

Using Modified Contour Features and SVM Based Classifier for the Recognition of Persian/Arabic Handwritten Numerals

Alireza Alaei¹, Umapada Pal² and P. Nagabhushan³

^{1,3}Department of Studies in Computer Science, University of Mysore, Mysore, 570 006, India

²Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata-108, India

¹alireza20alaei@yahoo.com, ²umapada@isical.ac.in, ³pnagabhushan@hotmail.com

Abstract

In this paper, we propose a robust and efficient feature set based on modified contour chain code to achieve higher recognition accuracy of Persian/Arabic numerals. In classification part, we employ support vector machine (SVM) as classifier. Feature set consists of 196 dimensions, which are the chain-code direction frequencies in the contour pixels of input image. We evaluated our scheme on 80,000 handwritten samples of Persian numerals. Using 60,000 samples for training, we tested our scheme on other 20,000 samples and obtained 98.71% correct recognition rate. Further, we obtained 99.37% accuracy using five-fold cross validation technique on 80,000 dataset.

Keywords: Persian/Arabic Numeral Recognition, Chain Code, Support Vector Machine

1. Introduction

With the increased use of Persian/Arabic writing in many day-to-day businesses in Persian/Arabian countries, it has become necessary for the machines to understand handwritten materials written in Persian/Arabic. As a part of Persian/Arabic scripts, numeral strings and isolated numerals play an enormous role. OCR for handwritten documents in some languages (English, Chinese, Japanese, etc.) has reached to a promising level [6]. The OCR for Persian/Arabic has not grown up like those languages because of the cursive-ness of handwritten in Persian/Arabic and multiple forms of each character with respect to its position in words.

Towards this end, we studied the effect of directional frequencies in the contour pixels of numeral image as features, which kept shape information of input and then applied SVM as classifier.

1.1. The Persian numerals characteristics

Persian numerals are used in Iran and in some of its neighboring countries. Comparable to other scripts, in Persian also there are 10 numerals. In Persian/Arabic scripts, alphabets are written from right to left but digits are written from left to right. Persian and Arabic numerals are almost the same; but there are some important differences between handwriting of digits of these two scripts [15]. Generally, in Persian digits, there are two types of writing for the digits 0, 2, 4, 5 and 6. These characteristics make the recognition of Persian numerals more complicated than in other languages. Examples of Printed and handwritten Persian digits are shown in Fig.1 and Fig. 2.

0	1	2	3	4	5	6	7	8	9
٠, ۰	۱	۲, ۲	۳	۴, ۴	۵, ۵	۶, ۶	۷	۸	۹

Fig. 1. Printed samples of Persian digits

1.2. A brief survey on Persian/Arabic numeral recognition

In the literature survey particularly relevant to Persian/Arabic languages, there are many methods for feature extraction and classification. As feature extraction methods; segmentation and shadow code [1, 7, 11], fractal code [13], profiles [3, 15], moment [5], template [8], structural feature (points, primitives) [14] and wavelet [9, 12] have been used. For classification, different types of Neural Networks [1, 5, 7, 8, 11, 12, 13], SVM's [3, 9, 15] and Nearest Neighbor [14] have been applied.

From the literature survey of the existing pieces of works on Persian/Arabic numerals recognition, it is evident that not much effort was extended to identify a more efficient feature set (most of them are time consuming process and some of them cannot preserve

the shape of the input image for feature extraction step), which could more appropriately be reacted to the recognition phase. To overcome such problem, we proposed to find out a more effective feature set based on modified contour chain code of each window-map, and then apply SVM for classification. This type of feature set, which expressed the physical shape of input image and extracted local information of the input image in each window-map, provided very good accuracy in experimental part.

We should mention the proposed system did not use any preprocessing techniques (skew and slant detection/correction, thinning, smoothing, noise removal, etc) that were expensive operations. Moreover, robustness of our feature set was taken care of some of these issues like skew and slant reasonably.

The organization of rest of the paper is as follows: In Section 2 we illustrate feature extraction technique, Section 3 describes classification, experimental results and comparative analysis are described in Section 4 and finally in last section we present conclusion.

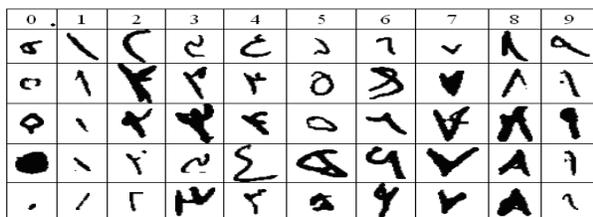


Fig. 2. Handwritten samples of Persian digits

2. Feature extraction

Directional chain code information of the contour points of the input image can be used as features for different purposes like character segmentation, recognition, etc [16]. In our system we computed features based on chain-code directional frequencies of contour pixels of the images as follows: First we found the bounding box (minimum rectangle containing the numeral) of each input image which is a two-tone image. Then for better result and independency of features to size, font and position (invariant to scale and translation), we converted each image (located in bounding box) to a normal size of 49×49 pixels. We chose this normalized value based of various experiments and a statistical study. With a statistical study on behavior of training dataset, we found that more than 96% of the images have a width/length less than 30 pixels. To obtain numeral shapes more clear we normalized them into 49×49 . In Fig. 3, a normalized image with its bounding box is shown. We extracted the contour of the normalized image [Fig. 4].

We scanned the image contour horizontally by keeping a window-map of size 7×7 on the image from the top left most point to down right most point (49 non-overlapped blocks). For each block the chain code frequencies for all 8 directions were computed (8 directions were shown in Fig. 5). Instead of expressing the features in terms of 8 directions, we have proposed to simplify the features into 4 sets corresponding to 4 directions [Fig. 6]: i) horizontal direction code (direction 0 and 4), ii) vertical direction code (direction 2 and 6) iii) principal diagonal direction code (direction 1 and 5) and iv) off diagonal direction code (direction 3 and 7). Thus, in each block, we got four values representing the frequencies of these four directions and these values were used as features (local contour direction values). To extract features, we considered 49 (7×7) uniform blocks in each image and we computed four features in each block so we got $49 \times 4 = 196$ features for each image.



Fig. 3. Bounding box of a normalized image



Fig. 4. Digit '5' in Persian and its contour



Fig. 5. Point P and its 8-direction codes for its 8 neighbors



Fig. 6. 4 directions obtained from 8 directions

3. Classification

The use of SVMs has frequently been reported in literature to be a very suitable classifier [3, 9, 15] especially for problems with small number of classes like numeral recognition. We used SVM for the recognition purpose.

The SVM was defined for two-class problem and it looked for the optimal hyper-plane, which maximized the distance, the margin, between the nearest examples of both classes, named support vectors (SVs) [2].

The linear SVM can be extended to a non-linear classifier by using kernel functions like polynomial and Gaussian kernels. We have tested linear, Gaussian and polynomial kernels during our experiments and we received the best result using Gaussian kernel thus we employed SVMs with Gaussian kernel as classifier. Details of SVM can be found elsewhere [2, 4].

We utilized one-against-all SVMs for the proposed classifier architecture. The input feature sets were the directional features (196-dimension). All the SVMs trained with the respective training feature sets and the results explored by using separate test data. We

obtained the best results with the Gaussian kernel of $\gamma=0.002$.

4. Experimental and comparative results

In this study, for experimental analysis, we considered a standard Persian numeral dataset [10]. Training and testing sets and experimental results describe in the next subsections.

4.1. Data set

For experimental analysis, we considered 60,000 samples (6,000 samples per class) for training and 20,000 samples (2,000 samples per class) for testing as mentioned in [10]. These samples were extracted from different registration forms of entrance examinations of universities in Iran containing Iranian Postal and National Codes. The images were scanned at 200 dpi resolution [10]. Because of writing styles of different individuals, samples sizes were very different, and as we discussed in Section 2 we normalized them.

4.2. Performance of the proposed system

Using 60,000 samples for training, we tested our scheme on other 20,000 samples and obtained 98.71% accuracy. From the experiment, we got an accuracy of 99.99% when the 60,000 data were used as training and the same data set was used for testing. In another experiment, we used 5-fold cross validation scheme for recognition result calculation. We divided our database (80,000 samples) into 5 subsets and testing is done on each subset using rest of the 4 subsets for training. The recognition rates for all the five test subsets of dataset are averaged to get the accuracy. We got the average accuracy of 99.37%. Further, we considered some noisy (salt and pepper) images (Fig.7) in our test data. The result showed the effectiveness of the proposed feature extraction technique (Table I).

4.3. Confusion pairs

In our experiment (with the 98.71% accuracy), we observed confusing numerals in the recognition phase between some digits. In Table I, we showed Detail of confusing results. The major confusions were amongst 2, 3 and 4. This happened because 2, 3 and 4 look like each other. From the Table I it may be noted that out of 2000 samples of number three (3), 59(2.95%) samples misrecognized to numeral 2 and 24(1.2%) samples misrecognized to numeral 4. In some of the samples, little confusions were also between 0 and 1. Some of these similar shapes are shown in Table II.

4.4. Erroneous samples

From the experiment, we noted erroneous samples and to get the idea of erroneous samples some of them are given in Table III. This table clearly shows that a part of the error set has resulted because of very poor quality samples (for example, the sample of numeral 2 has many discontinuities). In some cases, even human being cannot easily recognize them. In our experiment, we found that in some samples, there were two, three and sometimes more broken parts, which made recognition task more difficult. As we mentioned earlier, we did not use any preprocessing technique to connect isolated components in the input images, if any. We got some recognition errors because of this.

Table I. Confusion matrix of the result

	0	1	2	3	4	5	6	7	8	9
0	1956	32	0	0	0	9	0	3	0	0
1	12	1986	0	0	1	0	0	0	0	1
2	0	1	1981	10	4	0	1	1	0	2
3	0	1	59	1914	24	2	0	0	0	0
4	0	0	7	13	1979	1	0	0	0	0
5	4	0	0	0	5	1987	0	0	3	1
6	2	8	1	0	2	2	1974	0	0	11
7	0	4	3	1	0	0	0	1992	0	0
8	0	1	0	0	0	0	0	0	1997	2
9	3	9	0	0	0	3	8	0	1	1976

Table II. Similar samples of Persian digits

Numerals	0	1	2	3	4
Similar Shapes					

Table III. Some errors of our system

Class	0	1	2	3	4	5	6	8	9
Samples									
Recognized class									
Correct class									



Fig. 7. The noisy samples of Persian digits

4.5. Comparison of results

To compare the performance of our method we noted the performances of most of the works that were available for Persian numeral recognition. See Table IV for details of comparison. It may be noted from Table IV that all the existing works were evaluated on smaller datasets. The highest dataset of size 10,000

was used by a recent work due to Ziaratban et al. [8], where as we used 80,000 data for our experiment. The highest accuracy was obtained from the work due to Soltanzadeh et al. [3] but they have experimented with only 8,918 samples and used 257 dimensional features. We considered 80,000 data for our system and we obtained 98.71% and 99.37% accuracies using only 196 dimensional features.

Table IV. Comparison of different algorithms

Algorithms	Dataset size		Accuracy (%)	
	Train	Test	Train	Test
Shirali-shahreza et al.[1]	2600	1300	-	97.80
Soltanzadeh, Rahmati [3]	4979	3939	-	99.57
Dehghan, Faez [5]	6000	4000	-	97.01
Harifi., Aghagolzadeh [7]	230	500	-	97.60
Ziaratban et al. [8]	6000	4000	100	97.65
Mowlaei, Faez [9]	2240	1600	100	92.44
Hosseini, Bouzardum [11]	480	480	-	92.00
Mowlaei et al. [12]	2240	1600	99.29	91.88
Mozaffari et al. [13]	2240	1600	98.00	91.37
Mozaffari et al. [14]	2240	1600	100	94.44
Sadri et al. [15]	7390	3035	-	94.14
Proposed Algorithm	60000	20000	99.99	98.71
Proposed Algo. (5-fold)	80000		-	99.37

5. Conclusion

In this paper, an efficient feature extracting technique is proposed. From experimental results, it is evident that our features resulted good performances (98.71%, 99.37%). We noted that most of misclassified samples were from classes of 2, 3 and 4, which are similar shape. The recognition of such similar numerals was difficult even by human being. It is obvious that by removing confusion among few classes, we can achieve better performance. Further, to the best of our knowledge, this work is the first work, towards the recognition of Persian handwritten numerals on a huge dataset. In future, we plan to handle confusions amongst similar classes by removing the common part from the numerals that introduce confusion and apply principal component analysis to achieve feature set with lesser number of features.

Acknowledgment

A. Alaei would like to thank Mr. S. Chanda for his help and kind cooperation.

References

[1] M. H. Shirali-Shahreza, K. Faez and A. Khotanzad, "Recognition of Hand-written Persian/Arabic Numerals by Shadow Coding and an Edited Probabilistic Neural

Network", Proceedings of International Conference on Image Processing, Vol. 3, 1995, pp. 436-439.

[2] C. Burges, "A Tutorial on support Vector machines for pattern recognition", Data Mining & Knowledge Discovery, Vol. 2, 1998, pp. 1-43.

[3] H. Soltanzadeh and M. Rahmati, "Recognition of Persian handwritten digits using image profiles of multiple orientations", *Pattern Recognition Letters* 25, 2004, pp. 1569-1576.

[4] V. N. Vapnik, "*The Nature of Statistical Learning Theory*", Springer Verlag, 1995.

[5] M. Dehghan and K. Faez, "Farsi Handwritten Character Recognition With Moment Invariants", Proceedings of 13th International Conference on Digital Signal Processing, Volume 2, 1997, pp. 507-510.

[6] S. N. Srihari and G. Ball, "An Assessment of Arabic Handwriting Recognition Technology", CEDAR Technical Report TR-03-07, 2007.

[7] A. Harifi and A. Aghagolzadeh, "A New Pattern for Handwritten Persian/Arabic Digit Recognition", *Journal of Information Technology* Vol. 3, 2004, pp. 249-252.

[8] M. Ziaratban, K. Faez and F. Faradji, "Language-Based Feature Extraction Using Template-Matching in Farsi/Arabic Handwritten Numeral Recognition", Proceedings of 9th International Conference on Document Analysis and Recognition, Vol.1, 2007, pp. 297-301.

[9] A. Mowlaei and K. Faez, "Recognition Of Isolated Handwritten Persiawarabic Characters And Numerals Using Support Vector Machines", Proceedings of XIII Workshop on Neural Networks for Signal Processing, 2003, pp. 547-554.

[10] H. Khosravi and E. Kabir, "Introducing a very large dataset of handwritten Farsi digits and a study on the variety of handwriting styles", *Pattern Recognition Letters* Vol.28, Issue 10, 2007, pp. 1133-1141.

[11] H. Mir Mohammad Hosseini and A. Bouzardum, "A Combined Method for Persian and Arabic Handwritten Digit Recognition", Australian New Zealand Conference on Intelligent Information System, 1996, pp. 80 - 83.

[12] A. Mowlaei, K. Faez & A. Haghghat, "Feature Extraction with Wavelet Transform for Recognition of Isolated Handwritten Farsi/Arabic Characters and Numerals", *Digital Signal Processing* Vol. 2, 2002, pp. 923- 926.

[13] S. Mozaffari, K. Faez & H. Rashidy Kanan, "Recognition of Isolated Handwritten Farsi/Arabic Alphanumeric Using Fractal Codes", *Image Analysis and Interpretation*, 6th Southwest Symposium, 2004, pp. 104-108.

[14] S. Mozaffari, K. Faez and M. Ziaratban, "Structural Decomposition and Statistical Description of Farsi/Arabic Handwritten Numeric Characters", Proceedings of the 8th Intl. Conference on Document Analysis and Recognition, Vol. 1, 2005, pp. 237- 241.

[15] J. Sadri, C. Y. Suen and T. D. Bui, "Application of Support Vector Machines for Recognition of Handwritten Arabic/Persian Digits", Proceedings of the 2nd Conference on Machine Vision and Image Processing & Applications, Vol. 1, 2003, pp. 300-307.

[16] N. Sharma, U. Pal and F. Kimura, "Recognition of Handwritten Kannada Numerals", Proceedings of the 9th International Conference on Information Technology, Vol. 1, 2006, pp. 133-136.