

Current Arabic (Hindi) Hand Written Numbers Segmentation and Recognition Advance Image Processing and Neural Network

¹Marwan A. Abu-Zanona, ²Bassam M. El-Zaghmouri

¹Department of Computer Science - Imam Muhammad Ibn Saud Islamic - KSA

²Department of Computer Information Systems - Jearsh private University - Amman, Jordan

mascpd@yahoo.com, El_zaghmouri@yahoo.com

ABSTRACT

As a completion of Arabic hand written recognition, the Arabic numbers which are commonly known as “Old-Indian Numbers” are taken place in full text recognition and applications. Hand written Arabic numbers have less complexity than Arabic characters, but such researches are has few interests. The curves shape of these numbers make its characteristics difficult to be recognized by intelligent systems. This paper adapts and implements neural network to takes the features of number segment and recognize it in high reliability and accuracy. The numbers are being segmented by morphological approach. A MATLAB based program is developed to validity testing of the proposed system enhance the accuracy to by 98% over the testing data set.

Keywords: Hand written; Arabic; number recognition; morphology; neural network, LVQ, radon transform.

1. INTRODUCTION.

In the rise of hand written recognition researches in modern image processing and character recognition techniques, the English languages becomes not the single language of researches. Current researches is being interested in complex (non-Latin based) languages, like eastern languages. The main characteristics of Latin languages is that the character shape in almost doesn't depends on the writing style. Figure-1 shows the number two in Arabic written in two different styles.

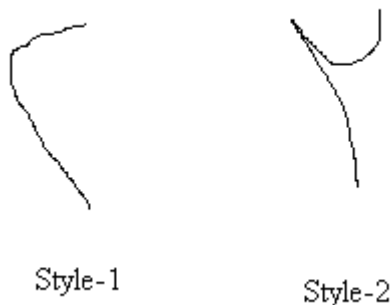


Fig 1: the number “TWO” in Arabic, written handy in two different styles

Even though, the Arabic language has very complex writing structure, starts from the old eastern line drawing format of the character, including the dot marking of most characters, but the Arabic numbers is being written from left to write and has no dot marking for any number. Table-1 shows the basic Arabic numbers – which is commonly known as old Indian numbers - .

Table 1: Arabic numbers and its equivalents in English

Arabic Number	Corresponding English number
٠	0
١	1
٢	2
٣	3
٤	4
٥	5
٦	6
٧	7
٨	8
٩	9

Arabic numbers have a wide importance in the regions that write in Arabic, starting from Arabic world, including Iran, and not ending by Pakistan. The wide use of those characters makes the hand written recognition of these characters to be important, also, it comes from the fact that, the Arabic hand written has no wide researches yet.

Numbers that should be recognized is almost grouped in a multi-digit number as shown in figure-2:

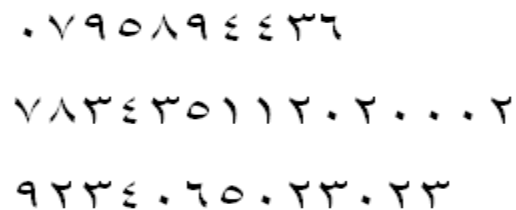


Fig 2: Samples of multi-digit Arabic number series

<http://www.cisjournal.org>

$$B_{\text{thresh}} = \frac{\sum_{i=0}^{r,c} (255 - I(i,j))}{r \cdot c} \quad (1)$$

These numbers could be recognized initially by separating all digits and thus, single digit recognition systems will be applied. But the fact that, these numbers are written by hand in different ways, makes separating of them to be really not easy. Figure-3 demonstrates samples of grouping number digits in hand written. It clear from that, the connectivity of some digits to another makes the segmentation somewhat difficult.



Hand Written Sample



As it is Printable

Fig 3: Interconnected hand written Arabic number digits (above) and printing style of those numbers (bellow)

This paper implement a morphological approach to separate those numbers in cooperation with binary projections transform (Radon Transform).

This paper implements a rigid approach that recognizes 98% percent of Arabic hand written numbers where those numbers may written in worst case. The steps of the proposed and implemented algorithm are:

- Apply Preprocessing of converting image into binary in addition to dilation and filling.
- Apply Radon transform in (-90°) degree.
- Calculate the Projection threshold.
- Generate all points those considered to be separation between numbers.
- Segment all numbers.
- Extracting features from number segments.
- Apply Linear Vector Quantizer (LVQ) neural network to recognize each individual number digit.
- Encode the total recognized number.

First of all, the input image is almost RGB, it should be converted to gray and the gray image is being converted to binary by using a binary threshold in equation-1:

2. RADON TRANSFORM

The radon transform form the few of image processing, is that, calculating the projection values of the image pixels in a specified angle direction. Figure-4 shows the basic radon transform of an image that contains a solid square, located in angle 45°. Where, the transform is obtained in projection angle of 90° degree.



a) Original Shape



b) Projection

Fig 4: Projection transformation (b) of a solid square object (a)

The radon transform computes the number of pixels that the light narrow passes through in the specified angle of projection θ . Thus, the more projection value means the more pixels intensity.

The projection principal is applied to hand written Arabic numbers in order to get the areas between numbers. Those areas assumed to be the minimal pixel intensity, thus, the minimal projection value.

Figure-5.a shows sample number series written handy, and displays the separation areas between those numbers. Where, figure-5.b illustrated the radon transform of that number in angle of (-90).

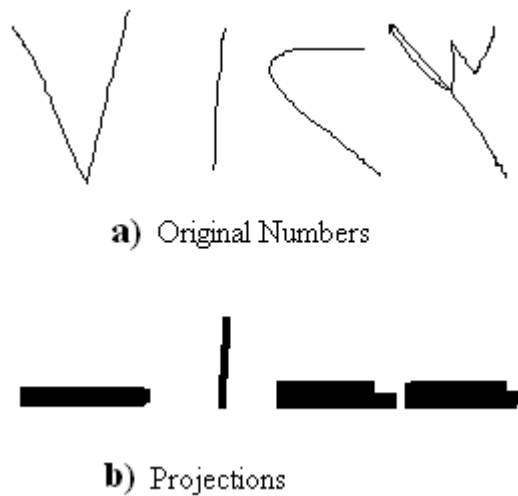
<http://www.cisjournal.org>

Fig 5: Illustration of projection transformation of hand written 4-digits Arabic number. The separations areas are clear between the projections are sets.

It clear form that figure, the areas that located between number digits are two types; the first – which is marked in red – is empty areas where no contact between number digits occurs. The projections of such areas ideally have 0 value. But actually, it may have a value that is near to zero by the cause of noise. The second areas – marked in blue – got connectivity of two numbers. Such areas have projection value more than zero.

The projection has a threshold that could be used to discriminate between the area that located between number digits and the numbers itself. That threshold has maximum value of intersection between two lines. The threshold value is adaptive and calculated from the equation-2:

$$Pth = \sqrt{Lx^2 + Ly^2} \quad (2)$$

Where Lx and Ly is the line width and height respectively.

3. MORPHOLOGICAL OPERATIONS

The morphological operations are divided into two phases; the first one is preprocessing before applying the radon transformation. It consists of dilation and filling of the objects in the image. The added value of the dilation and filling is that the pixels of the number may has discontinuity in some pixels gotten form imaging of the number's document. The dilation is connecting such continuities, where the filling makes no holes in the solid object. The object filling is important for two reasons; the fact that, radon transform will be applied, makes the holes in objects as weakness points of the projections, by the mean of that it results wrong projections value. Figure-6 shows the

projection value of a specified object, (a) the object has wholes in its area, and (b) the object is filled.

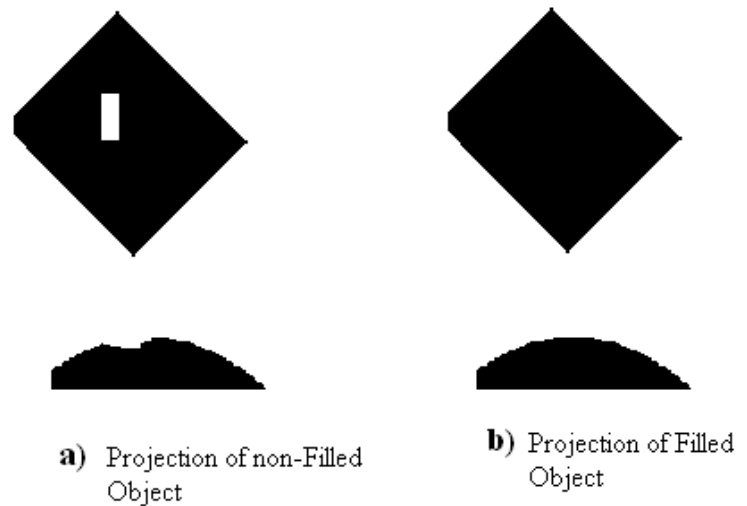


Fig 6: Projection of rotated square that has oblong clearance inside it (a) in comparison with the same square object results when it placed to filling process (b).

Also, the filled object has minimal features and static to be used in the following recognition phase. It represents the unique drawing of the number only, whereas, the hole add many extra features that represents noise and uncertainties.

The second phase of implementing morphological operations placed after radon transform computations; the segmentation phase. The connected areas between numbers that gotten by projection calculations by using of radon transformation will be removed by placing 0 value in all pixels of that area, thus, the resulted number digits all will be separated. Figure-7 shows three connected numbers (a) and the removal of detected area connection between these numbers (b).

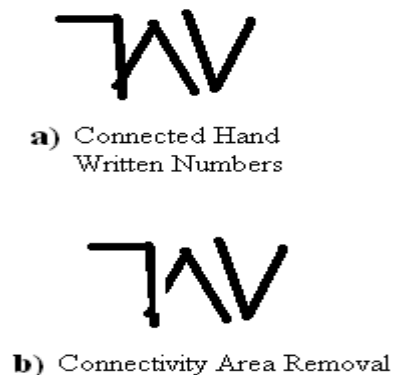


Fig 7: Sample connectivity removal (b) between a set of Arabic number digits (a)

<http://www.cisjournal.org>

In segmentation, normal black and white labeling is being used, to label each object in the image. The properties of each labeled object are being calculated basically. The basic features are normally:

- Centroid, which is the center of the object.
- Bounding Box.

The segmented objects may have noise represented such forms (i.e. dots in the page of numbers o). The removal of that noise is being done by checking the area of all objects and comparing it with threshold value of area. The area is calculated using equation-3, while the threshold area is calculated adaptively by implementing equation-4.

$$\text{Area} = \text{Width} \cdot \text{Height} \quad (3)$$

$$\text{Ath} = \frac{\sum_{k=0}^n A_k}{3n} \quad (4)$$

Where Width and Height is gotten from the bounding box, n is the number of generated objects on the image, A_k is the area of K^{th} object in the image.

4. EXTRATING FEATURES

Firstly, all number digits have been parsed / scanned and isolated from the original image that contains a series of numbers, a feature extracting algorithm should be implemented in order to get unique features those have both, minimal data size, and maximum recognize ability.

This paper implements spatial dividing of the number image, where all images are being resized to 32x24 pixels. Then, the image is divided to 12 segment of size 8x8 of each one, as shown in figure-8. The, 5 statistical features will be calculated for each segment resulting of 60 feature for each image segment. Those features are:

- The ratio of white pixels to black pixels in binary converted image of the segment. This is calculated as equation-5.
- The ratio between the two farthest pixels with respect to diagonal of the segment. This is calculated as equation-6.
- The normalized average of the spatial segment. That is calculated as equation-7.
- The variance of the farthest two vertical pixels divided into the variance of the farthest two horizontal pixels. This is calculated as equations-8 and equation-9.
- The average value of the line that is orthogonal to the main diagonal of the special segment. This is calculated as equation-10

$$Rwb = \frac{\sum_{i=0}^r \sum_{j=0}^c I(i,j)}{1 - \sum_{i=0}^r \sum_{j=0}^c I(i,j)} \quad (5)$$

$$Rfbd = \frac{I(ior,jor) - I(jol,jol)}{\sqrt{(rs^2 - cs^2)}} \quad (6)$$

$$Mn = \frac{\sum_{i=0}^r \sum_{j=0}^c I(i,j)}{(rs \times cs)} \quad (7)$$

$$\text{Var}(x) = \int (x - u)^2 \cdot f(x) dx \quad (8)$$

$$u = \int x f(x) dx \quad (9)$$

$$Mn = \frac{\sum_{i=0}^{rs} I(r-i, c-i)}{\sqrt{(rs^2 - cs^2)}} \quad (10)$$

Where "u" is defined in equation-9, "rs" and "cs" are the character's segment number of row and number of columns respectively, $I(i,j)$ is a pixel of the image. "jor", "jor", "jol, and "jol, are the coordination's of the farthest pixels of the character object in the image.

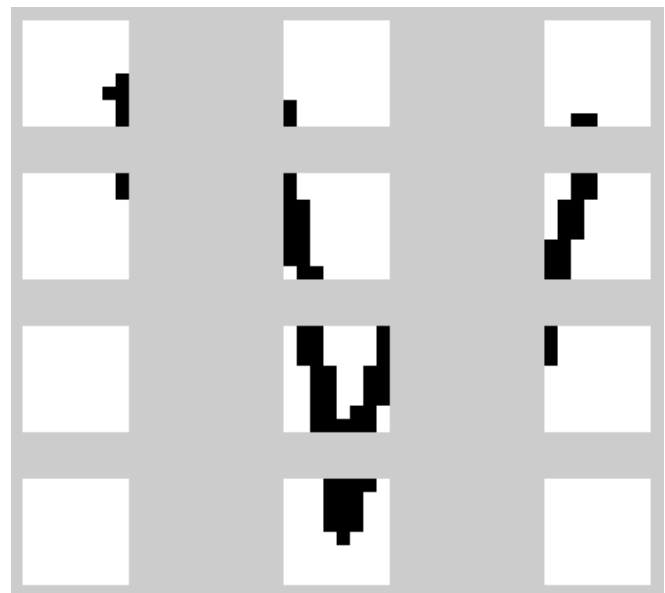


Fig 8: Image of sample character, it's segmented to 12 of 8x8pixel segments.

5. RECOGNITION

Lastly, the features that had been extracted from character segments will be input to Linear Vector Quantizer Neural Network. The output of the network will be the recognized character itself.

The use of extracted features will minimize the input data size from one complete segment of size 32x24 pixel (768 byte) to 60 extracted features. This minimizing of input data size optimizes the use of memory in addition to speed up the processing (it represents minimal computational power). Also, this data optimization represents optimized input for the neural network, whereas, the minimal inputs lead to minimal structure and minimal complexity. But the larger input numbers leads to larger structure and more complexity. In neural networks, the minimal complexity ensures better training and thus, better recognition results.

Linear Vector Quantizer N Net is suitable for limited data or limited outputs. The in this paper, the output consists of one node its output is the recognized number itself (i.e. 0, 1, 3 ...). LVQ in such data ensures fast processing in run mode of the network.

The neural network design and operation consists normally from two phases; training phase and simulation / run phase. In simulation or run phase, the network is assumed to start input the features and generates the recognized result. Where, in training phase, the network structure was built by programmer's design in cooperation with self training algorithm. The training phase stands from historical data and results a ready neural network for running / simulation.

6. RESULTS

The data set that is being used in order to test and measure the proposed system's performance consists of 100 different set of multi-digit numbers. Each number consists of 1 to 15 digits. This data set is being recorded manually by 20 different persons. A MATLAB program was been developed to implement this paper's algorithm.

Table-2 shows the results of recognition for each single number. Where, table-3 shows the Mean Absolute Percentage Error (MAPE) for the whole data – including the segmentation process – of validation test. In table-3, the hole data was been tested separately at first. Then, 2% noise was added and also, 4% noise was also added. The recognition result shows that, the proposed and implemented algorithm has immunity for noise.

Table 2: Recognition results for each character

Input separated number	Testing Accuracy of the Character
٠	98%
١	100%
٢	97%
٣	98%
٤	100%
٥	99%
٦	99%
٧	100%
٨	100%
٩	98%

Table 3: MAPE for different testing datasets

Trial	MAPE
Without noise	99%
2% noise	99%
4% noise	98%

REFERENCES

- [1] H. Boubaker, M. Kherallah, and A. M. Alimi. New algorithm of straight or curved baseline detection for short arabic handwritten writing. In 10th Inter. Conf. on Document Analysis and Recognition, 2009.
- [2] A. Elbaati, M. Kherallah, A. M. Alimi, and A. Ennaji. Restoration of the temporal order of the off-line handwriting using genetic algorithm. In 4th International Conference on Computer Science Practice in Arabic (CSPA), 2008.
- [3] ligne par reconstruction de l'ordre du trace en vue de l'indexation de documents d'archives. PhD thesis, INSA of Rennes, 2007.
- [4] J. Sadri, et.al , "Application of Support Vector Machines for recognition of handwritten Arabic/Persian digits," Proceeding of the Second Conference on Machine Vision and Image Processing & Applications (MVIP), Vol. 1, Feb. 2003, Iran, pp. 300-307.
- [5] S.V. Rajashekaradhyia et.al "Efficient Zone Feature Extraction Algorithm for Handwritten Numerals Recognition of Four South Indian Scripts", Journal of Theoretical and Applied Information Technology 2008

- [6] M. Hanmandlu et.al , “Input fuzzy for the recognition of handwritten Hindi numeral:”a, International Conference on Informational Technology 2007
- [7] Al-Taani Ahmad et.al, “Recognition of On-line Handwritten Arabic Digits Using Structural Features and Transition Network” Informatica 2008.