A Novel Baseline Estimation Method for Arabic Handwritten Text Based on Exploited Components of Voronoi Diagrams

Atallah AL-Shatnawi
Department of Computer Information System, Al-albayt University, Jordan

Abstract: *The goal of this paper is to present an efficient novel baseline estimation method for Arabic handwritten text based on the exploited components of Voronoi Diagrams (VD). The proposed based-VD method is constructed from three stages including: Preliminary stages, VD construction and baseline estimation process. The edges of the text are firstly extracted and then both inner and outer contour are traced in order to be converted into a set of sampling points. These sampling points are used to be the VD generators. Then, the baseline is detected from those edges and vertices which are positioned within the text boundaries. The proposed method is implemented, verified and validated on the IFN/ENIT Arabic handwritten dataset. It is discussed and compared with the horizontal projection method against the IFN/ENIT dataset based on affecting by noise, working properly with or without diacritics, working properly with the skewed images and running time efficiently. The results obtained are shown in graphs and output images. The results demonstrated that the proposed method works properly with the skewed and noisy images and with or without diacritics. It also, able to estimate the Arabic text baseline in straight or in curved line.*

Keywords: *Arabic text recognition, baseline estimation, VD, horizontal projection.*

Received December 7, 2013; accepted July 9, 2014; published online June 11, 2015

1. Introduction

Human reading capabilities are a benchmark for all character recognition systems, the main aim being able to match the reading, understanding and editing capabilities of a human. Arabic Character Recognition (ACR) has advanced on par with all other character recognition systems despite the fact that language has complex characters with cursive structure. The recognition rate achieved in ACR system is quite good considering [2, 4] the above fact. The recognition rate is a simple function of the five processing stages as image acquisition, pre-processing, segmentation, feature extraction and recognition. However, the pre-processing stage is considered as a heart of ACR system. It is an amalgamation of various sub processes like smoothing, noise removal, skew detection, noise removal, baseline estimation and many more. Therefore it is directly influence the dependability and efficiency of the other ACR stages [1, 2, 4, 6, 14].

The Arabic text is cursively written and the baseline is defined as the imaginary line which connects all the characters of a word [2, 4]. Based on this definition the characters and Shape of Arabic text are categorized as ascenders, descenders and diacritics, as shown in Figure 1. Ascenders and descenders lie above and below the baseline respectively. Diacritics are vital in the sense a diacritic marking between words changes the possible meaning [4]. The main uniqueness of ACR starts with baseline estimation. Hence, the baseline can be utilized in either Arabic text segmentation or for skew normalization as well as fordependent features extraction [9, 10].

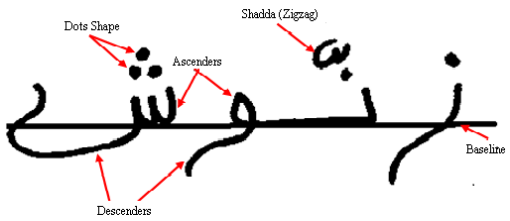


Figure 1. Ascenders, descenders and special marks in Arabic text.

2. The State of the Art

In this paper, VD is used as a tool for estimating the Arabic handwritten text baseline. This section provides and presents an overview about the Arabic baseline estimation methods and about VD background, definitions and components. Furthermore the dataset used in this paper is presented.

2.1. Arabic Text Baseline Estimation Methods

Enough research has been done focussed on Arabic baseline estimation for over three decades. The beginning was marked by Parhami and Taraghi [13] where in the horizontal projection method was applied in calculating baseline. This work was later carried on by Timsari and Fahimi [17]. Since, then research has seen a paradigm shift with advent of new techniques depending on contour tracing, skeletal extraction and principle component analysis [4].

The horizontal projection method gives a basic understanding of the underlying concept of baseline estimation. The histogram of pixel count in different rows is used to find the baseline. In fact, the peak of the histogram itself is considered baseline. The histogram projection method reduced the analysis of two dimensional images to one dimensional. The horizontal projection method used as the first attempt for Arabic text baseline estimation by [13]. The thickness of baseline is calculated using the most frequent column-height [17] or by taking into account the location of iteration as a reference since they are often near close to the baseline as illustrated in [12]. Pechwitz and Maergner [15] proposed an Arabic text baseline estimation method based on selecting baseline relevant features which are taken from the text skeleton.

Farooq *et al.* [9] proposed an Arabic text baseline estimation method based on selecting local minimal points which are generated during the text contour representation. Burrow [7] proposed a system to estimate the Arabic word baseline direction based on the Principal Components Analysis (PCA) according on distribution the foreground or the background pixels. In this paper, a novel baseline estimation method for Arabic handwritten text depending on the exploited component of the Voronoi Diagrams (VD) is presented.

2.2. Voronoi Diagram

VD have innumerable applications not limited to 3D space and surface mapping, microstructure modelling, matter disposition in space and environment. VD is induced by a set of points (called sites) subdivision of the plane where the faces correspond to the regions of close site. The VD concept was defined as late as 1644. Later, dirchlet and voronoi have reinvented voronoi concept and extended to 3D space. The VD's were given different names, generalized and studied before finding their application in various fields include "marketing, physics, astronomy, chemistry, image processing, medicine, microbiology, networks, imagery, geography and telecommunication and many others" [3, 5, 18]. VD tends to be involved in scenarios where a space should be partitioned into irregular lattices called spheres of influence. Thus, in addition to geometrical construction, it broadly used in many other fields [3, 5, 11].

When the perpendicular bisectors of the edges and the central point are connected together, this is called VD. In Figure 2, consider P is a plane and consider n different points which are called voronoi generators. VD is formed when plane P is divided into cells or geometric objects, each of these cells correspond to one point [3, 5]. A point q will be located in the cell p_x if and only if its distance (i.e., q) from point p_x not greater than its distance from any other point p_y. The voronoi region is the cell P_x [3, 5]. Mathematically [3]:

$$q \in Pi \text{ iff } |q - Pi| < |q - Pj|, \quad \text{where } i \neq j$$

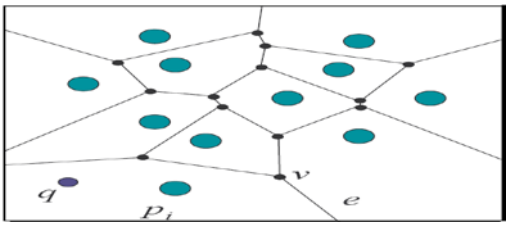


Figure 2. VD plane of a set of points P_x, where e is voronoi edge, q is free point and v is voronoi vertex [3, 5].

VD has the following properties:

1. *A Voronoi Vertex*: Is the central of a sphere of influence.
2. *Voronoi Regions*: Are finite and infinite convex polygons.
3. *Voronoi Edges*: Are line segments, infinite lines or half lines, which are boundaries of the regions of the voronoi. Voronoi edges are formed by lining perpendicular bisectors to the segment of line that link couple of points in the plane. While voronoi neighbors are sites of the two points that develop the voronoi edge [3, 5].

2.3. The IFN/ENIT Datasets

Database of 569 Arabic images were used as a dataset for validation and examination the proposed methods and which implemented in this paper. The images of Tunisian town's names are handwritten. The database is an example version of the complete IFN/ENIT handwriting database which is determined the most popular database for testing of Arabic handwriting text recognition system. The text and size of the file, resolution and dimensions of the data to be fed to the database is of a predefined format. The file format is BMP with a resolution of 300 dpi with a size of 18KB. The image dimensions are also capped with lower and higher boundaries set for the width /height as 90/50 to 976/161 respectively. No specific restriction was placed on the number of characters in the text except that all of them should be less than 3 words [16].

3. The Proposed Baseline Estimation Based VD Method

The proposed baseline estimation method is a multi-stage technique. It consists of three main stages including: Preliminary stages, VD construction and baseline estimation process stages. Figure 3 presents the improvement of the voronoi based method which is proposed in this paper.

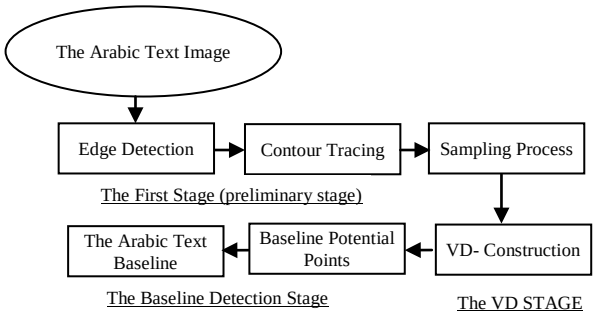


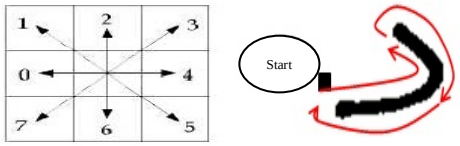
Figure 3. The proposed baseline estimation based-VD method development framework.

1.1. Preliminary Stage

In this stage the data is prepared for the next stage. It is built from three image pre-processing operations. These operations are: Edge estimation, contour tracking and identification and sampling process. In this phase, the input script pattern edges are identified, and then both the inner and outer contours are tracked, in order to be converted into a set of sampling points. These sampling points are used to be the VD generators. The first stage is implemented through the following steps.

1.1.1. Edge Detection and Contour Tracing

In this step, the edges are detected by generally searching for a 0-1 or 1-0 transition in pixel values and generally this is done using a window size of 3×3. For any given pixel, 8 neighbours are to be compared in all the possible directions (south, south-west, west and north-west, north, north-east, east, south-east). The contour representation is just an extension of this method which looks at marking the edges and removing the other content of the image by changing the 1's to 0's for the centre pixel. Figure 4 presents the suggested 8-neighbourhod contour representation and the direction of process [5].



a) direction. b) process.

Figure 4. The proposed 8-neighbourhood contour representation: a, b.

1.1.2. Sampling Process

This stage requires the down sampling of the pixels generated from contour tracing in the previous step. The first pixel is chosen based on special function which looks at the image from left to right column wise. Given a starting point the subsequent operation involves down sampling of contour by choosing every R^{th} pixel which is retained while removing all other pixels in between, R is a sampling argument, R is greater than 1. This function traverses in the clockwise direction. When R=1, all contour pixels are chosen. The sampling interval R is chosen based on the VD construction.

1.2. VD Construction

VD is usually built from a set of VD generators (i.e., sampling points). It alters the text into arithmetical shapes of convex holes and polygons, involves edges and vertices. In this stage, the point-VD construction will be described through an example. The Arabic word 'Mothleen' (مثلين) presented in Figure 5 is firstly read. Then, the proposed 8-nigbourhood contour representation method is used to extract the edges and to trace the contour. Finally, the VD generators been chosen using R=4 along the contour and the VD is then built using all the generated samples, as shown in Figure 6.

a) Binary image. b) Edge detection and contour tracing. c) Sampling using R=4 (VD- Generators).

Figure 5. The Arabic handwritten word 'Methleen' (مثلين): a, b, c.

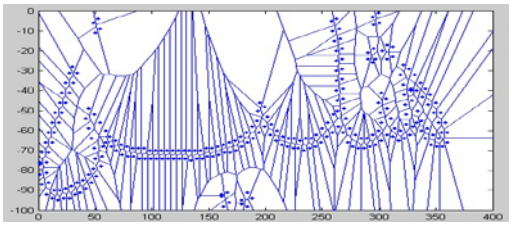


Figure 6. VD built from the generators shown in Figure 5-c.

A part of the VD in Figure 6 is zoomed in Figure 7, in order to present the information of VD-components. Voronoi edges of point VD have been categorized into four following types based on their position in the base text (the edges are identified in Figure 7 accordingly):

1. Edges that located fully inside the text (i.e., linked component or subword).
2. Edges that located fully outside the text.
3. Edges that located partially inside the text.
4. Edges go to the infinite.

Depending on Figure 7, the voronoi vertices are also categorized into the following types depending on their position with respect to text object (the vertices are identified in Figure 7 accordingly).

a. Vertices coordinate positioned in the text.
b. Vertices coordinate positioned outside of the text.
c. Vertices with unknown coordinates (virtually considered to have coordinates at infinity).

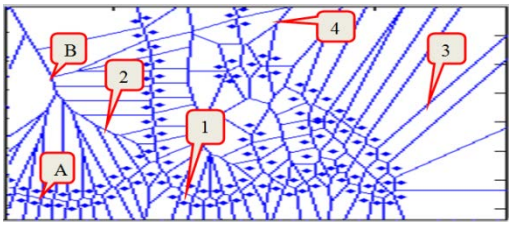


Figure 7. Types of voronoi vertices.

The above recorded types of voronoi vertices and edges will be utilized in estimating the Arabic handwritten text baseline in the next section.

1.3. Baseline Estimation

The voronoi generators are the sampling points which are resulted from the text contour tracing. The VD is built from those extracted generators. The proposed baseline based VD estimation method estimates the baseline in two steps. In the first step, the baseline potential points are selected based on those voronoi vertices and edges which are positioned inside the text boundaries (i.e., type "A" and "1" of the voronoi vertices and edges). Any vertexes satisfy one or more of the following conditions, it will be a baseline potential points:

1. If three or more edges connected in each other in the same vertex.
2. If the angle between two adjacent edges (are connected by the same vertex) is approximated between 70 to 90 degree.
3. If only one potential point is detected in any connected component, this point will be ignored.

The potential points of Figure 6 are presented in Figure 8. Then, in the second step the baseline can be estimated in curved or in straight baseline. The curved baseline is estimated by joining all the edges between two or more potential points located in the same connected component, while the straight baseline is estimated by joining all the potential points by using the liner line. Figure 9 presents the curved and the straight baseline estimated of Figure 8.

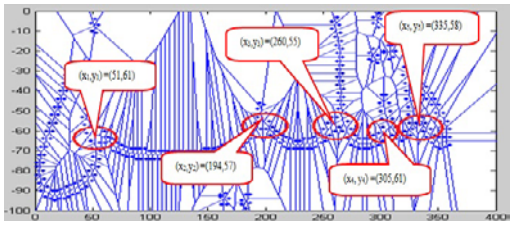


Figure 8. The potential points of the Figure 6.

a) Curved baseline. b) Straight baseline .

Figure 9. Arabic handwritten text baseline estimated of Figure 8-a, b.

Algorithm 1 proposed baseline based-VD algorithm steps are implemented to estimate the Arabic text baseline by choosing those edges and vertices of type A and 1:

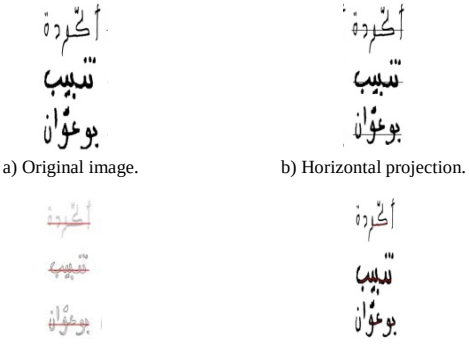
Algorithm 1: The proposed baseline estimation method based on the exploited components of VD.

Input: *Binary text image.*
Output: *Baseline image.*

{*a number identifying the connected components is given to each pixel in text image, while background pixels are identified by a zero value.*
inner and outer contours are tracked.
the voronoi generators are identified during the contour tracking using fixed interval R.
point_VD is constructed using all the generated samples.
study type "A" of the voronoi vertices and the type "1" of the voronoi edges as well as ignore the others VD components.
if three or more edges are connected to each other in the same vertex.
or if the angle between two adjacent edges approximated between 45 to 90 degree.
then consider these vertices as baseline potential points.
delete all the individuals potential points (*i.e., the single potential point in one connected component*).
join the selected vertices by a straight line or make it in curved form.
End
}

4. Experiment Results

The horizontal projection method (which is selected as benchmarking algorithm) and the proposed baseline estimation based VD method have been implemented using MATLAB programming language installed in the personal computer with processor speed core i3 2.13GHZ. The proposed and horizontal methods are applied on the example version of the IFN/ENIT dataset. Figure 10 presents examples of the Arabic text baseline obtained using the proposed based VD method (in straight and curved line) and the horizontal projection methods implemented in this paper.



a) Original image. b) Horizontal projection.

c) Proposed method (straight). d) Proposed method (curved).

Figure 10. Examples of the Arabic text baseline obtained by: a, b, c, d.

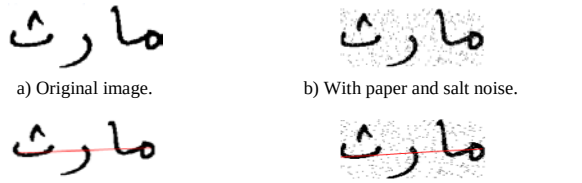
5. Discussion

An effective baseline estimation method for Arabic text should ideally deal with the following requests: working properly with or without diacritics and noise, working efficiently with the skewed images and running time properly. The results obtained by the proposed based VD and the horizontal projection methods were discussed and compared against the IFN/ENIT dataset depending on the following criteria.

5.1. Noise Removal

Arabic baseline estimation methods are negatively affected by noise. Therefore, the noise should be filtered prior to the baseline estimation process [6]. Many image processing operations can remove the

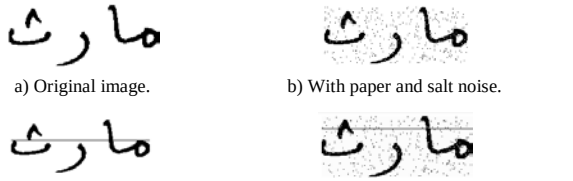
noise such as Gaussian or Median filters. The salt and pepper noise method is implemented on the Arabic word "Marth" (مارث) to verify that the proposed based VD and the horizontal projection methods are working probably with or without noise images. See Figures 11 and 12.



a) Original image. b) With paper and salt noise.

c) The text baseline estimated using the proposed method of image a. d) The text baseline estimated using the proposed method of the noise image b.

Figure 11. The Arabic handwritten word "Marth" (مارث): a, b, c, d.



a) Original image. b) With paper and salt noise.

d) The text baseline estimated using the horizontal projection method of image a. d) The text baseline estimated using the horizontal projection method of the noise image b.

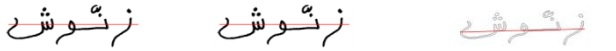
Figure 12. The Arabic handwritten word "Marth" (مارث): a, b, c, d.

Based on Figures 11 and 12 the noise does not effect the proposed baseline estimation based on VD method performance because it detects the baseline from those edges and vertices which are positioned within the text boundaries. On the other hand the horizontal projection method highly affected by the noise and the baseline cannot be detected.

5.2. The Diacritics (Dots, Madah, Shadda and Zigzag)

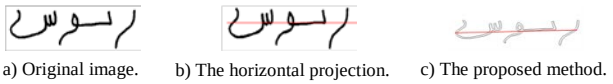
Not only due to the cursive nature of the Arabic text which poses complex questions to the ACR systems but also due to the diacritics which are often seen as dots or zigzag located upper or below the baseline. The ACR system becomes bulky and time consuming resulting in degraded accuracy. Researchers have found a workaround and suggested the removal of diacritics during post processing stages with specific recorded information on their location. This information can be used at a later point of time to make the OCR system more robust and accurate [8]. In this paper, an experiment is conducted to verify that the proposed based VD and the horizontal projection methods are working probably with or without diacritics. Therefore, the Arabic handwritten word 'Zanoush' (زنوش) which consist of five dots and "Shadda" has been chosen. Figure 13 presented the estimated baseline by the proposed based VD and the horizontal projection methods with diacritics. Figure 14 presents the results obtained by the two methods without diacritics.

Based on Figures 13 and 14 the proposed baseline estimation based on VD method can estimate the Arabic baseline estimation with and without diacritics, while the horizontal projection method did not do so.



a) Original image. b) The horizontal projection. c) The proposed method.

Figure 13. The estimated baseline with diacritics: a, b, c.



a) Original image. b) The horizontal projection. c) The proposed method.

Figure 14. The estimated baseline without diacritics: a, b, c.

Based on Figures 13 and 14 the proposed baseline estimation based on VD method can estimate the Arabic baseline estimation with and without diacritics, while the horizontal projection method did not do so.

5.3. The Skewness

The Arabic text skewness is determined as one of the main difficulties facing the Arabic baseline estimation methods performances. The robustness of the Arabic text baseline estimation algorithms to rotation should be tested by using various rotated images. An experiment is done to test the robustness of the proposed and the horizontal projection methods to rotation. A text image of Arabic word 'نقة' has been rotated by 90 and 180 degrees and the baseline is estimated from the resulting images using the proposed baseline estimation and the horizontal projection methods, the results is presented in Figure 15. Based on Figure 15 the proposed baseline estimation based on VD method can estimate the Arabic baseline estimation in the different rotated images, while the horizontal projection method did not do so.

a) Arabic text (rotated by 90 and 180 degrees). b) Horizontal projection method. c) Proposed based-VD method.

Figure 15. Baseline estimated for different rotation (90 and 180) using: a, b, c.

5.4. Speed

The processing times for the two methods are measured in seconds depending on the CPU time taken during the methods execution. Figure 16 presents the total average processing time obtained by the two baseline estimation methods in baseline processing the 569 text images of IFN/ENIT dataset. Based on Figure 16, the horizontal projection method is faster than the proposed method. As well as, estimating the straight line by using the proposed method is faster than the curved one.

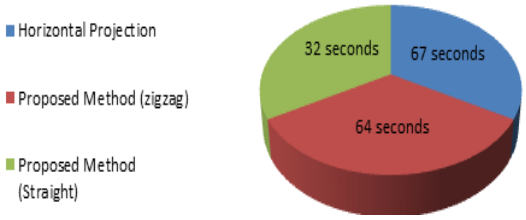


Figure 16. The total average processing time obtained by the two baseline estimation methods.

6. Conclusions

In this paper a novel baseline estimation method for Arabic text depending on the exploited component of the VD was presented, implemented and discussed. Baseline estimation is one of the most important in the pre-processing stage of ACR. It has a direct impact on the performance and dependability of the segmentation, skew normalization and feature extraction stages of ACR.

The proposed Arabic baseline estimation method consists of three main phases: Preliminary phases, VD constriction, and baseline estimation process phases. The contour was traced at the long of the text and the sampling points were recorded to be the VD generators. Then, the VD is constructed from those sampling and the baseline is detected from the vertices and edges which are positioned within the text boundaries. The proposed method was executed, verified and validated on the IFN/ENIT Arabic handwritten dataset. It is discussed and compared with the horizontal projection method against the IFN/ENIT dataset based on affecting by noise, working properly with or without diacritics, working efficiently with the skewed images and running time properly. The yielded results are presented in output images and graphs.

The results proven that the proposed Arabic baseline estimation method works efficiently with the slanted and the skewed text images. Hence, it does not need normalization process or any slant, skew correction. As well as, it is also capable to handle various sizes of text images, as well as it does not affect by noise. Furthermore it works probably with or without diacritics. On other hand the horizontal projection baseline estimation method affected by noise and diacritics. Due to the fact, VD construction is a time consuming stage, the proposed method recommend using the VD components in other OCR stages. Finally this paper concluded that, the proposed based VD method is the first Arabic text baseline estimation method able to estimate the Arabic text baseline in straight or in curved line.

References

[1] Abdulla S., Al-Nassiri A., and Abdul Salam R., "Off-Line Arabic Handwritten Word Segmentation using Rotational Invariant Segments Features," *the International Arab Journal of Information Technology*, vol. 5, no. 2, pp. 200-208, 2008.

[2] Al-Shatnawi A. and Omar K., "A Comparative Study between Methods of Arabic Baseline Detection," *in Proceedings of International Conference on Electrical Engineering and Informatics*, Malaysia, pp. 73-77, 2009.

[3] Al-Shatnawi A. and Omar K., "Detecting Arabic Handwritten Word Baseline using Voronoi Diagram," *in Proceedings of International Conference on Electrical Engineering and Informatics*, Malaysia, pp. 18-22, 2009.

[4] Al-Shatnawi A. and Omar K., "Methods of Arabic Baseline Detection-The State of Art," *International Journal of Computer Science and Network Security*, vol. 8, no. 10, pp. 137-142, 2008.

[5] Al-Shatnawi A., "A Non-Iterative Thinning Method Based on Exploited Vertices of Voronoi Diagrams," *PhD Thesis*, University Kebangsaan Malaysia, Malysia, 2010.

[6] Bouafif F., Maddouri S., and Ellouze N., "A Hybrid Method for Three Segmentation Level of Handwritten Arabic Script," *the International Arab Journal of Information Technology*, vol. 9, no. 2, pp. 117- 123, 2012.

[7] Burrow P., "Arabic Handwriting Recognition," *M.Sc Thesis*, University of Edinburgh, UK, 2004.

[8] Cowell J. and Hussain F., "Thinning Arabic Characters for Feature Extraction," *in Proceedings of the 5th International Conference on Information Visualization*, London, pp. 181-185, 2001.

[9] Farooq F., Govindaraju V., and Perrone M., "Pre-Processing Methods for Handwritten Arabic Documents," *in Proceedings of the 8th International Conference on Document Analysis and Recognition*, pp. 267-271, 2005.

[10] Nawaz S., Sarfraz M., Zidouri A., and Al-Khatib W., "An Approach to Offline Arabic Character Recognition using Neural Networks," *in Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems*, pp. 1328-1331, 2003.

[11] Okabe A., Boots B., and Sugihara K., *Spatial Tessellations*: *Concepts and Applications of Voronoi Diagrams*, USA: Prentice Hall, 1992.

[12] Olivier C., Miled H., Romeo K., and Lecourtier Y., "Segmentation and Coding of Arabic Handwritten Words," *in Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, pp. 264-268, 1996.

[13] Parhami B. and Taraghi M., "Automatic Recognition of Printed Farsi Texts," *Pattern Recognition*, vol. 14, no. 1, pp. 395-403, 1981.

[14] Parvez M. and Mahmoud S., "Offline Arabic Handwritten Text Recognition: A Survey," *ACM Computing Surveys*, vol. 45, no. 2, pp. 23-35, 2013.

[15] Pechwitz M. and Maergner V., "Baseline Estimation for Arabic Handwritten Words," *in Proceedings of the 8th International Workshop on in Frontiers in Handwriting Recognition*, pp. 479-484, 2002.

[16] Pechwitz M., Maddouri S., Märgner V., Ellouze N., and Amiri H., "IFN/ENIT-Database of Handwritte," available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.90.1245&rep=rep1&type=pdf, last visited 2002.

[17] Timsari B. and Fahimi H., "Morphological Approach to Character Recognition in Machine-Printed Persian Words," *in Proceedings of SPIE*, pp. 184-191, 1996.

[18] Zeki A., Ghyasi A., Mujahid M., Zainul N., Cheddad A., Zubayr M., and Zakaria M., "Design and Implementation of a Voronoi Diagrams Generator using Java," *in Proceedings of International Arab Conference on Information Technology*, Algeria, pp. 194-198, 2004.

Atallah AL-Shatnawi received his BSc in Computer Science from Yarmouk University (Jordan) in 2005, MSc in Computer Science from the University Science Malaysia (USM) and PhD in System Sciences and Management from the National University of Malaysia (UKM) in 2007 and 2010 respectively. Currently, he is an Assistant Professor at the Department of Information Systems, Prince Hussein Bin Abdullah College for Information Technology, Al Al-Bayt University (Jordan). His research interests include: Expert Systems, Pattern Recognition, Image Analysis and Processing as well as Embedded Systems. He has published numerous papers related to these areas.