Android based Smart Phone Data Image Recognition

G.Kundavinachiyar , I.Muthumani

Department of Electronics & Communication Engineering,
Alagappa Chettiar College of Engineering & Technology,
Karaikudi, Tamil Nadu, India

Abstract - **This paper focuses on an application that performs cursive handwritten English character recognition in hand held devices. The objective is to make use of the visual capabilities of the built in camera of Android devices to extract cursive handwritten English character. The image taken is further processed by preprocessing technique which is segmented based on the stroke and is recognized with the help of Support vector machine (SVM).**

Key-Words: Support Vector Machines, Character recognition, Android.

I. INTRODUCTION

"Character Recognition" is developed to identify either printed characters or discrete run-on handwritten characters. It is a part of pattern recognition that usually deals with the realization of the written scripts or printed material into digital form. The main advantage of storing these written texts in digital form is that, it requires less space for storage and can be maintained for further references without referring to the actual script again and again. Character recognition has wide applications such as in postal services to sort the mails according to their destination using the addresses that are written on the envelope, in restoring old manuscripts, in digital signature verification and much more.

For cursive character recognition first we have to do Data acquisition. After that Pre-processing and segmentation will make character in proper manner. This is useful for find out the starting and ending boundary of particular character. Next process of this mechanism is to design a dataset for feature extraction and another data set is for to train the classifier using Support vector machine (SVM). In next process different feature extraction methods are applied on input data set and extract feature from it. This extracted feature applied on different classifier which was already trained trough input data set which matches between feature and trained data [1].

II. BLOCK DIAGRAM OF RECOGNITION SYSTEM

The schematic block diagram of cursive handwritten English character recognition system consists of various stages as shown in figure 1.

2.1 Data acquisition
This is the stage where the data under consideration is taken. In this method, the data have been acquired through a camera in the Android mobile. The input data are saved in JPEG formats for further processing [2].
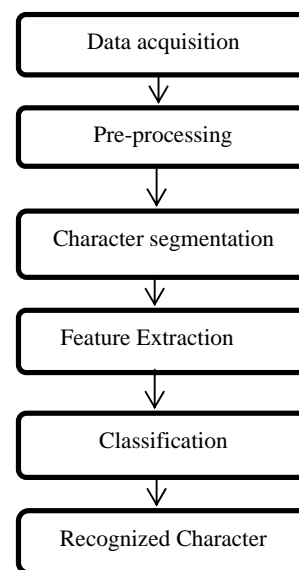

Figure 1. Block diagram of Recognition system

2.2 Preprocessing
The preprocessing stage plays an important role in the handwritten character recognition process. Preprocessing converts the acquired image into a more usable form for the next stages. The major objectives of our preprocessing stage are to reduce the amount of noise present in the document and to reduce the amount of data to be retained. The preprocessing stage includes a number of techniques in order to achieve these objectives, which are now explained [3] [4] [5].

2.2.1 Normalization
The preprocessing begins with data normalization. This is an important step in handwriting recognition because writing styles differ greatly with respect to the skew, slant, height and width of the characters. The acquired image is processed by normalization and the results in regulating the size, position and shape of the characters.

2.2.2 Smoothing
The cursive English handwritten word is smoothed to modify and remove rough edge to obtain more enhanced writing with less noisy shape. The output will be more stable and recognizable as it gives a straight form of the character.

2.2.3Interpolation

The final step in preprocessing is where the smoothened stroke is interpolated to give a fixed number of points, equally spaced along the curve length. The number of points is chosen based on the average number of points per stroke in the given dataset.

2.3 Segmentation

In segmentation, the input image is segmented into individual character and then each character is resized into m*n pixels towards the extracting the features. In this method the characters are segmented based on stroke.

Each character is represented as a combination of strokes. A stroke is defined as the trajectory traced by the pen from a pen-down event to a pen-up event and is represented using the data captured as the stroke is written. The number of points collected varies with the stroke and the speed of writing. The representation of the stroke can be chosen to be of variable length or of fixed length [1] [6] [7] [8] [9].

2.4 Feature extraction and Classification

The two essential sub-stages of recognition phase are feature extraction and classification. The feature extraction stage analyzes a text segment and selects a set of features that can be used to uniquely identify the text segment. The derived features are then used as input to the character classifier [10].

The classification stage is the main decision making stage of character recognition system and uses the extracted feature as input to identify the text segment. Performance of the system largely depends upon the type of the classifier used. Classification is usually accomplished by comparing the feature vectors corresponding to the input text/character with the representatives of each character class, using a distance metric. The classifier which has been used by our system is Support Vector Machine (SVM).
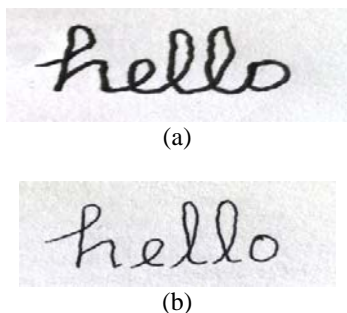

(a)

(b)

Figure.2. (a) Acquired image (b) Image after segmentation

2.4.1 SVM Classifiers

The SVM (Support Vector Machine) was introduced first by Vapnik and co-workers in 1992 [11]. Support Vector Machines are a group of supervised learning methods which can be applied to classification or regression. The SVM classifier accepts the set of input data and predicts to classify them in one of the only two distinct classes. Basically, SVM

classifier is trained by a given set of training data and a model is prepared to classify test data. Depending on how all the samples can be classified in dissimilar classes with appropriate margin, different types of kernel in SVM classifier are used. Frequently used kernels are: Linear kernel, Polynomial kernel, Gaussian Radial Basis Function (RBF) and Sigmoid (hyperbolic tangent) [6] [10] [12] [13] [14].

III. ANDROID IMPLEMENTATION

For the fast evolution of the science and technology, people in the present day have more requirements on mobile communication, Internet for wireless and multimedia amusement. Android is an open-source platform developed by Google and the Open Handset Alliance on which interesting and powerful new applications can be quickly developed and distributed to many mobile device users. For an android implementation the Google Android SDK, the Java Runtime, the Eclipse IDE are all set up on the computer and it will be ready to start writing above programs that can run an Android compatible mobile device [15] [16].

IV. CONCLUSION

In this paper, handwritten cursive character recognition system was implemented for Android smart phones. The implementation process of the system was described to recognize the characters in the document using the camera screen. Photo data taken by a smart phone was compared with the database of the system and the characters were recognized using SVM. The trained support vector machine classifier has shown higher efficiency in terms of speed, memory, and classification accuracy as compared to other related approaches dealing the handwritten character recognition problem. The recognized character was used to create a text file and it is editable by user for their application.

REFERENCES

[1] Parikh Nirav Tushar, Dr. Saurabh Upadhyay, " Chain Code Based Handwritten Cursive Character Recognition System with Better Segmentation Using Neural Network", International Journal of Computational Engineering Research||Vol, 03||Issue, 5||

[2] Zhu Dan, Chen Xu, "The Recognition of Handwritten Digits Based on BP Neural Network and the Implementation on Android", 2013 Third International Conference on Intelligent System Design and Engineering Applications.

[3] Sara Izadi, Mehdi Haji, Ching Y. Suen, "A New Segmentation Algorithm for Online Handwritten Word Recognition in Persian Script", Centre for Pattern Recognition and Machine Intelligence.

[4] Muhammad Naeem Ayyaz, Imran Javed and Waqar Mahmood, "Handwritten Character Recognition Using Multiclass SVM Classification with Hybrid Feature Extraction" Pak. J. Engg. & Appl. Sci. Vol. 10, Jan., 2012 (p. 57-67)

[5] Rajkumar.J, Mariraja K., Kanakapriya,K., Nishanthini, S., Chakravarthy, V.S.," Two Schemas for Online Character Recognition of Telugu script based on Support Vector

Machines" 2012 International Conference on Frontiers in Handwriting Recognition.

[6] Tanzila Saba, Amjad Rehman and Ghazali Sulong, "CURSIVE SCRIPT SEGMENTATION WITH NEURAL CONFIDENCE" International Journal of Innovative Computing, Information and Control -Volume 7, Number 8, August 2011.

[7] Rahul KALA, Harsh VAZIRANI, Anupam SHUKLA and Ritu TIWARI, "Offline Handwriting Recognition using Genetic Algorithm", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010.

[8] Sunny Kumar, Pratibha Sharma, "Offline Handwritten & Typewritten Character Recognition using Template Matching", Sunny Kumar et al./ International Journal of Computer Science & Engineering Technology (IJCSET)

[9] Tanzila Saba, Amjad Rehman and Ghazali Sulong, "CURSIVE SCRIPT SEGMENTATION WITH NEURAL CONFIDENCE" International Journal of Innovative Computing, Information and Control -Volume 7, Number 8, August 2011

[10] R. Dineshkumar, Prof. Dr.J.Suganthi, "A Research Survey on Sanskrit Offline Handwritten Character Recognition" International Journal of Scientific and Research Publications, Volume 3, Issue 1, January 2013

[11] V. Vapnik, The nature of statistical learning theory, Springer, (1995).

[12] Hyeran Byun and Seong-Whan Lee, "Applications of Support Vector Machines for Pattern Recognition: A Survey" S.-W. Lee and A. Verri (Eds.): SVM 2002, LNCS 2388, pp. 213-236, 2002.

[13] Javad Sadri, Ching Y. Suen, Tien D. Bui, "Application of Support Vector Machines for Recognition of Handwritten Arabic/Persian Digits" 2ndMVIP. Feb. 2003. Vol. 1

[14] Shailedra Kumar Shrivastava, Sanjay S. Gharde, "Support Vector Machine for Handwritten Devanagari Numeral Recognition", International Journal of Computer Applications (0975 – 8887), Volume 7– No.11, October 2010

[15] Soon-kak Kwon, Hyun-jun An, Young-hwan Choi, "Character Recognition System Based on Android Smart Phone" International Journal of Modern Engineering Research (IJMER), Vol.2, Issue.6, Nov-Dec. 2012 pp-4091-4093

[16] Zhu Dan, Chen Xu, "The Recognition of Handwritten Digits Based on BP Neural Network and the Implementation on Android", 2013 Third International Conference on Intelligent System Design and Engineering Applications.