International Journal of the Physical Sciences Vol. 6(13), pp. 3198–3204, 4 July, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.298
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Farsi digit recognition via features extraction

Mehdi Alirezanejad and Rasul Enayatifar*

Islamic Azad University, Firuzkooh Branch, Firuzkooh, Iran.

Accepted 14 March, 2011

In this paper, a new method is proposed to extract the features of a one-number Persian image in which for the final verification of the extracted features, a three-layer neural network (mesh) of Perceptron has been utilized. The method is capable of extracting some ideal features from a one-number image that are stable against rotation, movement, size change and noise. The method is examined on a database of 60000 discredited numbers, from which 40000 numbers were used in the training stage and 20000 ones were used for the experiment. The recognition percentage of 92.7% shows the great efficiency of the proposed method.

Key words: Features extraction, recognition of Persian numbers, perceptron neural network, standard deviance, average angle.

INTRODUCTION

Recognition of Persian handwritten numbers has many applications such as the recognition of a zip code on an envelope, the amount of money written on a check, the numbers written in examination forms (Khosravi and Kabir, 2007), etc. The main problems are dependence on the style of handwriting, image quality and handwriting size. After lots of investigations in the field, various methods were proposed to solve these problems (Khosravi and Kabir, 2007; Dehghan and Faez, 1997; Dehghan et al., 2000; Mozafari and Safabakhsh, 2004; Zhouchen and LiangWan, 2007; Mozaffari et al., 2004; Ghoochani and Kheirkhah, 2006; Goraine et al., 2006; Thierry et al., 2007; Alessandro et al., 2004).

Most of these methods were proposed for English, Japanese and Chinese numbers, and only few of them were for Persian numbers such as Markov hidden model (Dehghan et al., 2000), Neural networks (Mozafari and Safabakhsh, 2004), Constitutional features Dehghan et al., 2000; Mozafari and Safabakhsh, 2004; Zhouchen and LiangWan, 2007; Mozaffari et al., 2004; Ghoochani and Kheirkhah, 2006; Goraine et al., 2006; Thierry et al., 2007; Alessandro et al., 2004) and Momentums (Dehghan and Faez, 1997).

The proposed method in this paper is based on the extraction of new features of a number-narrowed image, and the final recognition is done through neural network.

The selection of features for proper classification is of great importance, because wrong selection of features leads to incorrect recognition of the classifier. Standard deviance and average angle, which are used to extract the features of image, is capable of extracting some new and appropriate features.

In this paper, image of each digit will be dividing in 24 parts. After that two features will extract from each part of digit's picture. As result, we obtain 48 features for final recognition via neural network.

Subsequently, the Persian numbers are introduced, and then the preprocessing stages are done. Later, the proposed method is introduced in details. Finally, the experimental results are explained and the conclusions are discussed.

Introduction of Persian numbers

In order to be familiar with Persian numbers, each English number is shown with its Persian equivalent (Figure 1).

Preprocessing

Preprocessing consists of two stages, firstly the remover of discrete points of the number and the transformation of the inlet image to a binary one, and secondly narrowing the inlet number.

*Corresponding author. E-mail: r.enayatifar@gmail.com

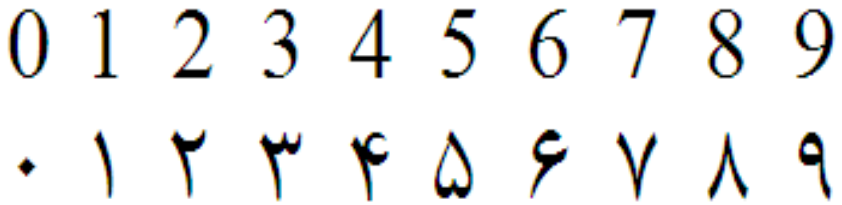


Figure 1. English numbers and their Persian equivalents.

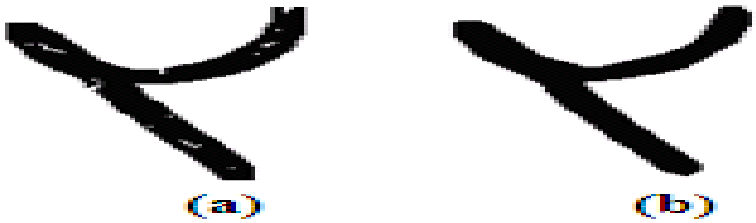


Figure 2 (a). The inlet image with discrete points, **(b)** the result of closing operator firstly with windows w_1 and then with w_2.

In order to have a narrowing algorithm with high efficiency, a binary image with minimum discrete points is needed.

To fill in the discrete points (Figure 2 (a)), morphological operators are utilized. Primarily, closing operator is exerted onto the image firstly by using window W_1 and then with W_2, which reduces the discrete points of the image. Morphological operator is:

$$W_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad W_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$Clo\sin g(I, W_1) = Erosion(Dilation(I, W_1), W_1)$$
$$Clo\sin g(I, W_2) = Erosion(Dilation(I, W_2), W_2) \quad (1)$$

Figure 2 (b) demonstrates how suitable morphological operators can eliminate the discrete points of the image.

Narrowing

What is meant by narrowing is to obtain the body of the number without eliminating fine details. The algorithm in Ghoochani and Kheirkhah (2006) is utilized to perform the narrowing process. In this algorithm, it is assumed that the points on the body are 1 and the points on the background are 0. This algorithm consists of two main parts, each one including four steps. In each step, this algorithm marks the point(s) to be eliminated, and eliminated them at the end of each part. This continues as long as there are no candidates to for the elimination.

If the octagonal neighborhood of P_1 is considered as in Table 1, $N(P_1)$ is the number of neighbors with the value 1 and $S(P_1)$ is the number of changes from 0 to 1 in the series $P_2, P_9, P_8, ..., P_3, P_2$, the elimination of P_1 in the first part is shown in Equation (2) and in the second part is shown in Equation (3).

$$\begin{cases} 2 \le N(P_1) \le 6 \\ S(P_1) = 1 \\ P_2 . P_4 . P_6 = 0 \\ P_4 . P_6 . P_8 = 0 \end{cases} \quad (2)$$

The elimination of P_1 in the first part:

$$\begin{cases} 2 \le N(P_1) \le 6 \\ S(P_1) = 1 \\ P_2 . P_4 . P_8 = 0 \\ P_2 . P_6 . P_8 = 0 \end{cases} \quad (3)$$

Table 1. Octagonal neighborhood of P_1.

P9	P2	P3
P8	P1	P4
P7	P6	P5

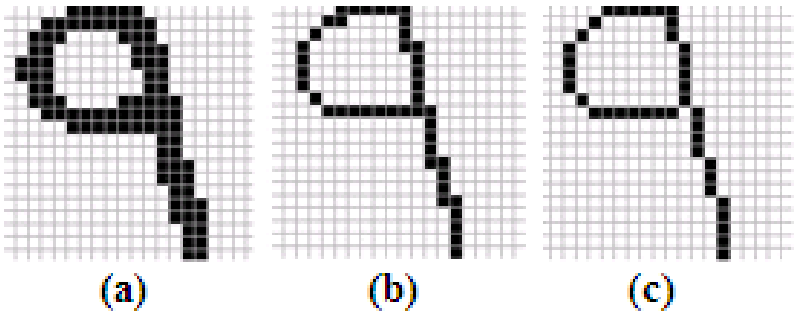


Figure 3 (a). The inlet image (b) The results of applying Equation (2) and (3) to the inlet image (c) The result of applying Equation (4) to section (b).

The elimination of P_1 in the second part:

It is necessary to know that all four conditions are to be satisfied to have a pixel marked for later elimination. Equation (2) is specifically for the marginal pixels of the binary zone, and Equation (3) is used for other pixels. The elimination of marked pixels at the end of each part is done simultaneously. These two parts are repeated as long as no more pixels are left to be omitted. Figure 3 (a) is an example of a number and Figure 3 (b) shows the results of the first and second stages of this algorithm.

In this method, to extract better features, narrowed inlet numbers are needed so that there is a connection with the pixel as long as possible, while for a number like nine, there is one pixel in some rows especially in the bottom part as seen in the narrowing algorithm shown in Figure 3 (b). In order to eliminate such pixels, which do not provide the user with any kind of data, a new section is added to the narrowing algorithm as in Equation (4), the results of Figure 3 (b) is seen in Figure 3 (c).

$$img(P_1)=0 \ \ if \begin{cases} img(P_2)=1 \ \ AND \ img(P_4)=1 \\ OR \\ img(P_4)=1 \ \ AND \ img(P_6)=1 \\ OR \\ img(P_6)=1 \ \ AND \ img(P_8)=1 \\ OR \\ img(P_8)=1 \ \ AND \ img(P_2)=1 \end{cases} \quad (4)$$

In Equation (4), the neighborhoods of Table 1 are utilized.

The proposed method

After pre-processing stage, it is time to describe proposed method.
Proposed method divides to 3 parts:

Frame finding
Classification
Feature extraction
Recognition using neural network

Frame finding

After narrowing the digits, the frame of narrowed digits is required in order of extracting features and executing classification algorithms. This frame includes just the important data of the input digit and some parts of the image that have unnecessary data is eliminated.

For finding narrowed digit frame, the coordinates of pixels with the quantity of one (black) in the farthest left, right, up and down side are found (Figure 4).

Classification

After the end of pre-processing and having the result of that, it is time to classify the narrowed digits to two category that are digits 1, 2, 3, 4, 6 and 9 that have a 1-like base and the other digits that are 0, 5, 7 and 8. To classify the narrowed digits to this two categories, first we need to have the height (H) of input digit. This could be achieved by having the coordinates of farthest pixels with

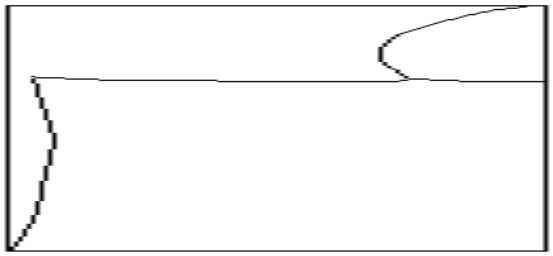


Figure 4. Digit's frame.

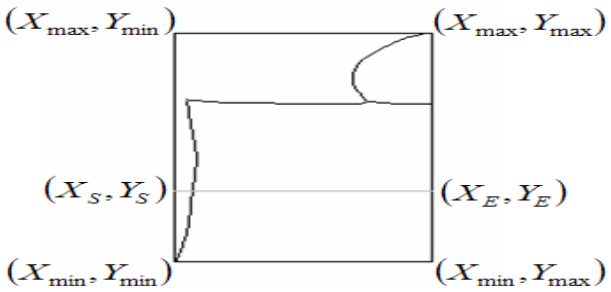


Figure 5. Picture of number 4.

the quantity of one in the up and down side from the previous step. Then, a line parallel to the horizontal edges of the frame in the proposed coordination system of Equation 5 is drawn (Figure 5).

$$H = X_{max} - X_{min}$$
$$\begin{cases} X_S = X_E = X_{min} + \dfrac{30}{100} \times H \\ Y_S = Y_{min} \\ Y_E = Y_{max} \end{cases} \quad (5)$$

After drawing the mentioned line, the digits are categorized due to their intersection with this line. If the line intersects with the digit in just one point it implies that the digit is of the category one. Otherwise the digit belongs to the category two. Since a new section was added to the narrowing algorithm, categorizing algorithm is working properly.

The framing algorithm divides the frame obtained from previous steps to 6 rows and 4 columns. In fact a 6*4 table is set for the frame according to its dimension (Figure 6).

In this work the image of input digit is not normalized there is a possibility that in the framing process the last row (column) of frame is not the same as the others. Subsequently, it would be noted that this would not affect the accuracy of the algorithm.

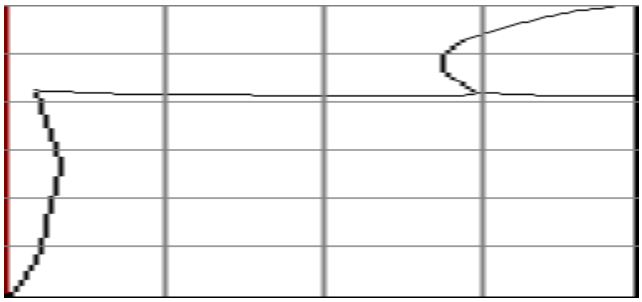


Figure 6. Framing of narrowing digit in 6*4 frame.

Feature extraction

After framing the image, 24 frame are obtained. Standard deviation and average angle are extracted as features of these frames.

Standard deviation

Standard deviation is a very efficient feature in different sciences. It is used to extract the following formula in any point.

$$\sigma = \frac{1}{n}\sum_{i=1}^{n}(X - X_i) \quad (6)$$

In the aforeseen formula, X stands for the average position of the pixels of the narrowed digit of every frame and X_i indicates the position of the pixel number I in every frame.

Average angle

One of the features of the effectiveness of the method in recognizing the hand-written digits is the average angle feature. For extracting this feature the framing used in the last step is used. In this method, the angle of every frame according to the horizontal level is extracted and then the average of these angles is calculated using formula 7. This average would be considered as the average angle feature.

$$\alpha_i = \frac{1}{n}\sum_{j=1}^{n}\theta_j \quad (7)$$

In the aforeseen equation, α_i is the average angle feature of the i frame. n indicates the number of the pixels of the image of the digit that lie in the frame i. θ_i is the angle of any pixel in the frame i that is calculated according to the horizontal level.

Table 2. The rate and the time of recognition based on the number of neurons in the middle layer.

The number neurons in the hidden layer	3	10	20	50
The rate of recognition (%)	61.28	75.19	83.27	91.71
The recognition time (seconds)	0.03	0.04	0.1	0.236

Table 3. The sensitivity of this method to noise.

Noise type	Noise rate (%)	Recognition percentage
Gaussian	2	84.42
	5	80.97
	8	76.03
	10	70.56
Pepper-salt	2	90.21
	5	88.11
	8	85.79
	10	84.34

The reasons of choosing the average angle feature can be investigated in the three different categories.

Stability against rotation: since the angle of the pixel in the frame is investigated this implies that the rotation cannot change the digit.
Stability against transform: transforming the image does not affect the average angle feature.
Stability against resizing: resizing the image cannot change the direction of the pixels so it cannot affect the angle of the pixels.

Recognition using neural network

After extracting the features introduced in Section 2.4 from the figure, using a multi-layer Perceptron neural network and a feed-forward xz algorithm, the final recognition of the number is performed. To determine the number of medial neurons, the network is simulation according to the number of different neurons in the medial layer, the results are shown in Table 2. As seen in this table, as the number of neurons in the middle layer increases, not only does the recognition percentage increase, but also the recognition times increases. The irregular increase of neurons in the middle layer leads to a progressive increase in the recognition time, this happens while there is no considerable change in the recognition quality.

The number of the neurons is pleasant as long as the recognition time does not increase progressively and the recognition percentage is not constant. According to the numbers recorded in Table 2, it is concluded that 16 neurons in the hidden layer provides appropriate conditions for recognition process.

Stability rate of feature arrow

Here, the stability of the method is studied against noise, rotation, image size and movement.

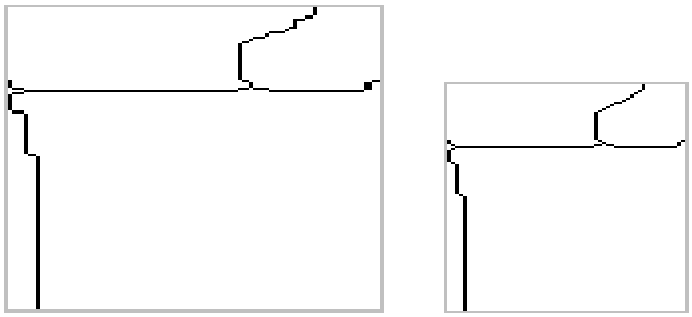
Stability against noise

The type of the extracted features in this work does not show any dependency on noise. To study the sensitivity of this method to noise, Gaussian and Pepper-Salt noises were exerted on the image, and the recognition results are shown in Table 3.

The recognition percentages presented in Table 3 show the strong resistance of the proposed method against noise conditions.

Change in input image size

On the contrary, of the some of the propsed algorithms (Ghoochani and Kheirkhah, 2006), one of the advantages of the extraction algorithm is the high stability of it against resizing the input image. Since two features of standards deviation and average angle are extracted taking into account the black pixels of the frame, all pixels of the frame (black and white), all black pixels of the image and all pixels of the image (image size), changing the size of the input image results in change in all these features so that the percentage extracted for these features in every frame stands the same. The image of digit 4 in two different positions (Figures 7 (a) and (b)) that are similar in shape but different in size are shown. These images are used to evaluate the stability of the algorithm against resizing for one of the extracting algorithms.



(a) **(b)**

Figure 7 (a). Picture number 4 in size (66*39), **(b)** Picture number 4 in size (116*64).

Table 4. Percentage of average angle for 7 (a).

0	0	40.38	21.42
43.09	12.43	36.54	8.39
93.81	0	0	0
91.22	0	0	0
90.36	0	0	0
90.36	0	0	0

Table 5. Percentage of average angle for 7 (b).

0	0	40.38	21.42
43.09	12.43	36.54	8.39
93.81	0	0	0
91.22	0	0	0
90.36	0	0	0
90.36	0	0	0

According to the Tables 4 and 5, Average angle have stability against change size.

CONCLUSIONS

In this paper, a new method has been proposed to extract image features through water filling algorithm. The number-narrowed image is the input of water filling method. And the extracted features are given to a Perceptron neural network for the final recognition. The relatively high stability of the extracted features against noise, rotation, image size changes and movement is the most important advantage of the proposed.

In this work, a database (Khosravi, 2005) of 60000 Persian numbers was used, out of which 40000 were used for education purposes, and 20000 were used for the experiment. Recognition percentage of the experiment was found to be 92.7%.

REFERENCES

Alessandro V, Samy B, Horst B (2004). “Offline Recognition of Unconstrained Handwritten Texts Using HMMs and Statistical Language Models”, IEEE Trans. Pattern Anal., 26(6): 709-720

Dehghan M, Faez K (1997). “Farsi Handwritten character Recognition with Moment Invariants,” Proceeding of 13th International Conference on Digital Signal Processing, 2: 507-510.

Dehghan M, Faez K, Ahmadi M, Shridhar M (2000). “Off-Line Unconstrained Farsi Handwritten Recognition Using Fuzzy Vector Quantization and Hidden Markov Word Models,” Proceeding of 15th International Conference on Pattern Recognition, 2: 351-354

Ghoochani SR, Kheirkhah AR (2006). “Letter recognition system by framing in feature extraction and based on potential learning algorithm and neural network”, The 4th Conference on Machine Vision and Image Processing of Iran.

Goraine H, Usheer M, Al-Emami S (2006). “Off-Line Arabic Character Recognition,” IEEE Computer, 25(7): 71-74.

Khosravi H (2005). “Persian number and letter recognition in public university entrance exam forms”, MS thesis of Electronic Engineering, Tarbiat Modares University, (2005).

Khosravi H, Kabir E (2007). “Introduction a Very Large Dataset of Handwritten Farsi Digits and a Study on Their Varieties,” Pattern Recognition Lett., 28: 1133-1141

Mozafari S, Safabakhsh R (2004). “Discrete Persian number and letter recognition based on the neural network of radial functions and fractal features”, The 3rd Conference on Machine Vision and Image Processing of Iran, pp.108-114.

Mozaffari S, Faez K, Rashidy KH (2004). “Recognition Of Isolate Handwritten Farsi/Arabic Alphanumeric Using Fractal Codes,”- IEEE Proceeding of Southwest Symposium on Image Analysis and Interpretation, pp.104-108.

Thierry A, Sanparith M, Patrick G (2007). “Online Handwritten Shape Recognition Using Segmental Hidden Markov Models” IEEE Trans. Pattern Anal., 29(2): 205-217.

Zhouchen L, Liang W (2007). “Style-preserving English handwriting synthesis”, Pattern Recognition, pp. 2097 – 2109.