International Journal of communication and computer Technologies, ISSN: 2278-9723

Available at http://www.ijccts.org

Advanced Approaches of Handwritten Digit Recognition Using Hybrid Algorithm

[1]Deepak Kumar Bishnoi , [2]Kamlesh Lakhwani

Suresh Gyan Vihar University, India

Received: 12-09-2012, **Revised:** 18-10-2012, **Accepted:** 23-11-2012 ,**Published online:** 29-12-2012

Abstract

Image processing is an interesting and widely used field in the research area. Handwritten digit recognition is a sub module of the image processing. Although lots of research has been done so for in this field and still continues. This paper suggested a novel approach of off-line handwritten digit recognition. The paper classified the digits into four regions: the left part, right part, upper part and lower part. In these four parts the images are identified on the basis of curve. The curve of a left part image is converted into pixels and as well as rest of all parts. Then these pixels are compared through decision tree and the result of comparison describes the format of digit. This method is used to test the various handwritten digit form modified NIST and MNIST databases, which shows the great success rate.

Keyword: - Digit recognition, Curve matching, Thinning, Smoothing, Regions Classification.

1. Introduction

Writing is an art that has been used for communication not only for human to human but also used for human to machine. Handwritten recognition has an active research field since the advancement of the digitalization. There have been lots of technologies that are under the research field such as networking, image processing, artificial intelligence etc. to increase the speed, accuracy of work. Handwritten recognition is a sub module of image processing which have lots of application for automatic recognizing the handwritten number like zip code reading, bank check reading, licensed plate recognition, security based system, reading bar code and many more[4]. For digit recognition there are two types of techniques applied [5].

1. On-Line Recognition
2. Off-Line Recognition

Here we work on the off-line technique for digit recognition. My work is based on statistical approach.

Offline handwritten digit recognition has been achieved through the four regions algorithm and decision tree algorithm. Basically, there are three parts of my

proposed work which is used to recognition the digits. Pre-processing, segmentation has performed on scanned image. After getting the segmented image my algorithm has been applied, this is used to divide the image into four regions. The next section is used to convert the image of each region into pixels format. The last section uses the decision tree algorithm for the comparison of each region. The tools is used in this process is java; this provides better security and increase the efficiency of the work.

There are four sections to describe the paper. The second section describes the methodology that is applied. The third section is used for the experimental results & discussion and in last section the conclusions is described.

2. Methodology

In English language there are ten numerals and each numeral has a different shapes and characteristics [3]. The main issue is that how to recognize these numeral/digit. The recognition is achieved on the basis of different shapes, similarities of same type of digit and features of the digit.

The numerals are divided into four regions by using the following algorithm which is called as four region algorithm. The following steps are used to divide the segmented image into four regions:

- Selected image (72x72)

- Now, we divide the image into four parts which is start with the coordinates X,Y(0,0) like

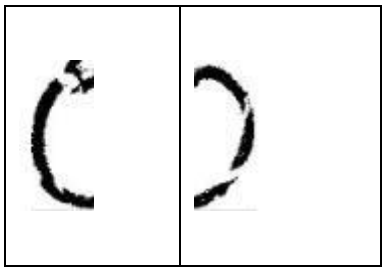

Figure:- Left and Right Part

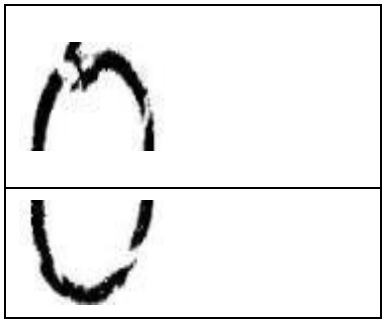

Figure:- Top and Bottom Part

- In order to get these four part of a image (72x72) we used java by applying these logics:-
  - Img = imag.getSubImage (0, 0, 71, 36) for the top part of image
  - Img = imag.getSubImage (0, 36, 71, 71) for the bottom part of image
  - Img = imag.getSubImage (0, 0, 36, 71) for the left part of image

- Img = imag.getSubImage (36, 0, 71, 71) for the right part of image

- After applying this algorithm we got the following patterns which are as follows:-

S.No.	Left Part	Right Part	Top Part	Bottom Part
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				

By analyzing these regions we find out following characteristics:

- Holes/Circle
- Vertical up & down cavities
- Horizontal left &Right cavities
- Right Curve
- Left Curve
- Upper Curve
- Bottom Curve

2.1 Similar regions

Now on the basis of these four regions a decision tree will be made. First of all, the similar regions are identified. The followings regions are identified which is shown in the following table:-

S.No.	Digit	Part	Attribute
1	3,5	Bottom	Right Curve
2	1,7,9	Bottom	Vertical Line
3	2,3	Top	Right Curve
4	2,6,8	Right	Right Curve
5	0,9	Left	Left Curve

To matching these image following algorithm will be used:

1. If we require the pre-processing of image to enhance contrast, noise removal
2. Now segmentation is performed which divide the image into regions and contains the pixels for the region that is similar and different. This has been done with the help of clustering algorithms.
3. A descriptor is needed to calculate the region as well as it includes the

shape, perimeter, circle etc. of the image.

4. If required, then perform Re-Segmentation of the image, where the regions are merged can be representing the same object.
5. If needed, then we can eliminate the regions that are useless for the work.
6. Accumulate the image's regions' descriptor used for auxiliary processing.

Now, a decision tree will be created on the basis of this information's which is as follows:-

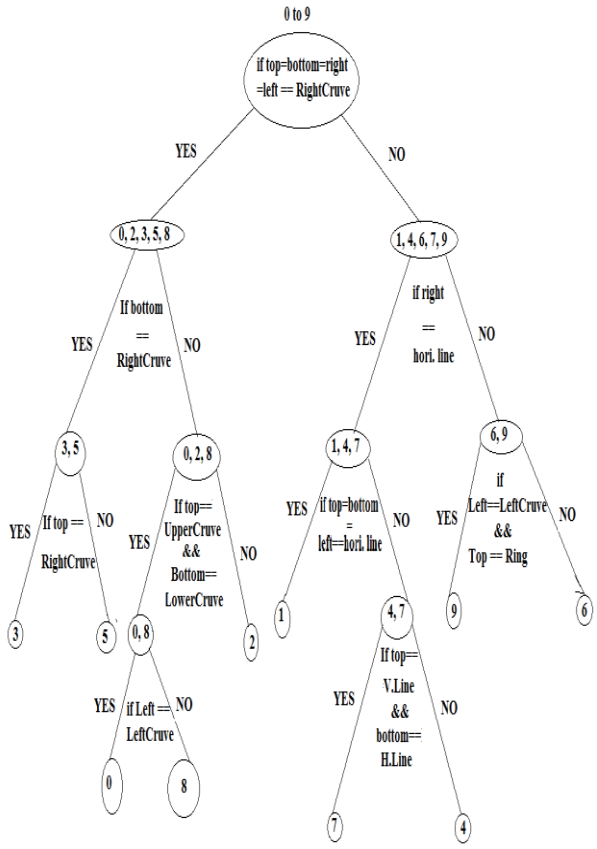

Figure:- A decision tree for recognition of the numerals

3. Results & Discussions

The digits are recognized by the help of decision tree. There are some common bit patterns which are used to classify the numerals into two groups. The region which have the right Curve in the top part, bottom part, left part and right part make a group (0, 2, 3, 5, and 8) called as first group. The group which does not have any curve in any parts (1, 4, 6, 7, and 9) called as second group.

First group which have the following digits: 0, 2, 3, 5, and 8 is further divided into two parts on the basis of right curve found at bottom part. 3 and 5 these digits make a parts and the second part have the 0, 2, and 8. First part (3, 5): if right curve found in the top part then the digit is 3 otherwise the digit is 5.

Second part (0, 2, 8): if we found upper curve at top part and lower curve at the bottom part which have 0 and 8 digits and the other part have the digit 2. Sub module which has 0 and 8: have the left curve in left then it is 0 otherwise it is 8.

Second group having the following digits: 1, 4, 6, 7 and 9 again divided into two sub module. A Module which has horizontal line/cavities at the right part is first sub-module (1, 4 and 7). Second sub-module has 6 and 9 digits. First sub-module which have (1, 4 and 7) again divided into two sub parts in which one part has the vertical line at top,

bottom, left and right part is recognized digit 1 and the other sub part have (4 and 7). If we found horizontal line at top and vertical line at bottom then it is 7 and otherwise it is 4. Second sub-module which have 6 and 9: if we found circle at top and left curve at the left then it is 9 otherwise it is 6.

In this paper we use the Modified national institute of standard and technology (MNIST) database and it have 60,000 training set and 10,000 testing numerals. For these paper 30 handwritten digits sets of MNIST is used. There are 6 training & 6 testing digit are taken form MNIST database. This provides the total number of 300 test numerals for experiments. The recognition of each digit which is successfully done was shown in table.

Table 3: Recognitions of handwritten digits for 30 users

Digits	Success rate of recognition (%)
0	**94.59%**
1	**97.85%**
2	**81.23%**
3	**92.17%**
4	**83.25%**
5	**95.67%**
6	**98.00%**
7	**95.49%**
8	**99.98%**
9	**97.99%**

4. Conclusions

This algorithm shows a good recognition rate of the English numerals in comparison with previous algorithms. A tool is made by using this algorithm which is implemented in java.

This algorithm uses a novel technology of regions which shows a great efficiency in the digits recognition because it work only on some part of image not on whole image which shows the faster recognitions. Regarding this specifics rules are describe in second section. All digits average success rate is above 91.21% which is a good recognition rate. This mechanism may decrease the proficiency of the algorithm when it found a rough written digits or digit with extra storks. This work can be implemented in the future.

5. References

1. Miran Karic, Goran Martinovic "Improving Offline Handwritten Digit Recognition Using Concavity-Based Features" INT J COMPUT COMMUN, ISSN 1841-9836 8(2):220-234, April, 2013.
2. Nishant Gupta1, Arun Sharma "A Hybrid Approach To Optical Numeral Recognition (Composition Of PCA & LVQ)" in International Journal of Engineering Research & Technology (IJERT) ,Vol. 2 Issue 1, January- 2013.

3. V VIJAYA KUMAR, A SRIKRISHNA, B RAVEENDRA BABU and M RADHIKA MANI "Classification and recognition of handwritten digits by using mathematical morphology" in S¯adhan¯a Vol. 35, Part 4, August 2010, pp. 419–426. © Indian Academy of Sciences.
4. Rahul Kala, Harsh Vazirani, Anupam Shukla And Ritu Tiwari "Offline Handwriting Recognition using Genetic Algorithm" in IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010.
5. ReÂ jean Plamondon, Sargur N. Srihari "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey" in IEEE transactions on pattern analysis and machine intelligence, vol. 22, no. 1, january 2000.
6. Faisal Tehseen Shah, Kamran Yousaf, "Handwritten Digit Recognition Using Image Processing and Neural Networks" in Proceedings of the World Congress on Engineering 2007 Vol I WCE 2007, July 2 - 4, 2007.
7. El Kessab, C. Daoui, B. Bouikhalene, M. Fakir2 and K. Moro "Extraction Method of Handwritten Digit Recognition Tested on the MNIST Database" in International Journal of Advanced Science and Technology Vol. 50, January, 2013.
8. Albiol Antonio, Mossi J M, Albiol Alberto, Naranjo V 2004 Automatic license plate reading using mathematical morphology.
9. Neeraj Pratap and Dr. Shwetank Arya "A Review of Devnagari Character Recognition from Past to Future" in International Journal of Computer Science and Telecommunications [Volume 3, Issue 6, June 2012].
10. Marwan A. Abu-Zanona, Bassam M. El-Zaghmouri "Current Arabic (Hindi) Hand Written Numbers Segmentation and Recognition Advance Image Processing and Neural Network in Journal of Emerging Trends" in Computing and Information Sciences [VOL. 3, NO. 6, June 2012].