Touchless Written English Characters Recognition using Neural Network

Bikash Chandra Karmokar[1], M. A. Parvez Mahmud[2], Md. Kibria Siddiquee[3], Kawser Wazed Nafi[4], Tonny Shekha Kar[5]
Department of Computer Science and Engineering[1, 3, 4, 5]
Department of Electrical and Electronic Engineering[2]
Khulna University of Engineering and Technology
Khulna-9203, Bangladesh

Abstract—Touchless written English character recognizer (TER), a new touchless approach to write and an intelligent approach to recognize English characters has been proposed in this paper. In TER, the inputs of English characters have been taken by touchless fashion i.e. by sensing specific color object with a moving hand tracking in front of a webcam. Then they have been recognized by efficient Artificial Neural Network (ANN). Like the application of other traditional computer input devices such as mouse or keyboard, TER can be extended to write and recognize English words and sentences by adding characters one by one to the text editor. Proposed TER has been applied for several different forms of touchless writings, namely 26 English characters and 10 English digits. Here for training, ANN with Scale Conjugate Gradient (SCG) method has been used that converges the training time faster and recognizes with good generalization ability. TER can be useful for the disabled persons.

Keywords- *Scale Conjugate Gradient, Back Propagation, Principal component analysis, Character recognition.*

I. INTRODUCTION

Character recognition has become an acute research area in recent years for the ease of access of computer applications. Numerous approaches have been proposed for character recognition and considerable successes have been reported. Traditional handwritten character recognition techniques enable a computer to receive and interpret intelligible handwritten input from sources such as papers, documents, touch-screens or pictures [4]. Herein, usually they extract some defined characteristics called features to classify an unknown handwritten character into one of the known classes.

Until now, it is still a difficult task for a machine to recognize human handwritings with significant accuracy, especially under variable circumstances such as variations in writings, variable sizes, different patterns for different people etc. To recognize the handwritten characters of different languages, usually the existing approaches take inputs from sources like pictures, papers etc. Touchless screen has been rarely used in this purpose. None of the inputs of the existing approaches have been taken from other sources like using a webcam. An approach like this can help the disabled people who knew to write, but later on unable to write on paper by using hand because of some difficulties.

Some works have been done to recognize handwritten English characters [1], [2], [3], none of them are touchless writings i.e. these are written using pen or pencil on a white paper. Touchless written English character recognition is necessary for those people who use to write English using mouth movement holding an object on it. Therefore, touchless written English character recognizer (TER) has been proposed in this paper.

The outline of this paper is as follows: Section II describes existing character recognition techniques. Section III presents the proposed TER. Experimental studies have been discussed in Section IV. Finally, concluding remarks are explained in Section V.

II. EXISTING WORKS

Various approaches already have been proposed for handwritten character recognition. A typical handwriting recognition system consists of several steps, namely: preprocessing, segmentation, feature extraction, and classification. Several types of decision methods, including statistical method, artificial neural network (ANN), structural matching and stochastic process (Markov chain) etc have been used along with different types of features [15]. Many recent approaches mix several of these techniques together in order to obtain improved reliability, despite wide variation in handwriting. Most widely used approach is based on back-propagation (BP) ANN [10]. Here the ANN architecture is trained by a set of training data and then the input is classified by the trained ANN. Linear classification is also used to recognize handwritten characters [5]. Here the background basis of ANN has been implemented as a classification function.

The works of linear classification is very similar to ANN because the mapping of ANN cell or the one layer of ANN cell is equivalent to the linear discrimination function. Therefore if the ANN is two-layer i.e. consisting of an input and an output layer, it can act as a linear classifier. Multilayer ANN (MLANN) usually employs the BP algorithm and is also widely used in face recognition. In case of implicit segmentation approach [20], the words are recognized entirely without segmenting them into

letters. This is most effective and viable only when the set of possible words are small and known in advance, such as the recognition of bank checks, postal address etc. But it is not applicable to places where words come differently every time.

In structural approach [21], each pattern class is defined by structural description and the recognition is performed according to structural similarities [6]. Statistical approach [7] is also applied to character recognition. It is relatively insensitive to pattern noise and distortion, but modeling of statistical information is a tedious task [7]. Among other techniques, Hidden Markov Models, Fourier and Wavelet Descriptors [8], Fuzzy rules [9], tolerant rough set [10] are also used for recognizing handwritten character.

Principal component analysis (PCA) [14] is a well-known method for dimension reduction which plays a vital role for handwritten character recognition system. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated [11], [12], [13]. Recently Scale Conjugate Gradient (SCG) [16] is becoming a popular method for handwritten character recognition for its remarkable characteristics such as good generalization performance.

III. Proposed Touchless Written English Character Recognizer (TER)

This paper proposes a character recognition system for touchless written English characters. The input of the system is the images of the characters taken by the webcam. The proposed TER system combines PCA [14] based dimension reduction and SCG [17] algorithm for training for a better performance. Now-a-days, SCG has been applied in many engineering and scientific applications. We have selected SCG to develop TER for its following salient features:

- Avoids the line-search per learning iteration by using a Levenberg-Marquardt approach,
- An iterative method of learning,
- Learns faster because of locally tuned nodes, and
- Training and testing i.e. recognizing time is extremely low, and
- Its near-optimal parameters can estimate according to the properties of the feature space

The SCG algorithm along with efficient PCA based dimension reduction method enables TER to achieve faster training and recognition performance. Among the existing components that have been used to develop TER, only SCG is described below:

Scale Conjugate Gradient (SCG)

SCG is a second order conjugate gradient algorithm that helps to minimize the goal functions of several variables. This theoretical foundation was proved by Moller [17] which remains first order techniques in first derivatives like standard BP and finds the better way to a local minimum in second order techniques in second derivatives. SCG uses a step size scaling mechanism that avoids a time consuming line-search per learning iteration, which makes it faster than other recently proposed second order algorithms. Based on Moller [17], SCG method shows super-linear convergence in most problems.

A. Major Stages of TER

Major stages of the proposed scheme proceed as follows:

1) Preprocessing:

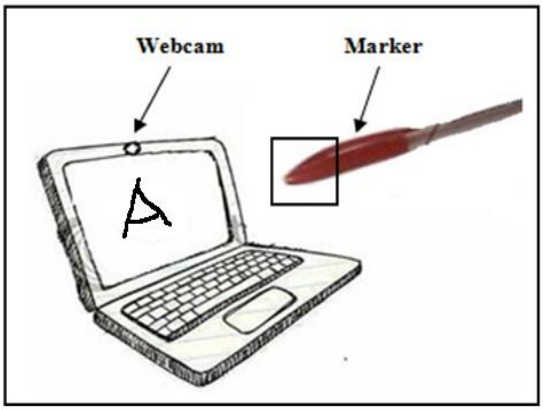
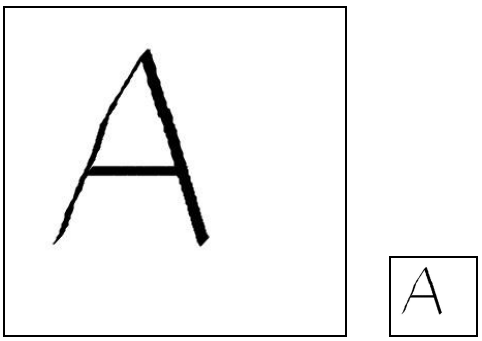


Figure 1. Marker tracking in action.



(a) (b)

Figure 2. Pre-recognition processing of an image.

This process starts through the tracking of a color marker in front of a webcam to capture the images of English characters as shown in Figure 1. Here, a pen with a red head has been used as a marker. The red color of the marker head has been kept unique in the environment to ensure uniform tracking. The movement of the marker has been done in such a way, which is usually used to write a character on a white board or like this. The tracked path has been drawn at the TER writing place which is the desired image for

training and recognition. As shown in Figure 1, the square area on the head of the marker is being tracked. Thus, the moving path has been drawn as a character which is shown in Figure 2(a). The scaled image has been shown in Figure 2(b). When touchless writings are captured, they may exist in jagged forms. Therefore it may contain some hooks and duplicated sampled points caused by hesitated writings. Furthermore, some points of a character may be missing, and some wild points might exist.

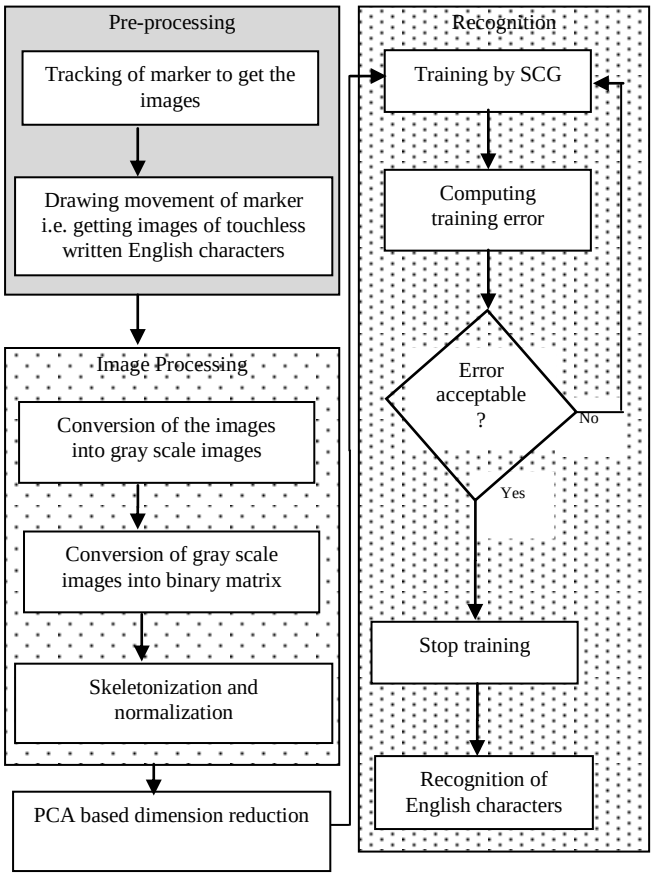


Figure 3. Block diagram of the proposed TER

2) Skeletonization and Normalization

In TER, skeletonization and normalization has been used to reduce the dimension of the images. Before skeletonization, unnecessary noises existing in the images have been removed for their smoothness. Spatial filters have been used to reduce the noises and the images are made smooth by Gaussian filter.

Before skeletonization process, the captured images have been converted into gray scale images and then binary matrix because skeletonization process works on binary pixel image. The redundant pixels which do not belong to the backbone of the character, were deleted and the broad strokes were reduced.

After skeletonization, normalization process has been applied to normalize the images into desired size of pixels. In training phase, least important units of the ANN are estimated and deleted. Thus the input data is transformed into a set of feature. The PCA based dimension reduction mechanism used in TER is an unsupervised learning algorithm.

While reducing dimension, if it is carefully chosen, it is expected that the feature set will extract only the relevant information from the input data in order to perform the desired task. In the proposed scheme, the sizes of the original drawn images are very large. After reduction, the size of an image has been converted into a 30*33 scale matrix as shown in Figure 2(b). Thereby the input of the SCG is 990.

3) PCA based Dimension Reduction

For dimension reduction, the mechanism of PCA [14] is well-known. It has been applied in TER.

4) SCG algorithm for Training

The characters captured by the webcam have been trained by SCG [17] which is discussed earlier.

5) Major Steps of TER

The block diagram of the proposed TER is shown in Figure 3. Its major steps are described as follows:

a) Tracking of a marker i.e. the head of a pen in front of a webcam.

b) Drawing the path of the movement of the marker. Thus the images of English characters are generated.

c) Converting the images into gray scale images.

d) Converting the gray scale images into a binary matrix data.

e) Applying Skeletonization and normalization process to get the desired size of pixels.

f) Dimension reduction and selection from the obtained data.

g) Training of these obtained data i.e. patterns by using SCG algorithm up to a expected error level.

h) Testing i.e. recognizing touchless written English characters by taking inputs later on.

IV. EXPERIMENTAL STUDIES

In the experiment, we have used 36 English characters as input where each character has 10 samples that make 36 x 10 = 360 samples. All these 360 samples have been used for training. As shown in Figure 4, the error rate decreases as the number of training cycles increases and the curve becomes steady after 2500 training cycles. To gain the best performance, we have chosen 3000 training cycles for our experiment. Table I shows the performance of the proposed TER system.

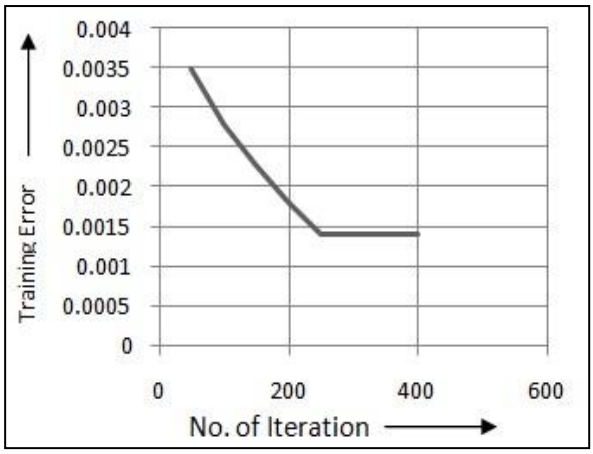


Figure 4. Training error of TER for 360 samples of 36 English characters.

SCG and BP, two different ANN learning algorithms have been used alternatively to develop TER. In the domain of same problem i.e. for three different sizes of inputs, the comparison of training time of TER has been stated in Table II and has been shown in Figure 5. While the no. of pixels in the input domain increases, the training time also increases. Here, BP takes much time than SCG both in training and testing phases. As mentioned in section III, the superiority of SCG method over BP has been found in these figures.

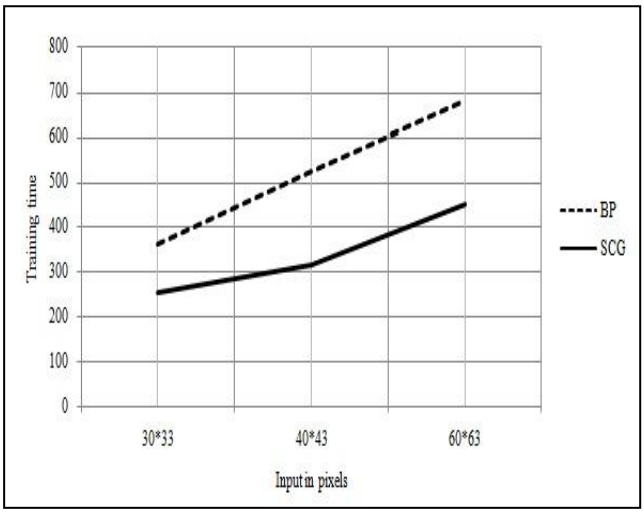


Figure5. Training time comparison of TER between SCG and BP.

A. *Performance of Recognition*

As already mentioned, to evaluate the performance, we have trained TER with 360 samples of pre-defined images of different pixel sizes. Then TER can recognize any 36 of re-drawn English characters.

While training the characters, the SCG ANN learns through adjusting its weight which is a supervised learning algorithm. For example for characters "A", "B", "C", "D", "1", "2" etc, we have sufficiently trained the ANN in an iterative process through adjusting its weights. In TER, the chosen architecture of ANN has two hidden layers with flexible no. of nodes. Its output layer contains 36 nodes because of 26 alphabets and 10 digits of English characters. The network was trained by the SCG algorithm until the root mean square error remains below 0.0015. The weights were updated after each pattern presentation.

The training error of ANN has been estimated according to the following equation.

$$E = \sum_{\tau}^{T} \sum_{k=1}^{N^L} e_k^2(\tau) \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where $e_k(\tau) = d_k(\tau) - o_k(\tau)$ is the difference between the desired and the calculated values at output k at time step τ, and the sums are taken over all T time instants in the training sequence and overall N^L outputs. The simplest method of adapting weights towards the minimum of E is the steepest descent method in which weight movement is in the direction of the negative error gradient scaled with a "learning rate" η:

$$\Delta w_{ij} = -\eta \Delta w_{ij} E$$

Weight adaptation in TER has been performed by SCG [17] which has significantly improved the speed and the convergence of the training as compared to BP. The results are shown in Table I.

TABLE I. RECOGNITION PERFORMANCE OF TER USING SCG AND BP.

No. of outputs = 36	**Iterations**	**Training time (seconds)**	**Accuracy (%)**	
No. of samples 36x10 = 360			***SCG***	***BP***
			Testing	***Testing***
Input pixel size = 30x33	3000	256	97	88
		252	95	84
		251	93	86
	Average	253	95	86

B. *Comparison of TER between SCG and BP*

Table II shows the performance comparison of TER between BP and SCG for three different input pixel sizes i.e. 30*33, 45*48 and 60*63. From here it is found that, in training phase, the performance of SCG is much better than BP.

TABLE II. COMPARISON OF PERFORMANCE OF TER BETWEEN SCGAND BP

Input pixel size	**Approach (time in seconds)**	
	BP	***SCG***
	Training Time	***Training Time***
30*33	360	253
40*43	525	315
60*63	680	450

C. Discussions

In this experiment the drawing process of the images by tracking the marker in front of the webcam is little bit cumbersome and tedious. Better performance can be achieved by the following steps:

- The defined color for tracking and the color of the marker should match perfectly for better tracking.
- Camera pixels should be high enough for a better quality of images.
- Adequate light should be provided while capturing the images.
- Defined tracking color should be unique in the environment to avoid unexpected noise.

V. CONCLUSION

TER system proposed in this paper is an initiative work to recognize touchless written English characters. It is expected that disabled people who knew to write English but later on unable to write by their hand can be benefitted by TER. To implement TER, we have exploited PCA based dimension reduction and SCG algorithm for training that is capable enough to perform a good recognition. TER has been implemented by using SCG and BP alternatively. The experimental results show that SCG method yields a good recognition accuracy approximately 95% which is much better than BP. So far, TER can recognize only touchless written English characters. A future plan of improvement is to employ TER to recognize touchless written English words as well as sentences.

REFERENCES

[1] J.Pradeep, E.Srinivasan and S.Himavathi," Diagonal based feature extraction for handwritten alphabets recognition system using neural network", International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.

[2] Anita Pal1 & Dayashankar Singh, "Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science & CommunicationVol. 1, No. 2, July-December 2010, pp. 141-144.

[3] U. Pal, and B. B. Chaudhuri, "Indian Script Character Recognition: a Survey", Pattern Recognition, 37(9): 1887–1899, 2004.

[4] C. Sureshkumar and Dr. T. Ravichandran, "Character recognition using RCS with neural network", International Journal of Computer Science Issues, Vol. 7, Issue 5, September 2010.

[5] Miguel Po-Hsien Wu, "Handwritten Character Recognition", BSC Thesis, The University of Queensland.

[6] J. Cai and Z. Q. Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition", IEEE Trans. On PAMI, 21(3): 263-270, 1999.

[7] R. Plamondon and S. N. Srihari, "On-line and off-line handwritten recognition: A comprehensive survey", IEEE Trans. on PAMI, 22(1): 62-84, 2000.

[8] P. Wunsch and A. F. Laine, "Wavelet Descriptors for Multi resolution Recognition of Hand-printed Digits", Pattern Recognition, 28(8):1237-1226, 1995.

[9] Z. Chi and H. Yan, "Handwritten numeral recognition using self-organizing maps and fuzzy rules", Pattern Recognition,28(1):36-66, 1995.

[10] K. Kim and S. Y. Bang, "A handwritten numeral character classification using tolerant Rough set", IEEE Trans. On PAMI, 22(9): 923-937, 2000.

[11] M. Turk and A. Pentland, "Eigen faces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[12] K. Ohba and K. Ikeuchi, "Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 9, pp. 1043-1048, 1997.

[13] H. Murase and S. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance", International Journal of Computer Vision, vol 14, pp. 5-24, 1995.

[14] M. Mudrova, A. Prochazka,"Principal component analysis in image processing", Institute of Chemical Technology, Prague.

[15] Md. Saidur Rahman, G.M. Atiqur Rahaman, Asif Ahmed and G.M. Salahuddin, "An approach to recognize handwritten English Numerals for postal automation", ICCIT 2008.

[16] J. Lunden and V. Koivunen, "Robust estimation of radar pulse modulation," in Proc. 6th IEEE Int. Symp. on Signal Processing and Inform. Technology, Vancouver, Aug. 27–30, 2006.

[17] Martin Fodslette Moller. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. Neural Networks, 6:525-533, 1993.

[18] Mozer M. C.; Smolensky P., (1989): Skeletonization: A technique for trimming the fat from a network via relevance assessment. in D. S. Touretzky, Advances in Neural Information Processing Systems 1, pp. 107–115, Ed. Denver.

[19] Chakraborty, G.; Chakraborty, B.; "A novel normalization technique for unsupervised learning in ANN", IEE Transaction on Jan 2000, Volume 11, Issue 1.

[20] Amjad Rehman, Dzulkifli Mohamad and Ghazali Sulong, "Implicit Vs Explicit based Script Segmentation and Recognition: A Performance Comparison on Benchmark Database", Int. J. Open Problems Compt. Math.,Vol.2,No.3,September 2009.

[21] J. Cai and Z. Q. Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition", IEEE Trans. On PAMI, 21(3): 263-270, 1999.

[22] www.wikipedia.org/wiki/English_language