A Statistical Approach For Latin Handwritten Digit Recognition

Ihab Zaqout
Dept. of Information Technology
Faculty of Engineering & Information technology
Al-Azhar University – Gaza
Gaza Strip, Palestine

***Abstract*— A simple method based on some statistical measurements for Latin handwritten digit recognition is proposed in this paper. Firstly, a preprocess step is started with thresholding the gray-scale digit image into a binary image, and then noise removal, spurring and thinning are performed. Secondly, by reducing the search space, the region-of-interest (ROI) is cropped from the preprocessed image, then a freeman chain code template is applied and five feature sets are extracted from each digit image. Counting the number of termination points, their coordinates with relation to the center of the ROI, Euclidian distances, orientations in terms of angles, and other statistical properties such as minor-to-major axis length ratio, area and others. Finally, six categories are created based on the relation between number of termination points and possible digits. The present method is applied and tested on training set (60,000 images) and test set (10,000 images) of MNIST handwritten digit database. Our experiments report a correct classification of 92.9041% for the testing set and 95.0953% for the training set.**

Keywords- Digit recognition; freeman chain coding; feature extraction; classification.

I. Introduction

The significant task of handwritten digit recognition has great importance in the recognition of postcodes sort mail, bank check amounts and so on. Since three decades, there is no single classifier performs the best for all pattern classification problems consistently. There are different challenges faced while attempting to solve this problem. The handwritten digits are not always of the same thickness, size, orientation or position relative to margins.

There are several approaches for handwritten digit recognition problem have been reported in the literature in the past. They include SVM (support vector machine) [1, 2, 3], NN (neural network) [4, 5, 6, 7], deformable template matching [8, 9], hybrid method [10, 11, 12, 13] and others. In SVMs for digit classification problems, the training of a large data set is still a bottle-neck and is comparatively slow. NNs have been widely used to solve complex classification problems. However, A single NN often exhibits with the over fitting behavior which results in a weak generalization performance when trained on a limited set of training data. A better deformation algorithms and proper selection of representative prototypes along with its computational requirements are required for deformable template matching method. Hybrid method has been widely used in pattern recognition applications. It combines two or more of the above mentioned methods or others to overcome their individual weakness and to preserve their individual advantages. But it is still an open problem to obtain a superior hybrid method.

A comprehensive benchmark of handwritten digit recognition with several state-of-the-art approaches, datasets, and feature representations had been reported by [14]. Several classifiers and feature vectors are evaluated on MNIST handwritten digit database [15].

In this paper, our research is focused on an accurate and feasible method applied and tested on training set (60,000 images) and test set (10,000 images) of MNIST handwritten digit database. We agree with others that the preprocess stage is a crucial and it reflects the accuracy of the classification process. Firstly, a preprocess step is started with thresholding the gray-scale digit image into a binary image, and then noise removal, spurring and thinning are performed. Secondly, by reducing the search space, the region-of-interest (ROI) is cropped from the preprocessed image, then a freeman chain code template is applied and five feature sets are extracted from each digit image. Counting the number of termination points, their coordinates with relation to the center of the ROI, Euclidian distances, orientations in terms of angles, and other statistical properties such as minor-to-major axis length ratio, area and others.

Finally, six categories are created based on the relation between number of termination points and possible digits. The advantage of method is that it does not require training, which can save a lot of training time. Experimental results on MNIST database will be reported in the paper to support the feasibility of our method. The remainder of this paper is organized as follows. The system overview is presented in Sec. 2. The feature extraction is proposed in Sec. 3. In Sec. 4, the classification is discussed. Experimental results are shown in Sec. 5 to demonstrate the reliability of our method. Finally our conclusion and future work is given in Sec 6.

II. Sysyem Overview

In this section, the system overview is introduced. The recognition system includes three dependent stages: the preprocess, the feature extraction and classification as shown in Figure 1.

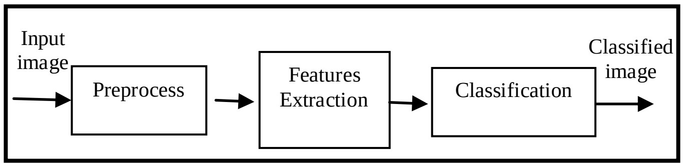


Figure 1. The System Overview

In this paper, a MNIST digit database is used as a dataset to the proposed classification processes. A preprocessed step is started with thresholding the gray-scale digit image into a binary image, and then noise removal, spurring and thinning [16] are performed. During the process of image thinning, we introduce Freeman chain code tracking [17] as shown in Figure 2. To reduce the search space, the region-of-interest (ROI) is cropped from the preprocessed image, then a freeman chain code template is applied and five feature sets are extracted from each digit image. Counting the number of termination points, their coordinates with relation to the center of the ROI, Euclidian distances, orientations in terms of angles, and other statistical properties such as minor-to-major axis length ratio, area and others. Six categories are created based on the relation between number of termination points and possible digits.

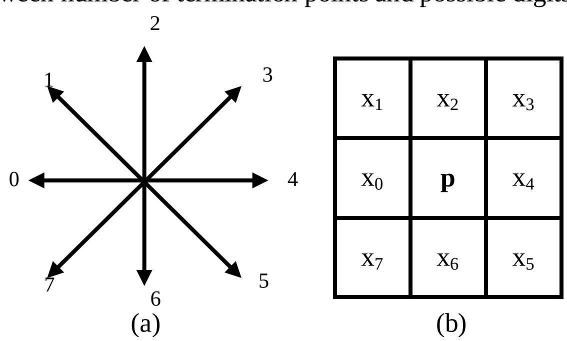


Figure 2. Freeman chain code and 3x3 template

III. FEATURE EXTRACTION

A fast recognition system has to consider the processing time and its complexity is primarily determined by the feature extraction. In this section, the feature extraction is performed on the thinned image. Five feature sets are extracted from each ROI digit image:

- Number of termination points, xy-coordinates and orientations,
- Minimum-to-maximum distance ratio (Euclidean distance between the xy-coordinate of the center of the ROI and xy-coordinates of termination points), minor-to-major axis length ratio and their spatial distribution compared to the center of the ROI,
- Sum of white pixels after partitioning the thinned digit image into horizontal blocks for 0, 2, 3, 5, 6 and 8,
- Other statistical measurements such as area, minor-to-major axis length ratio, filled area to area ratio, filled area to convex area ratio and number of holes, and
- Sum Freeman chain code.

As shown in Fig. 2(b), the pixels x_1, x_2, …, x_8 are the 8-neighbours of p in its 3×3 template and said to be 8-adjacent to p. We will use x_i to denote both the pixel and its value 0 or 1, x_i is called white or black, accordingly. In the thinned image, the pixel p is examined for termination point is detected if T (summation of neighbors of the central pixel p, where p is a white pixel and its value equal to 1) is equal to 1 as shown in in the following equation:

$$T = \sum_{i=0}^{7} x_i = 1 \qquad (1)$$

IV. CLASSIFICATION

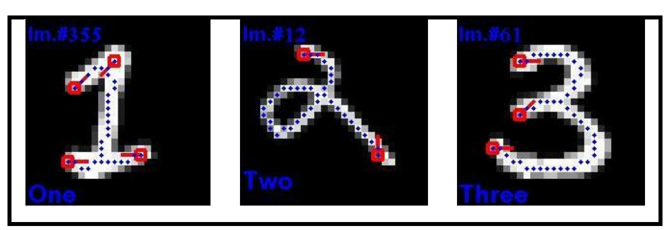
After the features extraction in the previous section, the classification process in our recognition system is divided into two stages. In the first stage, mainly counting the number of termination points, the digits are classified into six categories. The relation between the number of the termination points and possible digits is shown in Table I. In the second stage, other features such as freeman chain code, orientation, positions, distances and others, all the digits are recognized.

TABLE I. THE RELATION BETWEEN NUMBER OF TERMINATION POINTS AND DIGITS

Number of termination points	Possible digits									
	0	1	2	3	4	5	6	7	8	9
0	√								√	
1	√		√				√		√	√
2	√	√	√	√	√	√	√	√	√	√
3		√	√	√	√	√		√		√
4		√	√	√	√	√		√		
5					√			√		
> 5	Rejected									

The implementation of the above mentioned two-staged classification is depicted in Figure 3. We can obviously notice the following:

1. The result of the thinned operation appears as blue colored dots,
2. Each termination point is bounded with a red square color including its orientation information,
3. The sequence of each digit image as exists in its test set appears on the top-left corner with blue color, and
4. The classification result appears on the bottom-left corner with blue color.



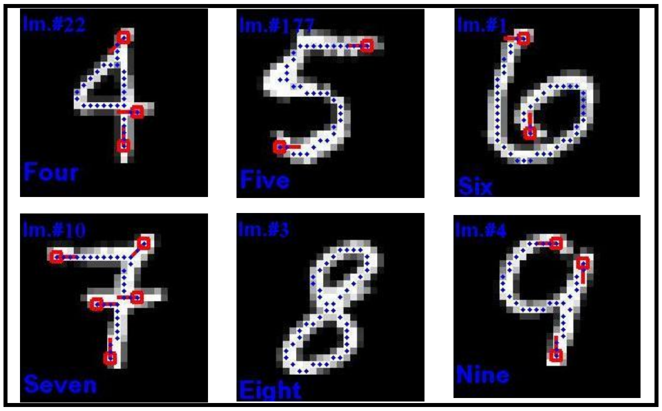


Figure 3. The implementation of the two-staged classification process

V. EXPERIEMENTAL RESULTS

Our experiments are performed on MNIST digit database by using MATLAB 6.5.1 release 13. This database consists of 10,000 gray-scale digit patterns in testing set and 60,000 gray-scale digit patterns in training set. The size of each digit image is 28×28 pixels. All results are obtained by using 2.40 GHz P4 processor under Windows XP. The processing time of each single digit is around 0.488 s.

The results of our recognition system tested on the testing set are summarized in Table II and on the training set are shown in Table III. It's shown that for each class, there exists a true positive (TP) means that the digit is correctly classified, false positive (FP) means that the digit is wrongly classified, rejected (RJ) means that the digit is not classified or unknown and reliability (RL) is calculated as:

$$RL\,(\%) = \frac{TP(\%)}{100 - RJ\,(\%)} * 100 \qquad (2)$$

TABLE II. THE PERFORMANCE OF THE CLASSIFIERS ON THE TESTING SET

#	No. of Images	True Positive (TP%)	False Positive (FP%)	Rejected (RJ%)	Reliability (RL%)
0	980	95.5102	2.8571	1.6327	97.0955
1	1135	97.0925	2.5551	0.3524	97.4359
2	1032	92.3450	4.6512	3.0039	95.2049
3	1010	94.3564	4.4554	1.1881	95.4909
4	982	94.3992	4.5825	1.0183	95.3704
5	592	90.5405	6.9257	2.5338	92.8943
6	958	91.0230	6.0543	2.9228	93.7635
7	1028	91.1479	6.4202	2.4319	93.4198
8	974	90.5544	8.9322	0.5133	91.0216
9	1009	92.0714	4.9554	2.9732	94.8928
Overall Performance %		**92.9041**	**5.2389**	**1.8570**	**94.6589**

TABLE III. THE PERFORMANCE OF THE CLASSIFIERS ON THE TRAINING SET

#	True positive		False positive		Rejection		Reliability
	No. of Images	TP%	No. of Images	FP%	No. of Images	RJ%	RL%
0	5698	96.2175	159	2.6849	65	1.6327	97.2853
1	6498	97.7584	113	1.7000	36	0.3524	98.2907
2	5803	94.2658	246	3.9799	108	3.0039	95.9491
3	5918	94.9615	237	3.8030	77	1.1881	96.1495
4	6433	95.2332	259	3.8342	63	1.0183	96.1297
5	5126	93.0309	253	4.5917	131	2.5338	95.2965
6	7810	94.7011	301	3,6498	136	2.9228	96.2890
7	5150	94.8260	196	3.6089	85	2.4319	96.3337
8	3235	95.2872	119	3.5052	41	0.5133	96.4520
9	5401	94.6713	211	3.6985	93	2.9732	96.2402
Σ	57072		2093		835		
Overall Performance %	**95.0953**		**3.5056**		**1.3991**		**96.4416**

As shown in the above tables, the overall performance for Latin handwritten digit recognition is 92.9041% and 95.0953, consecutively.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have described a simple approach for feature extraction and classification for Latin handwritten recognition. Five feature sets are obtained from the thinned digit image using the concept of the freeman chain code template. A two-staged classification is implemented. Firstly, counting the number of termination points, the digits are classified into six categories. The relation between the number of the termination points and possible digits is easily detected. Secondly, other features such as freeman chain code, orientation, positions, distances and others are calculated. The overall performance is 92.9041% for the recognition of 10,000 digit images in the test set and 95.0953% for the recognition of 60,000 digit images in the training set.

In the future work, a tapped operation for disconnected digits will be tested in the preprocess stage. Furthermore, we will test the above mentioned approach in the classification of the formation of the characters in the Arabic language.

REFERENCES

[1] D. Gorgevik, D. Cakmakov and V. Radevski, Handwritten Digit Recognition by Combining Support Vector Machines Using Rule-Based Reasoning, *Proc. of the 23rd International Conference on Information Technology Interfaces*. **19**(12) (2001) 139–144.

[2] D. Gorgevik and D. Cakmakov, Combining SVM Classifiers for Handwritten Digit Recognition, *Proc. of the 16th International Conference on Pattern Recognition*. **3** (2002) 102–105.

[3] S. Maji and J. Malik, Fast and Accurate Digit Classification, *Technical Report No. UCB/EECS-2009-159*. (2009).

[4] Z. Chi, Z. Lu and F-H. Chan, Multi-Channel Handwritten Digit Recognition Using Neural Networks, *Proc. of IEEE International Symposium on Circuits and Systems*. **1**(1997) 625–628.

[5] B. Duerr, W. Haettich, H. Tropf and G. Winkler, A combination of statistical and syntactical pattern recognition applied to classification of unconstrained handwritten numerals, *Pattern Recognition*. **12** (1980) 189–199.

[6] F. Shah and K. Yousaf, Handwritten Digit Recognition Using Image Processing and Neural Networks, *Proceedings of the World Congress on Engineering*. **1** (2007).

[7] E. Kussul and T. Baidyk, Improved method of handwritten digit recognition tested on MNIST database, *Image and Vision Computing*. **22**(12) (2004) 971-981.

[8] A. Jain and D. Zongker, Representation and Recognition of Handwritten Digits Using Deformable Templates, *IEEE Trans. Pattern Analysis and Machine Intelligence*. **19**(19) (1997) 1386–1390.

[9] V. Mane and L. Ragha, Handwritten digit recognition using elastic matching, *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*. (2011), pp. 1364.

[10] D. Gorgevik and D. Cakmakov, D, An efficient Three-Stage Classifier for Handwritten Digit Recognition, *Proc. of the 17th International Conference on Pattern Recognition*. **4** (2004) 507–510.

[11] A. Bellili, M. Gilloux and P. Gallinari, An Hybrid MLP-SVM handwritten digit recognizer, *Proc. of the 6th International Conference on Document Analysis and Recognition*. (2001), pp. 28–32.

[12] C. Pereira and G. Cavalcanti, Prototype Selection for Handwritten Connected Digits Classification, *The 10th International Conference on Document Analysis and Recognition*. (2009), pp. 1021-1025.

[13] R. Ebrahimpour and S. Hamedi, Hand Written Digit Recognition by Multiple Classifier Fusion based on Decision Templates Approach, *In Proc. of the International Conference on Computer, Electrical, Systems Science, and Engineering (CESSE)*. (2009), pp. 245-250.

[14] L. Cheng-Lin, N. Kazuki, H. Sako and H. Fujisawa, Handwritten Digit Recognition: Benchmarking of State-of-art Techniques, *Pattern Recognition*. **36** (2003) 2271-2285.

[15] MNIST Database of Handwritten digits: http://yann.lecun.com/exdb/mnist/.

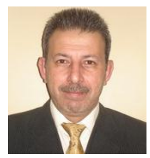
[16] L. Lam, S-W. Lee and C. Suen, Thinning methodologies-a comprehensive survey, *IEEE Trans. Pattern Analysis and Machine Intelligence*. **14**(9) (1992) 869–885.

[17] H. Freeman, On the Encoding of Arbitrary Geometric Configurations, IRE Trans. Pattern Analysis and Machine Intelligence Electronics and Computers. **10**(1961) 260–268.

[18] Kekre, H. B. (2010). Texture Based Segmentation using Statistical Properties for Mammographic Images. *IJACSA - International Journal of Advanced Computer Science and Applications*, *1*(5), 102-107.

[19] Malhotra, R. (2011). Software Effort Prediction using Statistical and Machine Learning Methods. *IJACSA - International Journal of Advanced Computer Science and Applications*, *2*(1), 145-152.

AUTHOR PROFILE



Ihab Zaqout received the B. S. in Computer Science from the University of Al-Fateh, Libya, in 1987 and the M.S. degree in Computer Science from Jordan University, Jordan in 2000 and the Ph.D. in Computer Science from the University of Malaya, Malaysia in 2006. He is currently at the Dept. of Information Technology, Al-Azhar University - Gaza, Palestine as an assistant professor. His main research interests include image processing, pattern recognition, data mining and machine learning.