Identification of Character from an Image by using Android Mobile Phone

Dr. Vaibhav Sharma
Assistant Professor
S.S. Jain Subodh P.G. College
Jaipur, India
tamvaibhav238@gmail.com

Praveen Choudhary
Assistant Professor
S.S. Jain Subodh P.G. College
Jaipur, India
praveenchoudhary0@gmail.com

ABSTRACT - Identification of character recognition is a difficult assignment in the field of research on image processing, artificial intelligence as machine vision since the handwriting fluctuates from individual to individual. Also, the styles, sizes and its direction make it much progressively complex to interpret the text. The Zip Code recognition and in removing the issue of taking care of archives physically has made it important to secure carefully designed information. This paper introduces the recognition of transcribed characters utilizing android cell phone and direct obtaining of image utilizing Matlab, trailed by the usage of different other Matlab tool kits like Image Processing and so on to process the filtered or gained image. The image is procured by android cell phone that make framework progressively convenient and usable on the grounds that present time is have a place with cell phone.

KEYWORDS - *Image Acquisition, Image Rendering, Character Extraction, Image Processing, Edge Detection, Android*

1. INTRODUCTION

Character recognition is a basic, yet most challenging in the field of example recognition with enormous number of helpful applications. It has been an exceptional field of research since the beginning of software engineering because of it being a characteristic method for associations among PCs and people. All the more absolutely Character recognition is the way toward distinguishing and perceiving characters from the information image and changes over it into ASCII or other identical machine editable structure [1][2].

The procedure by which a computer framework can perceive characters and different images composed by deliver normal handwriting is called handwriting recognition framework. Handwriting recognition is grouped into offline handwriting recognition and online handwriting recognition [3]. On the off chance that handwriting is filtered and, at that point comprehended by the PC, it is called offline handwriting recognition. On the off chance that, the handwriting is perceived while composing through touch pad utilizing stylus pen, it is called online handwriting recognition From the classifier point of view, character recognition frameworks are ordered into two primary classifications for example segmentation (worldwide) and segmentation based (explanatory). The segmentation free otherwise called the comprehensive way to deal with perceive the character without portioning it into subunits or characters. Each word is spoken to as a lot of worldwide features, for example ascender, loops, cusp, and so on. While segmentation based methodology [4]; each word/ligature is sectioned into subunits either uniform or non-uniform and subunits are considered freely.

Handwritten character processing frameworks are area and application explicit, similar to it is not possible to design a generic framework which can process a wide range of manually written texts and language. Bunches of work has been done on European

dialects and Arabic (Urdu) language. While domestic dialects like Hindi, Punjabi, Bangla, Tamil, Gujarati and so forth are less investigated because of restricted utilization.

2. METHODS/ PRINCIPLE

Normally handwritten recognition is divided into six phases which are image acquisition, pre-processing, segmentation, feature extraction, classification and post processing.

a) Image Reading

Digitized/Digital Image is at first taken as information. The most widely recognized of these gadgets is the electronic tablet or digitizer. These gadgets utilize a pen that is computerized in nature. Information images for manually written characters can likewise be taken by utilizing different strategies, for example, scanners, photos or by directly writing in the PC by utilizing a stylus.

b) Preprocessing

Pre-processing is the essential period of character recognition and it's vital for good recognition rate. The principle target of pre-processing steps is to standardize strokes and evacuate varieties that would some way or another confuse recognition and reduce the recognition rate. These varieties or contortions incorporate the irregular size of text, missing focuses during pen movement collections, jitter present in text, left or right

twist in handwriting and uneven separations of focuses from neighboring positions. Pre-processing incorporates five basic steps, in particular, size normalization and focusing, interpolating missing focuses, smoothing, incline correction and re-sampling of points.

c) Segmentation

Segmentation is finished by partition of the individual characters of a image. By and large record is rocessed in a various leveled way. At first dimension lines are portioned utilizing column histogram. From each line, words are separated utilizing section histogram lastly characters are removed from words.

d) Feature Extraction

The fundamental point of feature extraction stage is to remove that example which is most appropriate for classification. Feature extraction procedures like Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Chain Code (CC), Scale Invariant Feature Extraction (SIFT), zoning, Gradient based features, Histogram may be connected to separate the features of individual characters. These features are utilized to prepare the framework.

3. LITERATURE REVIEW

An early striking endeavor in the region of character recognition research is by Grimsdale in 1959. The beginning of a lot of research work in the mid sixties depended on a methodology known as investigation by-combination technique proposed by Eden in 1968. The incredible significance of Eden's work was that he officially demonstrated that every single transcribed character are framed by a limited number of schematic features, a point that was verifiably incorporated into past works. This idea was later utilized in all techniques in syntactic (auxiliary) methodologies of character recognition.

K. Gaurav, Bhatia P. K. [5] Et al, this paper manages the different pre-processing procedures engaged with the character recognition with various sort of images ranges from a directly manually written structure based archives and reports containing shaded and complex foundation and shifted powers. In this, unique preprocessing procedures like slant recognition and adjustment, image improvement strategies of complexity extending, binarization, commotion evacuation strategies, standardization and segmentation, morphological handling systems are talked about. It was inferred that utilizing a solitary system for preprocessing, we can't totally process the image. Be that as it may, even subsequent to applying all the said procedures may impractical to accomplish the full exactness in a preprocessing framework.

Salvador España-Boquera et al [6], in this paper crossover Hidden Markov Model (HMM) model is proposed for perceiving unconstrained offline written by hand messages. In this, the auxiliary piece of the optical model has been displayed with Markov chains, and a Multilayer Perceptron is utilized to evaluate the discharge probabilities.

In [7], an altered quadratic classifier based plan to perceive the offline written by hand numerals of six well known Indian texts is proposed. Multilayer perceptron has been utilized for perceiving Handwritten English characters [8]. The features are removed from Boundary following and their Fourier Descriptors. The character is recognized by investigating its shape and looking at its features that recognize each character. Likewise an investigation has been completed to decide the quantity of shrouded layer hubs to accomplish elite of the back engendering system. An recognition precision of 94% has been ccounted for Handwritten English characters with less processing time.

In [9], corner to corner highlight extraction has been proposed for offline character recognition. It depends on ANN model. Two methodologies utilizing 54 features and 69 features are picked to assemble this Neural Network recognition framework. To look at the recognition proficiency of the proposed corner to corner strategy for feature extraction, the neural system recognition framework is prepared utilizing the level and vertical component extraction strategies. It is discovered that the askew technique for feature extraction yields the recognition precision of 97.8 % for 54 features and 98.5% for 69 features.

A. Brakensiek, J. Rottland, A. Kosmala, J. Rigoll [10] et al, in this paper a framework for offline cursive handwriting recognition is depicted which depends on Hidden Markov Models (HMM) utilizing discrete and half and half demonstrating strategies. Handwriting recognition investigations utilizing a discrete and two distinctive half and half methodologies, which comprise of a discrete and semi-consistent structures, are looked at. A segmentation free methodology is considered to build up the framework. It is discovered that the recognition rate execution can be improved of a half breed demonstrating strategy for HMMs, which relies upon a neural vector quantizer (crossover MMI), contrasted with discrete and crossover HMMs, in light of tired blend structure (mixture - TP), which might be brought about by a relative little informational collection.

R. Bajaj, L. Dey, S. Chaudhari et al [11], utilized three various types of features, to be specific, the thickness features, minute features and spellbinding segment features for arrangement of Devanagari Numerals. They proposed multi classifier connectionist engineering for

expanding the recognition unwavering quality and they acquired 89.6% precision for transcribed Devanagari numerals.

SandhyaArora in [12], utilized four component extraction strategies to be specific, convergence, shadow include, chain code histogram and straight line fitting features. Shadow features are figured internationally for character image while crossing point features, chain code histogram features and line fitting features are registered by partitioning the character image into various portions. On experimentation with a dataset of 4900 examples the general recognition rate watched was 92.80% for Devanagari characters. Mohammed Z. Khedher, Gheith A. Abandah, and Ahmed M. Al Khawaldeh [13] et al, this paper describes that Recognition of characters incredibly relies on the features utilized. A few features of the written by hand Arabic characters are chosen and examined. A disconnected recognition framework dependent on the chose features was manufactured. The framework was prepared and tried with practical examples of written by hand Arabic characters. Assessment of the significance and precision of the chose features is made. The recognition dependent on the chose features give normal exactnesses of 88% and 70% for the numbers and letters, separately. Further enhancements are accomplished by utilizing highlight loads dependent on experiences picked up from the exactness's of individual features.

4. ALGORITHMS

STEP 1. Start process.

STEP 2. Image acquisition.

STEP 3. Preprocessing(Finding Region of Interest)

STEP 4. Segmentation (Separation of characters in blocks)

STEP 5. Feature Extraction (Finding the characters for matching)

STEP 6. End of process.

5. PROCEDURE
a) Input Image

Here image is captured by Android Mobile Phone. This process live and continuous.

Now desktop/laptop in connect to Android Mobile Phone.

- This cell of codes read the image to MATLAB workspace.
- Conversion to gray scale image from RGB image format.

b) Edge Detection

Edges characterize object limits and are in this way valuable for segmentation, registration, and identification of objects. Edge discovery of a image altogether reduce the amount of data and filter through futile data, while protecting the significant basic properties in a image. In any case, most of various strategies might be assembled into two classes, gradient and Laplacian. The gradient technique identifies the edges by searching for the most extreme and least in the first derivative of the image. The Laplacian strategy which is being utilized in this test scans for zero crossing in the second derivative of the image to discover edges.

c) Object Location

Select a single number from the TEST image and the program distinguishes the number. On the off chance that choosing more than one number, the program will distinguish a wrong one, so you should choose a single number.

d) Binarization

Following stage is binarization of gray scale character images in offline character recognition. Binarization procedure changes over a gray scale image into a double image by utilizing the way that characters are normally made out of slight lines of uniform

e) Feature Extraction

The following part is the component extraction wherein manually written numbers must be edited exclusively and changed over to 7x5 bitmap images at that point (35x1) to be as the perfect numbers for the preparation procedure of the NN i.e reshaping the image. Feature extraction is the procedure to recover the most significant information from the crude information. The real job of Feature extraction is to separate a lot of features, which boosts the recognition rate with minimal measure of components. In feature extraction arrange each character is spoken to as a component vector, which turns into its personality. Because of the idea of handwriting with its high level of inconstancy and imprecision acquiring these features, is a troublesome assignment. Feature extraction techniques depend on 3 sorts of features Statistical, Structural, Global changes and minutes. Basic and measurable features seem, by all accounts, to be reciprocal in that they feature various properties of the characters.

6. MATLAB CODE
a) Capturing Image in Matlab by using Mobile Phone live

url = 'http://192.168.1.106:8080/shot.jpg';

ss =imread(url);

fh = image(ss);

```
while(1)

ss  =imread(url);

set(fh,'CData',ss);

drawnow;

end

I = imread('training.bmp'); imshow(I)

Igray = rgb2gray(I); imshow(Igray)

Ibw      =       im2bw(Igray,graythresh(Igray));
imshow(Ibw)

Iedge = edge(uint8(Ibw)); imshow(Iedge)

se = strel('square',2);

Iedge2 = imdilate(Iedge, se); imshow(Iedge2);

Image Filling

Ifill= imfill(Iedge2,'holes'); imshow(Ifill)

 [Ilabelnum] = bwlabel(Ifill); disp(num);

Iprops    =    regionprops(Ilabel);    Ibox    =
[Iprops.BoundingBox];

Ibox = reshape(Ibox,[4 50]); imshow(I)

forcnt              =              1:50
rectangle('position',Ibox(:,cnt),'edgecolor','r'); end
```

7. CONCLUSION

It is trusted this point by point exchange will be valuable knowledge into different ideas included, and lift further advances in the area. The exact recognition is directly relying upon the idea of the material to be perused and by its quality. From different investigations we have seen that choice of pertinent component extraction and characterization system assumes a significant job in execution of character recognition rate. This survey builds up a total framework that changes over filtered images of manually written characters to text records. This material fills in as a guide and update for per users working in the Character Recognition area.

8. REFERENCES

[1] Namrata Dave, "Segmentation Methods for Hand Written Character Recognition", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 8, No. 4 (2015), pp. 155-164.

[2] FlávioBortolozzi, Alceu de Souza Britto Jr., Luiz S. Oliveira and Marisa Morita, "Recent Advances in Handwriting Recognition ",Document Analysis, Editors: Umapada Pal, Swapan K. Parui, Bidyut B. Chaudhuri, pp.1-30.

[3] 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.

[4] Haider A. Alwzwazy , Hayder M. Albehadili2 , Younes S. Alwan3 , Naz E. Islam4, Handwritten Digit Recognition Using Convolutional Neural Networks, International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 2, February 2016.

[5] K. Gaurav and Bhatia P. K., "Analytical Review of Preprocessing Techniques for OfflineHandwritten Character Recognition", 2nd International Conference on Emerging Trends in Engineering & Management, ICETEM, 2013.

[6] Salvador España-Boquera, Maria J. C. B., Jorge G. M. and Francisco Z. M., "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, April 2011.

[7] U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts," Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2007.

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

[9] J. Pradeep, E. Srinivasan and S. Himavathi, "Diagonal Based Feature Extraction For Handwritten Alphabets Recognition System Using Neural Network", InternationalJournal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.

[10] A. Brakensiek, J. Rottland, A. Kosmala and J. Rigoll, "Offline Handwriting Recognition using various Hybrid Modeling Techniques & Character N-Grams", Available at http://irs.ub.rug.nl/dbi/4357a84695495.

[11] Reena Bajaj, LipikaDey, and S. Chaudhury, "Devnagari numeral recognition by combining decision of multiple connectionist classifiers", Sadhana, Vol.27, part.1, pp.-59-72, 2002.

[12] SandhyaArora, "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA, December 2008.

[13] Mohammed Z. Khedher, Gheith A. Abandah, and Ahmed M. AlKhawaldeh, "Optimizing Feature Selection for Recognizing Handwritten Arabic Characters", proceedings of World Academy of Science Engineering and Technology, vol. 4, February 2005 ISSN 1307-6884.