ratio of height-to-width of the vicinity bounding box, curliness, and the slop of writing trajectory. The results were an average about 94%.

In [28], a recognition system for online handwritten Arabic words based on Hidden Markov Models was introduced. Five features were used to identify each word which are chain code, turning angels, curvature, baseline, and vertical position between y value of every point and y value of the baseline. The system gave a high accuracy rate with average of 97%. On the other hand, for [29] a system to recognize online Arabic handwritten words which obtained from ADAB dataset was developed. The main contribution of this system was to avoid the delay strokes and remove it from the acquired word. Just one feature was used which is the local writing direction at every point of the trajectory. The system gave an accuracy rate reached up to 95.2%.

Researches on languages similar to the Arabic language were also conducted and caught the attention of several researches in this area like [30]. Here, an on-line system for recognizing handwritten Farsi (Persian) and Arabic digits based on statistical techniques was presented. After the Preprocessing phase and the representation of the writing by smaller units of windows or frames, a module extracts features of the data based on essential to the employed shape classification algorithm. The module uses the coordinates of the strokes points and order information of the points. The system was tested by 14 writers who wrote each digit five times. This produced an average accuracy of 93.14% for these writers.

Several researches works have been done for the same language such as [31] . The researchers presented an approach for recognizing on-line Farsi handwritten text based on the representation of input tokens with simple features and comparison of these representations with patterns of the same form. Four different features were derived for each segment: its type, direction, curvature direction, and relative length. The system gave output rate of 77% for all letters without any guidance to reach 95% with guidance by using pre-specified words and the accuracy rate up to 98% for letters which can be defined only with 3 or less pattern parts without any guidance.

In [32] the researchers proposed a hierarchical recognition system for online Farsi handwritten words based on tree decision combined with neural networks. A combination of three features which are Start-to-End Vector, Directional Vectors, and Sharp Edge Points, and Local Maxima were used for features extraction phase. The system which was tested by using 50 and 150 normalization size gave a recognition rate of 81% and 76% respectively. Also in [33] a deductive method based on finding the maximum and minimum local points of Farsi and Arabic script was implemented. The proposed online handwritten character recognition was initiated to recognize all Farsi and Arabic Script without any primary vowel points. Then, a comparison among different character shapes showed that it got higher recognition rate, that was 96.8% for isolated character shape and the lowest rate was 92.6% for final character shape.

Another study was done [34] to introduce online system for recognizing Farsi characters based on decision tree. The main aim was to identify and produce a group of features for 32 Farsi characters. For extracting features, nine groups were used which contained 24 sub features. 4000 handwritten Farsi character forms from TMU dataset were used and written by 117 writers. The system results were about an accuracy of 94%. Next Year, Khodadad et al [35] developed an approach to recognize isolated Arabic and Farsi characters based on Neural Networks. One feature vector was used in this system which constructed from the frequency domain coefficients using Discrete Cosine Transform (DCT). The eight lowest coefficients of each signal were just used to distinguish between the 32 isolated Arabic and Farsi characters. In the recognition phase, three layers feed forward Neural Network was used. The results of the system came out with an accuracy rate of 95.7%.

Recently [36], a recognition system was proposed to recognize Farsi and Arabic characters based on multilayer perceptron and back propagation Neural Networks as the classifier engine. They used three different features in their system which are the ratio of token length to its own stroke length, the direction of tokens of first stroke, and the type of middle coordinate point of tokens whether is counter-clockwise or clockwise. the result of this system were 96.8% for isolated form, 96.1% for beginning form, 95% for middle form, and the lowest was the final form with about 92.6%.

To sum up, Table II summarizes the features extraction techniques for Arabic texts in the past online Arabic handwriting systems and the accuracy rates of these systems.

V. CONCLUSION

Feature extraction is an important phase in text recognition systems and for many pattern recognition problems such as face, voice, signatures, and image processing recognition systems. It is used to distinguish between numbered patterns by using specific characteristics which are different from one pattern to another. Feature extraction for a text may be just one characteristic like recognizing numbers or isolated characters but for complicated situation like recognizing words and Arabic characters, here a combination of features should be used to distinguish between the pattern. The combination of features can give more general description of the character, word, digit, or stroke. Due to the nature of Arabic text, most of the systems and studies for Arabic online recognition have used the structure method of features. On the other hand, the success rate of any recognition system depends not only on the features extraction but it depends on several reasons such as the recognizer technique, the preprocessing stage, or the segmentation step. However, until now there is no specific method of choosing the feature way. It depends on the text and the system methodology. More researches are needed to find the best features techniques for online Arabic recognition systems.

1. TABLE II. TECHNIQUES USED FOR FEATURES EXTRACTION FOR ONLINE ARABIC TEXT RECOGNITION SYSTEMS

Features Type	Authors and Year	Features Technique	Text Type	System Rates
Structural	Almuallim & Yamaguchi, (1987) [10]	Pen movement	Words	81.25%
Structural	Al-Emami& Usher. (1990) [11]	Writing direction codes	Words	86%
Structural	Beigi et al. (1994) [30]	Differences between adjacent coordinates, sine and cosine of the angle between tangent line and stroke path, and the absolute Y- coordinate of each point with the baseline	Digits	93%
Combination	Bouslama & Amin (1998) [9]	24 local features related to the geometry, position, and size of a character represented by Fuzzy sets	Characters	Not mentioned
Structural	Halavati et al. (2005) [24]	Type, direction, curvature direction and relative length	Words and Characters	98%for characters 95% for words
Statistical	Rokbani et al. (2005) [13]	Heuristic method by using 13 visual codes	Words	92% dependent 70% independent test
Statistical	Al-Taani (2005) [12]	Extracting primary and secondary primitives	Digits	95%
Structural	Biadsy et al. (2006) [14]	Local-angle, super-segment, and loop-presence	Words	92%
Structural	Alsallakh & Safadi (2006) [15]	Coordinates series, tangent angles series, the winding value, and aspect ratio between height to the width of the letter.	Characters	98% dependent 91% independent test
Structural	Al-habian & Assaleh (2007) [17]	Pen-tip positions	Strokes and Characters	75%
Structural	Elanwar et al. (2007) [16]	Freeman chain codes	Characters and Words	92% for characters 77% for words
Structural	Faradji et al (2007) [25]	Start-to-end vector, directional vectors and sharp edge points, and local maxima	Words	81%
Statistical	Izadi & Suen, (2008) [19]	Relative Context (RC) feature	Arabic Characters	97.8%
Structural	Kherallah et al. (2009) [21]	Beta–elliptic model	Words	91%
Statistical	Daifallah et al. (2009) [23]	Hu's seven moments	Words and Characters	97% for words and 92% for letter
Combination	Saabni &El-Sana (2009) [22]	Global and local features	Words	Average of 89%
Structural	Sternby et al. (2009) [20]	Angle, arc type, length ratio, connection angle, and the relative positions Rx,Ry	Isolated Characters and Words	94.8% for characters 92% for words
Structural	Omer & Long, (2010) [25]	Freeman direction	Characters	97.6%
Structural	Al-Taani & Al-Haj, (2010) [26]	Number of segments, Left-right density ratio, and bottom-up density ratio	Isolated Characters	75.3%
Combination	Harouni et al (2010) [36]	Ratio of token length to its own stroke length. Direction of tokens of first stroke. Type of middle coordinate point of tokens.	Farsi/ Arabic Characters	96.8% isolated 96.1% beginning 95% middle 92.6% end form
Combination	Ghods & Kabir, (2010) [34]	Nine groups were used which contained 24 sub features	Farsi Characters	94%
Global Transforma-tions	Khodadad et al (2011)[35]	Discrete Cosine Transform (DCT)	Isolated Arabic/ Farsi Characters	95.7%
Structural	Ahmed & Azeem, (2011) [29]	Writing Direction	Arabic Words	95.2%
Structural	Hosny et al (2011) [28]	Chain code, Turning angels, Curvature, Baseline, and Vertical position	Arabic Words	97%
Structural	Abdelazeem & Eraqi, (2011) [27]	Local and Vicinity Features	Arabic Words (Names)	94%

REFERENCES

[1] K. Versteegh, M. Eid, A. Elgibali, M. Woidich and A. Zaborski, *Encyclopedia of Arabic Language and Linguistics*: Leiden Brill, 2006.

[2] M. A. Abuzaraida, A. M. Zeki and A. M. Zeki, "Segmentation Techniques for Online Arabic Handwriting Recognition: A survey," in *3rd International Conference on Information and Communication Technology for the Moslem World: ICT Connecting Cultures, ICT4M 2010*, Jakarta, Indonesia, 2010, pp. D37-D40.

[3] M. A. Abuzaraida, A. M. Zeki and A. M. Zeki, "Recognition Techniques for Online Arabic Handwriting Recognition Systems," in *International Conference on Advanced Computer Science Applications and Technologies (ACSAT2012)*, Kuala Lumpur, Malaysia, 2012.

[4] M. A. Abuzaraida, A. M. Zeki and A. M. Zeki, "Difficulties and Challenges of Recognizing Arabic Text," in *Computer Applications: Theories and Applications*, ed Kuala Lumpur: IIUM Press Malaysia, 2011.

[5] S. A. H. Al-Qahtani, M. Al-Sulaiman and M. S. Khorsheed, "Recognizing Cursive Arabic Script Using Hidden Markov Models," Master's Degree, College of Computer and Information Sciences,

Department of Computer Engineering, King Saud University, Riyadh, Saudi Arabia, 2004.

[6] A. D. Trier, A. K. Jain and T. Taxt, "Feature Extraction Methods for Character Recognition - A Survey," *Pattern Recognition,* vol. 29, pp. 641-662, 1996.

[7] I. Khodadadzadeh, "Recognition of Handwritten Arabic Characters," Master's thesis, Department of Electrical and Computer Engineering, The University of Windsor, Windsor, Faculty of Graduate Studies, Ontario, Canada, 2010.

[8] M. Z. Khedher, G. A. Abandah and A. M. Al-Khawaldeh, "Optimizing Feature Selection for Recognizing Handwritten Arabic Characters," *World Academy of Science, Engineering and Technology,* vol. 4, 2005.

[9] F. Bouslama and A. Amin, "Pen-based Recognition System of Arabic Character Utilizing Structural and Fuzzy Techniques," *International Conference on Knowledge-Based Intelligent Electronic Systems, Proceedings, KES,* vol. 3, pp. 76-85, 1998.

[10] H. Almuallim and S. Yamaguchi, "A Method of Recognition of Arabic Cursive Handwriting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. PAMI-9, pp. 715-722, 1987.

[11] S. Al-Emami and M. Usher, "On-line Recognition of Handwritten Arabic Characters," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 12, pp. 704-710, 1990.

[12] A. T. Al-Taani, "An Efficient Feature Extraction Algorithm for the Recognition of Handwritten Arabic Digits," *International Journal of Computational Intelligence,* vol. 2, pp. 107-111, 2005.

[13] N. Rokbani, M. Kherallah and M. Alimi, "On-line Handwriting Recognition System Based on Graph Matching and Genetic Algorithm," in *International Conference on Machine Intelligence*, Tozeur – Tunisia, 2005, pp. 959-963.

[14] F. Biadsy, J. El-Sana and N. Habash, "Online Arabic Handwriting Recognition using Hidden Markov Models," in *10th International Workshop on Frontiers in Handwriting Recognition*, La Baule (France), 2006, pp. 85-90.

[15] B. Alsallakh and H. Safadi, "AraPen: An Arabic Online Handwriting Recognition System," *Information and Communication Technologies,* vol. 1, pp. 1844-1849, 2006.

[16] R. I. Elanwar, M. A. Rashwan and S. A. Mashali, "Simultaneous Segmentation and Recognition of Arabic Characters in an Unconstrained on-line Cursive Handwritten Document," *International Journal of Computer and Information Science and Engineering,* vol. 4, pp. 203-206, 2007.

[17] G. Al-Habian and K. Assaleh, "Online Arabic Handwriting Recognition Using Continuous Gaussian Mixture HMMS," in *International Conference on Intelligent and Advanced Systems*, Kuala Lumpur, Malaysia, 2007, pp. 1183-1186.

[18] A. Biem, "Minimum Classification Error Training for Online Handwriting Recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 28, pp. 1041-1051, 2006.

[19] S. Izadi and C. Y. Suen, "Online Writer-Independent Character Recognition Using a Novel Relational Context Representation," in *Seventh International Conference on Machine Learning and Applications (ICMLA '08)*, 2008, pp. 867-870.

[20] J. Sternby, J. Morwing, J. Andersson and C. Friberg, "On-line Arabic Handwriting Recognition with Templates," *Pattern Recognition,* vol. 42, pp. 3278-3286, 2009.

[21] M. Kherallah, F. Bouri and A. M. Alimi, "On-line Arabic Handwriting Recognition System Based on Visual Encoding and Genetic Algorithm," *Engineering Applications of Artificial Intelligence,* vol. 22, pp. 153-170, 2009.

[22] R. Saabni and J. El-Sana, "Hierarchical On-line Arabic Handwriting Recognition," in *10th International Conference on Document Analysis and Recognition (ICDAR '09)* 2009, pp. 867-871.

[23] K. Daifallah, N. Zarka and H. Jamous, "Recognition-Based Segmentation Algorithm for On-Line Arabic Handwriting," in *10th International Conference on Document Analysis and Recognition*, Barcelona, Spain 2009, pp. 886-890.

[24] H. Ming-Kuei, "Visual Pattern Recognition by Moment Invariants," *IRE Transactions on Information Theory,* vol. 8, pp. 179-187, 1962.

[25] M. A. H. Omer and M. Shi Long, "Online Arabic Handwriting Character Recognition using Matching Algorithm," in *The 2nd International Conference on Computer and Automation Engineering*, 2010, pp. 259-262.

[26] A. T. Al-Taani and S. Al-Haj, "Recognition of on-line arabic handwritten characters using structural features," *Journal of Pattern Recognition Research,* vol. 1, pp. 23-37, 2010.

[27] S. Abdelazeem and H. M. Eraqi, "On-line Arabic handwritten personal names recognition system based on HMM," in *International Conference on Document Analysis and Recognition (ICDAR 2011)* 2011, pp. 1304-1308.

[28] I. Hosny, S. Abdou and A. Fahmy, "Using advanced Hidden Markov Models for online Arabic handwriting recognition," in *First Asian Conference on Pattern Recognition (ACPR)*, 2011, pp. 565-569.

[29] H. Ahmed and S. A. Azeem, "On-line Arabic Handwriting Recognition System Based on HMM," in *International Conference on Document Analysis and Recognition (ICDAR,2011)*, 2011, pp. 1324-1328.

[30] H. S. M. Beigi, K. Nathan, G. J. Clary and J. Subrahmonia, "Challenges of Handwriting Recognition in Farsi, Arabic and other Languages with Similar Writing Styles an on-line Digit Recognizer," in *Second Annual Conference on Technological Advancements in Developing Countries*, New York, 1994.

[31] R. Halavati, M. Jamzad and M. Soleymani, "A Novel Approach to Persian Online Hand Writing Recognition," in *Fourth World Enformatika Conference*, Istanbul, Turky, 2005, pp. 232-236.

[32] F. Faradji, K. Faez and M. S. Nosrati, "Online Farsi Handwritten Words Recognition Using a Combination of 3 Cascaded RBF Neural Networks," in *International Conference on Intelligent and Advanced Systems*, Kuala Lumpur, Malaysia, 2007, pp. 134-138.

[33] M. Harouni, D. Mohamad and A. Rasouli, "Deductive Method for Recognition of On-line Handwritten Persian/Arabic Characters," in *Second International Conference on Computer and Automation Engineering*, Singapore, 2010, pp. 791-795.

[34] V. Ghods and E. Kabir, "Feature Extraction for Online Farsi Characters," in *International Conference on Frontiers in Handwriting Recognition (ICFHR,2010)*, 2010, pp. 477-482.

[35] I. Khodadad, M. Sid-Ahmed and E. Abdel-Raheem, "Online Arabic/Persian character recognition using neural network classifier and DCT features," in *IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS,2011)* 2011, pp. 1-4.

[36] M. Harouni, D. Mohamad and A. Rasouli, "Deductive method for recognition of on-line handwritten Persian/Arabic characters," in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*, 2010, pp. 791-795.