Instant Arabic Translation System for Signboard Images Based on Printed Character Recognition

Rafeeq Abdul Rahman A. Al-Hashemi and Shoroq Almamon Alsharari

***Abstract*—Daily lives contain many of the menus and signboards which carry important information, but sometimes it may cause problems when we cannot understand it. This paper aimed to develop a new system that translates Arabic texts of the signboards into English text by using mobile phone camera. In spite of the diversity of translation products of text embedded in images for many languages, Arabic texts seem to be not yet well solved to address this problem. The system will automatically translate the Arabic text embedded in images into English language. Four subsystems used in the algorithm: preprocessing, segmentation (text detection), character recognition and translation. Dealing with Arabic language was the most important problem faced by the proposed system because it has a set of characteristics makes the identification very difficult, such as words interrelated. The system automatically detects the text and works well with different backgrounds, rotated images, skewed, font sizes and blurred images. The system was assessed using recall measurement to evaluate the performance of the developed system and the experimental results of character recognition show a rate of 81.82%, the word recognition subsystem gave a rate of (94.44%) and the word translation was about (83.33%).**

***Index Terms*—Arabic translation, information extraction, character recognition.**

I. Introduction

The proposed system comes to solve the problem of translate Arabic texts of the signboards into English text by using mobile phone camera. through using the Image translation system by Optical Character Recognition OCR. An image translation system generally consists of three core techniques: text extraction, OCR, and translation. The procedure of image translation is as follows: extracting text from an image taken by the camera .After selecting the text through the previous steps, the text is entered into the OCR system finally translating the results into the target language [1].

Bouslama [2] proposed a new recognition method for machine printed Arabic characters that combined the structural and statistical for feature extraction and used fuzzy rules for classification. Hamid [3] proposed algorithm that segments handwritten Arabic text. Through the segmentation of the text into blocks of characters, it generates pre-segmentation points for these blocks, and then it uses a neural network to verify the accuracy of these segmentation points.

The Arabic character recognition has received less attention than Latin character recognition so our paper scope focuses on Arabic OCR. Many works have been undertaken in the area of Arabic character recognition but with limited success, due to the nature of Arabic characters and other problem such as the Arabic alphabet consists of 28 basic characters. Some characters may have different shapes depending on their position within a word (beginning, middle, and end) and their different size (height and width). Furthermore, 16 Arab characters have one dot or two dots or three dots, or are zigzag, which are used to differentiate between characters. Park [4] presents a system for automatic detection of signboard texts which are captured from a mobile phone. The proposed system can detect and binaries texts from images. Mollah [5] developed a text/graphics separation methodology for camera captured business card images and implemented a fast skew correction technique for the text regions extracted from business card image at the begging, the background based on intensity variance is eliminated. This makes the foreground components distinct from each other. Then the non-text components are removed using various characteristic features of text and graphics. Finally, the skew text regions are corrected. Kumar [6] the Two - phase method reduces the skewed data before applying the skew detection and correction methods of accuracy and speed. This technique helps in reducing the computational time required for Hough Transform considerably with less compromise on the accuracy front and speeds up it.

The system is designed to translate different signboards from the Arabic language into the English language. The system starts by tacking the input image and ending up with translating the text. The image translation system is divided into four subsystems: image processing, segmentation, character recognition and translation.

II. Proposed System

The Image translation system based on Arabic OCR architecture is shown in Fig. 1.

A. *Preprocessing*

The importance of the preprocessing stage lies in preparing the words to be in their final shape before entering the recognition stage. After the image is captured, a number of preprocessing steps are performed for these images. These steps are:

1) Binarization.
2) Noise removing.
3) Inversing the binary image
4) Skew detection and correction.

Manuscript received March 19, 2013; revised July 23, 2013.

The authors are with Al-Hussein Bin Talal University (AHU), Jordan (e-mail: Rafiq_alhashimy@yahoo.com).

DOI: 10.7763/IJMLC.2013.V3.344

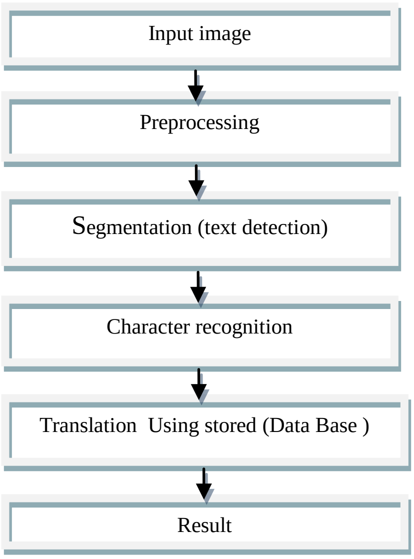

Input image
Preprocessing
Segmentation (text detection)
Character recognition
Translation Using stored (Data Base)
Result

Fig. 1. Steps of the proposed system for translation.

1) Binarization

This process aims to convert the image into binary image. In this step we first convert the image into grayscale. The rgb2gray converts the RGB image or color map into grayscale then converts the gray values into the binary values .This step is done by comparing the gray values of an image with a given threshold. This threshold value is measured by finding the dominant gray value in the input image, and then choosing the threshold value to be the center point between the dominant value and the maximum gray value. In this stage threshold = .80%.

2) Noise removal:

The image always contains noise that usually appears as an extra pixel (black or white) in the character image. If the noise is not taken into consideration, it could subvert the process and produce an incorrect result, so we use the median filter and fspecial filter to remove the noise from the image.

3) Image inversion:

Image inversion means inversion the pixels of an image: the zero pixels to ones and vice versa. Inversing the Image is necessary since after applying the binarization, the background of the image takes the ones pixels and the foreground of the image takes the zero pixels. At this stage the inversion function is applied to the image for this purpose.

4) Skew detection and correction:

Skew occurs when the Camera is not parallel with the images [5]. In this stage, at first system must calculate the amount of skew in the image and then correct skew. It is an important preprocessing step to increase the accuracy of the subsystem processes, such as character recognition. For skew correction, we will implement vertical histogram algorithm.

B. Segmentation

Segmentation step identifies the image area that contains the text. It is the most important process that improves the system results because when the system selects the text area accurately, the process of identifying words becomes easier and therefore gives the system good results. Text detection is the process of partitioning the digital image into multiple regions that can be associated with the properties of one or more criterion. These properties are gray level, color, texture shape, and others.

The system defines the text area (ROIs) by using a horizontal projection profile technique [7]. The algorithm scans the imput image horizontally to find the first and last black pixels in a line. Once these pixels are found, the area in between these pixels represents the line that may contain one or more character.

C. Character Recognition

Character recognition is one of the most critical subsystems in this research. The Arabic CR presents, including:

- The Arabic script is cursive.
- Connected characters, where characters can have different shapes in different positions of a word.
- 15 of the 28 Arabic letters include dots.
- A word is composed of sub-word (s).

These characteristics have made the progress of Arabic OCR more complex and difficult compared to other languages. Template matching algorithm was suggested in order to find the similarity between ROI and template of the letter. Excellent results were obtained that enabled the system to get excellent translation and the following algorithm is applied in order to recognize all characters in the original image:

1) Template is passed on the segment and the correlation between template and ROI (region of interest) in the segment is calculated.
2) Image resizing is applied to normalize size of the template to the size of the extracted segment.
3) The decision is made based on pre-assigned threshold that the candidate character is similar or not similar to the template. The threshold found to be 85% for acceptable performance. If the character is recognized , a code of the letter is obtained based on the number of the template.
4) An encoding subsysyem is applied her on each recognized character by using a matrix that contains the codes of the templates. For example "معان محافظة " starts with the letter "م" its code is 094. This process is repeated for all letters from the beginning to the end of each segment.
5) The process is repeated for all template characters from the beginning to the end of each segment. Notice that after coding the Arabic character in recognition phase, the letter is translated into english. As an example "هـ" letter in Arabic will be "H" in English.

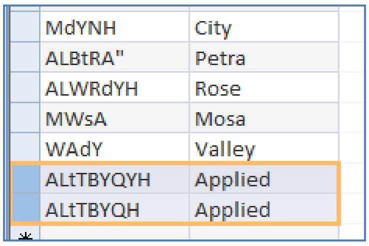
The algorithm is applied on the generated matrix and started by reading the matrix from right to left (like Arabic language) and concatenates the codes into a single line and multiple lines into one text for further processing. The text is divided into multiple words using the space character and fed to the translation subsystem.

D. Translation

As described in the previous section, at the end of recognition phase the output is passed to the translation subsystem. The system uses a database that contains dictionary. The database used in the system contains words translated for tourism *translation application. The system can easily be applied to other application by feeding the needed words and their matched translation of the database. The recognition subsystem will recognize the word and attempt to translate it using the same translation subsystem.

The translation subsystem takes the word from the recognition subsystem and searches the database. If the word exists in the table, the translation is obtained and appended to the output string. If the word is not found in the database word, the translation is assumed to be the same word and also fed to the output string. This will be helpful for identifing names of people or objects. For instance, the line “ مطعم أحمد- وجبات سريعة” in Arabic language is recognized as “MTAM AHMD WJBAT SRYAH” and after translation it becomes “Restaurant AHMD –FAST FOOD”. Here “AHMD” is translated also as “AHMD” because it is not found in the database. This will help the user of the system at least to identify the objects by its similar translation in English. Substitution of letters for translation.

Misrecognized characters can be optionally substituted to compensate and to improves the recognition process at the last stage of translation. This substitution is not arbitrary but based on the fact that many characters are very similar to each other in Arabic language, For instance “ـحـ” is very similar to “ـخـ” and “ـجـ”. For certain threshold (85%) that used for recognition subsystem one of these characters may be recognized as incorrect. For the system to substitute a letter for another, they both must be very similar, or the word recognized with the original character is very similar to one existed in the database. Fig. 2 depicts an example of using this enhancement. In Fig. 2, the system attempts to recognize the word “التطبيقية” from the Arabic language. Now instead of recognizing it completely “ALtTBYQYH”, the system recognizes it “ALtTBYQH” which misses the character “ـيـ” in the Arabic language or “Y” in recognition system. Because the system didn’t recognize it well, the substitution of this letter could be attempted by inserting another structure for the word “التطبيقية” in the Arabic language, and this will enhance the overall translation process for the English language which will be translated using this method to “applied” word in the English language for better translation.



	MdYNH	City
	ALBtRA"	Petra
	ALWRdYH	Rose
	MWsA	Mosa
	WAdY	Valley
	ALtTBYQYH	Applied
	ALtTBYQH	Applied
*		

Fig. 2. Coding and Substitution of missed character System.

III. Experimental Results and Discussion

A. Factors Measured in the System

The Evaluation factors that used for measuring the system performance are as the following:

- Rate of characters recognized.
- Rate of words recognized.
- Rate of words translated.

Rate of characters recognized is obtained by dividing the number of recognized characters through the sample image by the total number of characters that counted in the image by the user of the system. For convenience, the number of total characters is entered the system by a text file assigned to the image file as in equation (1).

$$\text{Recall} = \frac{\text{Number of correctly recognized characters}}{\text{Total number of characters}} \quad (1)$$

Rate of word recognized is calculated by counting the combined characters into word from the recognition systems divided by the total number of words that must be recognized by the system entered using text file assigned together with the sample image file as in equation (2).

$$\text{Recall} = \frac{\text{Number of correctly recognized characters}}{\text{Total number of words}} \quad (2)$$

Rate of the word translated is computed by dividing the total number of recognized words that translated successfully by the dictionary by the total number assigned in the text file attached to the sample image. The translation process is dictionary based translation neither lexical nor grammar analysis carried by the system as in equation (3).

$$\text{Recall} = \frac{\text{Number of correctly words translated}}{\text{Total number of words}} \quad (3)$$

B. Experimental Results for Various Font Sizes

A sample of the images that tested in the system are shown in Fig. 3 below. The system shows an acceptable level of character recognition rates as shown in Fig. 4. The translation rate for some font sizes fluctuates between (50%-100%) because the system recognized either one or two words out of two words existed in the tested images.

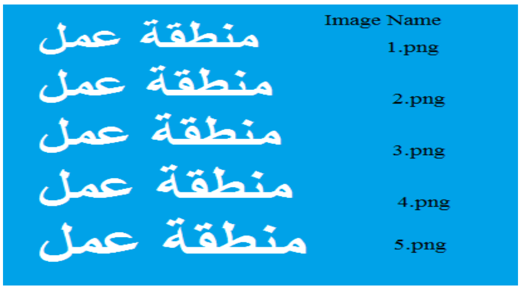


Fig. 3. Samples Images of 5 different font sizes.

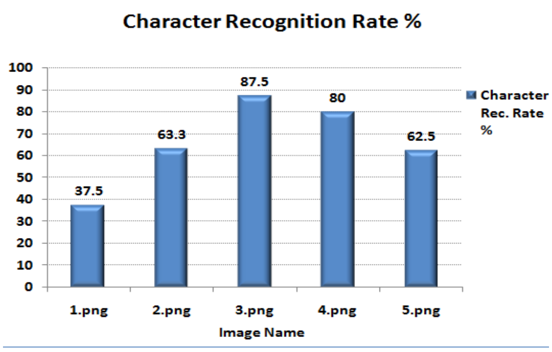
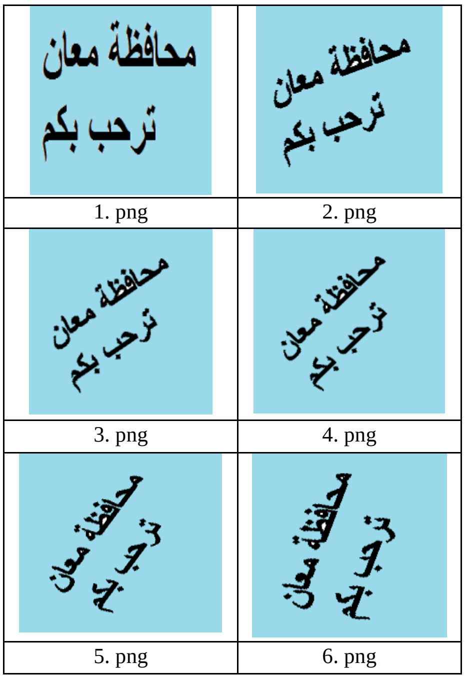


Fig. 4. Characters recognition rate for 5 different font sizes.

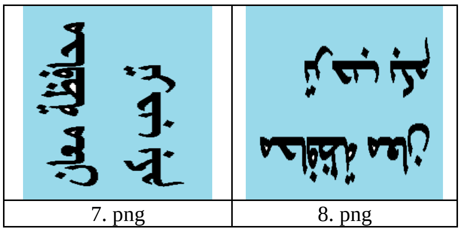
C. Performance of the System for Skew Detection

The system is tested for sample images with an embedded text that skewed within each image. The text in the image has a random angle of skew. The system shows excellent ability to recognize and translate for text of degrees 0^{o}, 90^{o}, 180^{o} and 270^{o}. For these images the character recognition rate is the same for the image of 0^o degree. Images 1. png, 7. png, 8. png and 9. png have the same character recognition. The system also provides a different response for skewed texts in the image with different angle of alignment of the text. The images that tested in the system are shown in Fig. 5 below.



1. png
2. png
3. png
4. png
5. png
6. png

Fig. 5. (a). Sample testing images



7. png
8. png

Fig. 5. (b). Sample testing images.
(The images display the words "Welcome in Maan governarate").

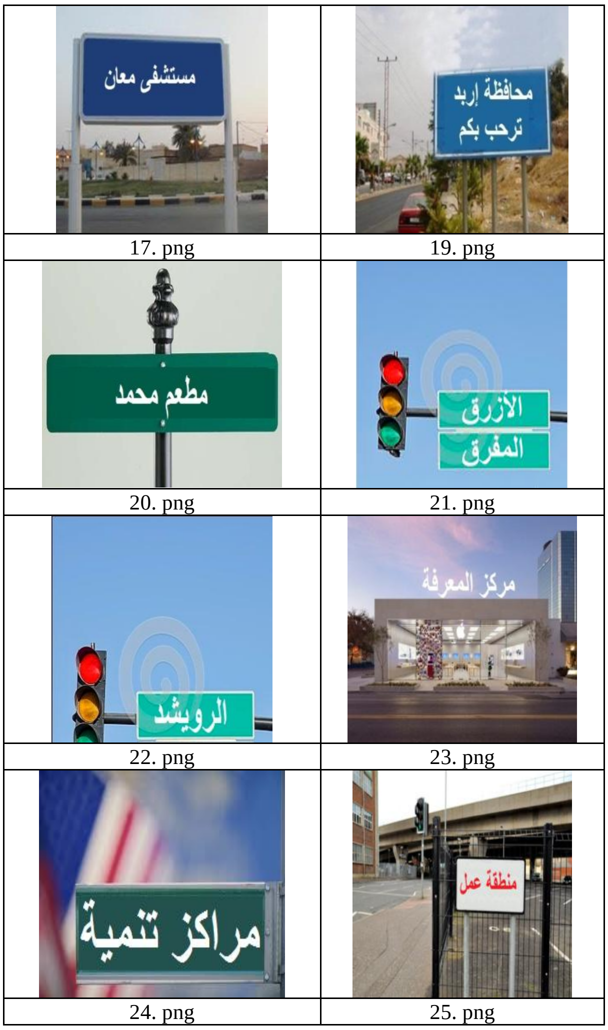
D. Experiment Results for General Image Samples

The system tested against different images with various texts within each image, for convenience images of eight of the sample images are shown in Fig. 6. The system is tested again different types of street sign images and computer based sign images. The system deals with different font size, skewed images and noisy images.

TABLE I: THE RECOGNIZED TEXT VERSUS THE TRANSLATED TEXT FOR INPUT IMAGES

Image Name	Recognized Text	Translated Text
25. png	MNTQHH gGMLL	'Area' 'Work'
24. png	MRZRKRZRZZ tNMYHH	' MRZRKRZRZZ' 'Development'
23. png	MRZRKRZRRA"LMgRFQHH	MRZRKRZRRA"LMgRFQHH
22 .png	ALRWYsdd	'ALRWYSHD'
21. png	AXRZRQQ LMRQQ	'ALARZRAQ' 'ALMAFRAQ'
20. png	MTeTgMM MJOhJOhOhOhMd	'MTeTgMM' 'MOhJOhOhOhMd'
19. png	MOJOhAFTeTeHH A"RBdd tRJh"BB BB	MOJOhAFTeTeHH 'Irbid' 'Welcome' 'you'
17. png	MsQAA MgANN	MsQAA maan

Table I shows the recognized text versus the translated text for images in the experiment. Recognized text is obtained after the words being combined from the recognized characters. The translated texts are the output of the translation system.



17. png
19. png
20. png
21. png
22. png
23. png
24. png
25. png

Fig. 6. Sample 2 of images for testing.

The system shows high rate for character recognition. This will aid the translation subsystem to deal efficiently with image contents Fig. 8 show Rate of Recognition of images

The word recognition rate is shown in Fig. 7 Note that the recognition rate is affected by the number of words in the image. For instance image 22.png contains only the Arabic word ‘الرويشد’. The number of words recognized by the system is ‘1’ which is ‘ALRWYsdd’ and the number of the words in the image also “1” so the rate became 100%. For the image number 23 only one out of two worlds was recognized. So the rate became 50%. The system shows stable word recognition rates for images that contain high number of words to be recognized.

Translation rate depends on the rate of character recognition rate and the content of the word recognized against the dictionary in the system. Fig. 8 shows the rate of translation system

The system shows an acceptable level of recognition and translation levels and the average of character recognition is 81.82%. This affects together with the accuracy of the word segmentation process aided the word recognition subsystem (94.44%). as consequence the word translation also has an acceptable level about (83.33%).

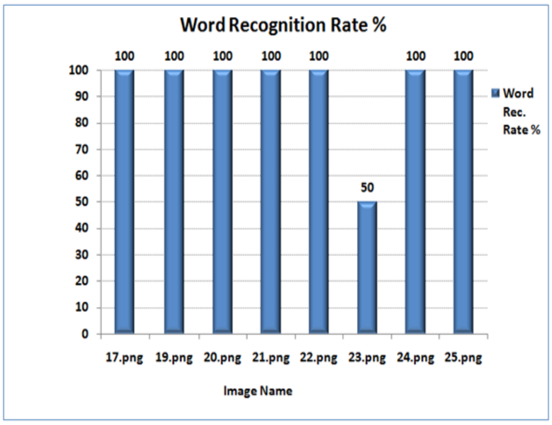


Fig. 7. Rate of word Recognition for images from 17-25.

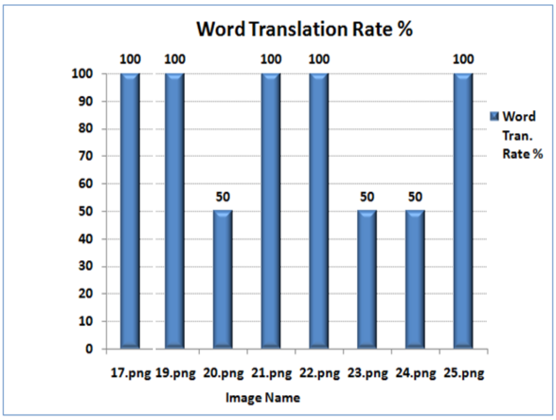


Fig. 8. Rate of Word Translation for images 17-25.

IV. Conclusion

The system is designed and implemented to make it easier for tourism to translate the Arabic text which could be difficult for them to understand in Arab countries during their trips.While the text detection method is robust and accurate, text recognition is good; however the system was able to obtain excellent results in translation Arabic texts into English.

The test images consist of a variety of variations such as text in different color, light, text on dark background, image with blurring, and image with the skew.

We calculate the Recall for characters ,words and translation to evaluate the performance of the developed system and experimental result shows character recognition is 81.82%, the word recognition subsystem (94.44%) and the word translation also has an acceptable level about (83.33%).

References

[1] J. Yan, J. Gao, Y. Zhang, and A. Waibel, "Towards automatic sign translation," in *Proc. First International Conf. on Human Language Technology Research*, 2001. pp. 1-6.

[2] F. Bouslama and H. Kishibe, “Fuzzy logic in the recognition of printed Arabic text,” *IEEE*, pp. 1150-1154, 1999.

[3] A. Hamid and R. Haraty, “A neuro-heuristic approach for segmenting handwritten Arabic text,” in *Proc. ACS/IEEE International Conf.*, 2001, pp. 110–113.

[4] J. Park, T. N. Dinh, and G. Lee. (2008). Binarization of text region based on fuzzy clustering and histogram distribution in signboards. *World Academy of Science, Engineering and Technology*. [Online]. 4(3). pp. 85-90. 2008. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.193.6347

[5] A. F. Mollah, S. Basu, and M. Nasipuri, “Text/Graphics separation and skew correction of text regions of business card images for mobile devices,” *Journal of Computing*, vol. 2, no. 2, 2010.

[6] S. Kumar and S. Gopinath, “Two-Phase data reduction for hough transformation to detect skew,” *International Journal on Computer Science and Engineering (IJCSE)*, vol. 3, no. 10, pp. 3403-3411, 2011.

[7] H. Almohri, J. S. Gray, and H. Alnajjar, "A real-time DSP-based optical character recognition system for isolated Arabic characters using the TI TMS320C6416T," in *Proc. IAJC-IJME International Conf.*, 2008.

Rafeeq Al-Hashemi is the associate professor of Computer Science at Al-Hussein Bin Talal University (AHU) in Jordan. He received his Ph.D. in Computer Science from the University of Technology in 2006. His current research focus on text mining, data mining, designing and implementing educational software and teaching practices in higher education.

Dr. Rafeeq Al-Hashemi is currently the dean of the College of Information Technology at AHU. He took part in many workshops on E-learning. He is a member of Colleges of Computing and Information Society.

Shoroq Almamon Alsharari received master of Computer Science (C.S) degree in 2013 from Middle East University, Jordan and bachelor of Software Engineering degree from Al Hussein Bin Talal University, Jordan. Presently she is working as a teacher. Her areas of interest include Image processing and software engineering.