International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:3, 2007

A New Pattern for Handwritten Persian/Arabic Digit Recognition

A. Harifi and A. Aghagolzadeh [1]

Faculty of Electrical Engineering, University of Tabriz
Tabriz, 51664, Iran

aghagol@tabrizu.ac.ir [1]

***Abstract*— The main problem for recognition of handwritten Persian digits using Neural Network is to extract an appropriate feature vector from image matrix. In this research an asymmetrical segmentation pattern is proposed to obtain the feature vector. This pattern can be adjusted as an optimum model thanks to its one degree of freedom as a control point. Since any chosen algorithm depends on digit identity, a Neural Network is used to prevail over this dependence. Inputs of this Network are the moment of inertia and the center of gravity which do not depend on digit identity. Recognizing the digit is carried out using another Neural Network. Simulation results indicate the high recognition rate of 97.6% for new introduced pattern in comparison to the previous models for recognition of digits.

***Keywords*— Pattern recognition – Persian digits – Neural Network

I. INTRODUCTION

One of the major problems in handwritten digit recognition is the difference between the shapes of each digit in different handwritings. Therefore, the most active current research is using intelligent algorithms such as neural network. For recognition of digits by Neural Network, a suitable feature vector shall be extracted via the image matrix of the digits. A feature vector should have the following conditions:

1. Feature vector of two different digits should contain major differences between them as the network could conclude these differences.
2. Feature vector of a specified digit written by different handwritings should be the same, as the network could not find a meaningful difference among them.

There are many different methods to extract feature vector. The coefficients of discrete wavelet transform have been used as a feature vector by Mowlaei and et al. [1]. Shirali and et al. has extracted a suitable feature vector by using shadow coding method and 32 different parts [2]. Another method for extracting feature vector has been proposed by Hosseini and Bouzerdoum by scanning the image matrix via 11 horizontal and vertical lines [3].

In this research, the idea of segmentation pattern for feature vector extraction from the image matrix has been investigated. Then for improving the feature vector characteristics, an asymmetrical segmentation pattern is introduced. Due to one degree of freedom (the control point) in the introduced asymmetrical pattern, an optimum point can be found by defining a cost function. Based on this procedure, a method of finding the optimum control point is given. Since this algorithm depends on the digit identity, for enhancing the classification results, a neural network optimizer is proposed. In this paper, all the above introduced procedures have been unified and implemented by the neural network. Part of the simulation results and the objective measure are given in this short presentation. The simulation results show very good performance and improvement with respect to other reported works.

II. SEGMENTATION PATTERN

The idea of segmentation pattern was taken from the so-called 7-Segment display. All Latin digits could be displayed using 7-Segment display system. Since Persian digits have oblique sections, there is a necessity for more segments. Furthermore some digits like 0 and 5 resemble each other and this makes the situation somehow laborious.

In this paper a 12-Segment pattern was used. Figure 1 depicts the pattern. Although this pattern doesn't look like to be an appropriate segmentation pattern, it is utilizable.

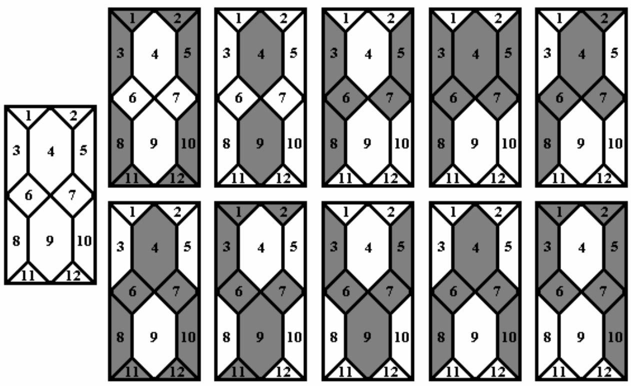

Figure 1: Segmentation pattern and mapping of Persian digits from 0 to 9 on it

To extract the feature vector out of an image using this pattern, a similar method "shadow coding" is used. First the clean out of the area surrounding the digit is carried out to obtain the bounding box. Then the 12-Segment pattern is mapped on the image matrix (figure 2). The ratio of the number of the enclosed pixels in one segment to the whole number of the pixels inscribed in that segment can be calculated for every 12 segments which are used to calculate the feature vector. Since the number of the enclosed pixels in each segment is divided by the whole area of that segment, the feature vector is independent from the scale, precisely speaking, to the size of the segment.

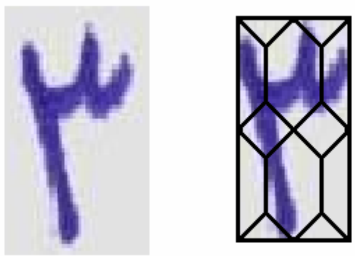

Figure 2: Feature vector extraction out of an image matrix for digit 3

III. Asymmetrical Segmentation Pattern

Elongated and skewed handwritten digits can extensively affect the feature vector. Therefore an asymmetrical segmentation pattern presentation is very helpful. With some modification to the current pattern, an appropriate asymmetrical segmentation pattern can be obtained (figure 3).

This pattern has just one degree of freedom as it is illustrated in figure 3(point 1). Hence the characteristics of the feature vector can be conspicuously improved by optimizing the location of this control point.

The extraction of the feature vector is the same as the previous pattern.

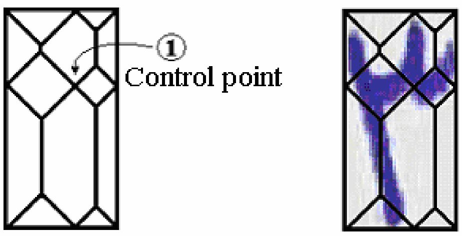

Figure 3: Asymmetrical segmentation pattern

IV. Finding the Optimum Control Point

To locate the optimum control point, the pattern is mapped on the image matrix. Then by moving the control point in the image space, feature vector is found for every location. If F is a feature vector, then a cost function can be defined:

Cost function = $C^T F$

where C is a weight vector. Any location that maximizes this function is the optimum location of the control point. By inspection of the sample digits, it's recognized that the optimum control point is different for every digit. Hence the different weight vectors are obtained for each digit. Figure 1 helps to find these weight vectors:

$$Digit\ 0:\quad C = [1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1]^T$$
$$Digit\ 1:\quad C = [0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0]^T$$
$$Digit\ 2:\quad C = [0\ \ 0\ \ 1\ \ -1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0]^T$$
$$Digit\ 3:\quad C = [0\ \ 0\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0]^T$$
$$Digit\ 4:\quad C = [0\ \ 1\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0]^T$$
$$Digit\ 5:\quad C = [0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1]^T$$
$$Digit\ 6:\quad C = [1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 1\ \ 0]^T$$
$$Digit\ 7:\quad C = [0\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0]^T$$
$$Digit\ 8:\quad C = [0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0]^T$$
$$Digit\ 9:\quad C = [1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0]^T$$

The reasoning behind the using of -1 as a coefficient in the weight vector of digit 2 is the significant difference between 2 and 3. This means that the optimum control point for digit 2 must be calculated in a manner that a few pixels reside in the segment 4. This method makes the recognition of the digit 2 easy.

V. Optimizer Neural Network

According to the previous sections, the method of locating the optimum control point for every digit is different; since the above mentioned method is dependent to the digit identity a method should be found which is not dependent on the digit identity. Hence a Neural Network model was chosen. First some handwritten samples were gathered. Then using the above mentioned method and the digit identity, an optimum control point is found. Finding the proper input for Neural Network was of the paramount importance that preserves relevance with the digit. In this research the center of gravity and the moment of inertia was considered to be the best choice.

VI. Proposed Algorithm

According to the explanations presented in the previous sections, the whole digit recognition process consists of two parts:

1-Offline Neural Networks training

2-Using the trained Neural Networks for online recognition of a digit

1. Neural Networks training:

A MLP (Multi Layer Perceptron) network consisting of 4 input neurons, 10 neurons in hidden layer and 2 output neurons was chosen to locate the optimum control point. Using several samples to calculate the gravity center and the moment of inertia led to obtain the network inputs after normalization. Network output is also found by normalizing the optimum control point using optimization algorithm. Now the optimum control point could be determined for every image matrix using the trained network.

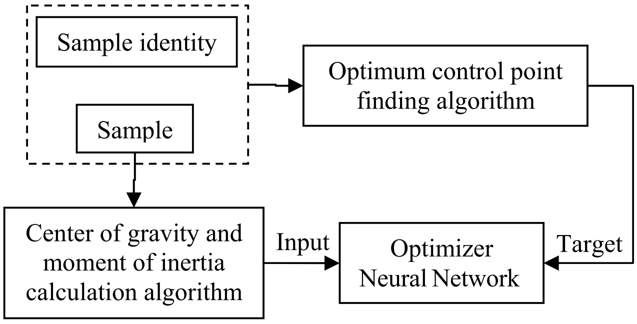

Figure 4: The offline training procedure for optimizer Neural Network

The necessary Neural Network for digit recognition was MLP with 12 input neurons, 20 neurons in hidden layer and 10 output neurons. To train this network, the control point was determined using the above mentioned trained Neural Network. Then feature vector was determined through mapping the asymmetrical pattern on the image matrix profiting from the optimum control point and fed into the Neural Network as inputs. The network behavior for the different digits is known. e.g. for digit 3, the output of the fourth neuron is 1 and that of the others is 0.

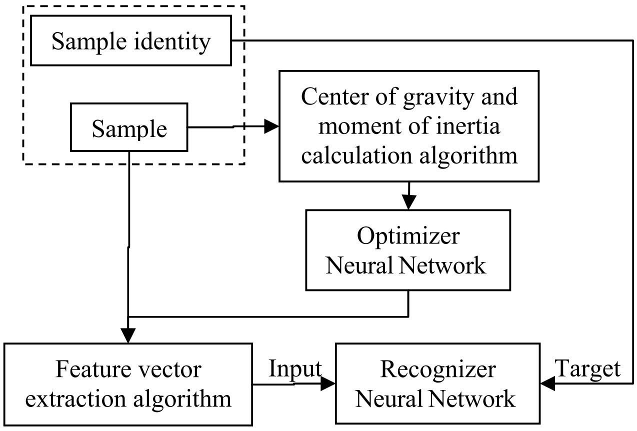

Figure 5: The offline training procedure for Recognizer Neural Network

2. Digit recognition:

Now it is possible to recognize the digits using two cascaded Neural Network. First the area surrounding the digit is cleaned out. Then the moment of inertia and the center of gravity are found for the image matrix. By applying these two values as inputs to the optimizer Neural Network will determine the optimum control point. Using this optimum control point, an asymmetrical pattern is mapped on the image matrix and the feature vector is calculated. Then applying this vector as an input to the second Neural Network, it returns an output which indicates the recognized digit.

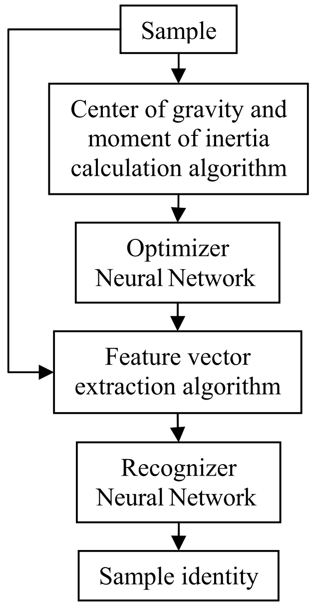

Figure 6: The block diagram of integrated proposed algorithm for handwritten Persian digits recognition

VII. Simulation Results

The proposed algorithm was simulated to determine its performance. For simulating the training procedure, more than 230 handwritten digits from 23 different people was used. Also 500 digits from 50 different people (also different from those 23 persons) were used for determine the recognition rate. Some of these samples of digit are shown in figure 7.

Table 1 shows the simulation results. The recognition rate of this algorithm is found to be 97.6%. The recognition rate of the algorithms presented in [1], [2], and [3] is reported to be 91.81%, 97.8%, 92%, respectively.

These results indicate that the recognition rate of this method is a little less than the recognition rate that reported in reference [2]. The most important reason of this problem is that in [2] 32 different segmentation parts are for feature vector extraction but we used 12 different segmentation parts. Another reason is that they have used handwritten samples from only 5 different people. So, it is evident that the proposed method would represent better recognition rate than other methods under similar conditions.

According to table 1, it is apparent that the output of Neural Network for some input is unknown. This is the case when no neuron outputs one. Since these outputs are easily distinguishable, another algorithm may be used to increase the accuracy of the current algorithm by finding the specified inputs.

Table 1: Output of the proposed algorithm for 500 digits

	0	1	2	3	4	5	6	7	8	9	U
0	48	0	0	0	0	2	0	0	0	0	0
1	0	50	0	0	0	0	0	0	0	0	0
2	0	0	48	1	0	0	0	0	0	0	1
3	0	0	2	48	0	0	0	0	0	0	0
4	0	0	0	0	48	0	0	0	0	0	2
5	1	0	0	0	0	49	0	0	0	0	0
6	0	0	0	0	1	0	48	0	0	0	1
7	0	0	0	0	0	0	0	50	0	0	0
8	0	0	0	0	0	0	0	0	50	0	0
9	0	0	0	0	0	0	0	0	0	48	2

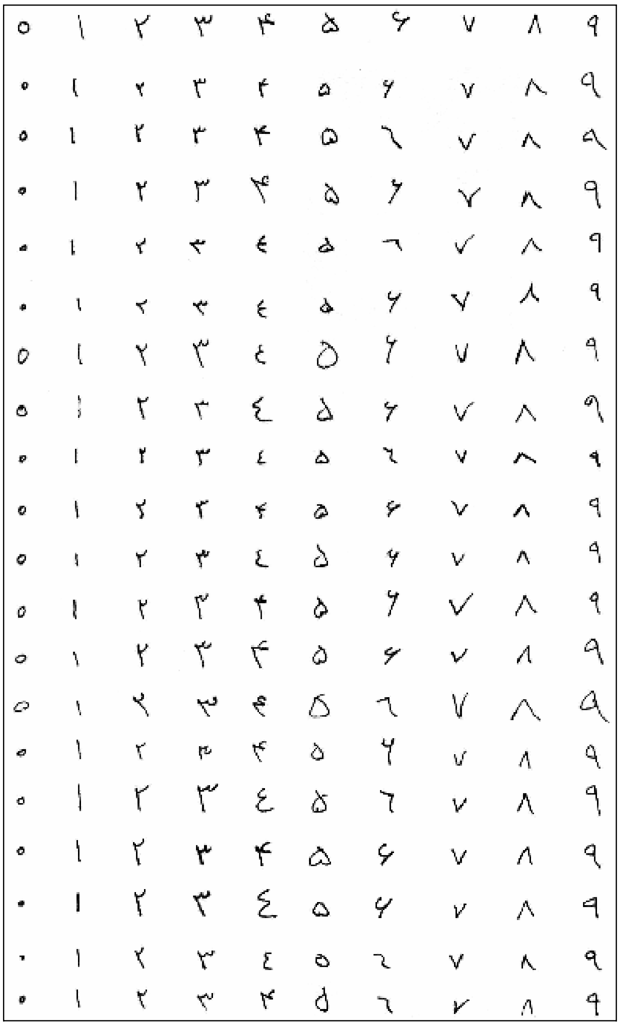

Figure 7: Handwritten digit samples, each row from 0 to 9 (from left to right)

VIII. Conclusions

In this paper an asymmetrical segmentation pattern was introduced and used to obtain feature vector from an image matrix. One of the benefits of using this pattern is to provide better characteristics, thanks to the control point (one degree of freedom); therefore an algorithm was designed to exploit this characteristic.

But, as far as this algorithm must be dependent and adjusted to the digit identity, a Neural Network with proper inputs is designed to resolve this problem. The simulation results show the high performance of this new model for recognizing the Persian digits, thanks to the characteristics of the new feature vector. Hence using a pattern with more flexibility (e.g. 2 or 3 degree of freedom) can even show better performance.

References

1. A. Mowlaei, K. Faez and A. T. Haghighat, "Feature extraction with Wavelet Transform for recognition of isolated handwritten Farsi/Arabic character an numerals", 14th International Conference on Digital Signal Processing, vol. 2, 1-3 July 2002
2. M. H. Shirali, K. Faez and A. Khotanzad, "Recognition of handwritten Persian/Arabic numerals by shadow coding and an edited probabilistic Neural Network", International Conference on Image Processing, vol. 3, 23-26 October 1995
3. H. M. M. Hosseini and A. Bouzerdoum, "A combined method for Persian and Arabic handwritten digit recognition", Australian New Zealand conf. on Intelligent Information Systems, Adelaide, Australia, 18-20 Nov. 1996.
4. D. J. Burr, "Experiments on Neural Net recognition of spoken and written text," IEEE Transactions on Acoustic, Speech, and Signal Processing, Vol. 36, No. 7, pp. 1162-1168, July 1988
5. S. Haykin, "Neural networks a comprehensive foundation", second edition, Prentice-Hall, Inc, 1999.