Rotation, Scale and Font Invariant Character Recognition System using Neural Networks

Nitin Khosla[1] and Dr.Sheifali[2] Gupta

[1]Chitkara University, Banur, Punjab, INDIA

[2]Chitkara University, Banur, Punjab, INDIA

Abstract:

It is observed that the text extraction suffers from the drawback of size and rotation of the text appeared on the images and the scanning device has to be focused on the textual area of the image. This engages the person using the application software. However, this can be automated if the algorithm is designed in way so as to identify the text area on the image either at some orientation or at different sizes. In this paper an algorithm is proposed which can very effectively extract characters from the image, irrespective of the size, rotation and font of the character. The system is first trained using extracted features of each character and then the system is tested, if the system properly identify all the characters no further training is required otherwise the system is trained again.

Index Terms— **Character Recognition, Feature Extraction, Neural Network, Back propagation.**

I. Introduction

Text extraction from images finds application in most of the documents related entries in offices and the most popular application is in libraries where no. of books are entered daily by typing the book title along with its author name and other attributes. This can be made easier by using a suitable algorithm or software application that can extract the text from the book cover and present in a text file thereby reducing the typing job of the user. Now he needs only to arrange the book title and authors name and etc. by formatting the material.

The problem with the traditional Character Recognition System is that the characters in the image have to be of a proper size and properly oriented and they should not be rotated, if the characters are rotated and varying in size it results in degradation of the performance of the system and system produce error. Also the system should identify characters of different fonts effectively. To overcome this problem an effective algorithm is required to make the system invariant of rotation, scale and font of character. In this research, a system is proposed which can detect characters effectively from an image irrespective of the size, font and rotation of the characters in the image .

The system can be design by using combination of Neural network and image processing techniques. In the proposed system the Neural Network is trained to remember the feature of each English character using multilayer perception (MLP) network model. The network undergoes supervised training, with a finite number of pattern pairs consisting of an input pattern and a desired or target output pattern. An input pattern is presented at the input layer. The neurons here pass the pattern activations to the next layer neurons, which are in a hidden layer. The outputs of the hidden layer neurons are obtained by using perhaps a bias, and also a threshold function with the activations determined by the weights and the inputs.

II. Related Works

Techniques for page segmentation and layout analysis are broadly divided in to three main categories: top-down, bottom-up and hybrid techniques [1]. Many bottom-up Approaches are used for page segmentation and block identification [5], [11].Yuan, Tan [2] designed method that makes use of edge information to extract textual blocks from gray scale document images. It aims at detecting only textual

ISSN: 2278 – 1323

regions on heavy noise infected newspaper images and separate them from non-textual regions.

The sequence of separating the text from the image is basically fall into two catagories:statical methods and syntactic methods. First category include techniques like template matching, measurement of density of point in a region, moments ,characteristic loci, and mathematical transformation .In the second category the process aims at collecting the effective shape of the numeral generally from contours.

The White Tiles Approach [3] described new approaches to page segmentation and classification. In this method, once the white tiles of each region have been gathered together and their total area is estimated, and regions are classified as text or images. George Nagy, Mukkai Krishnamurthy [4] have proposed two complementary methods for characterizing the spatial structure of digitized technical documents and labelling various logical components without using optical character recognition. Projection profile method [6], [13] is used for separating the text and images, which is only suitable for Devanagari Documents (Hindi document).

The main disadvantage of this method is that the irregular shaped images with non-rectangular shaped text blocks may result in loss of some text. They can be dealt with by adapting algorithms available for Roman script. Kuo-Chin Fan, Chi-Hwa Liu, Yuan-Kai Wang [15] have implemented a feature based document analysis system which utilizes domain knowledge to segment and classify mixed text/graphics/image documents. This method is only suitable for pure text or image document, i.e. a document which has only text region or image region. This method is good for text-image identification not for extraction.

The Constrained Run-Length Algorithm (CRLA) [14] is a well-known technique for page segmentation. The algorithm is very efficient for partitioning documents with Manhattan layouts but not suited to deal with complex layout pages, e.g. irregular graphics embedded in a text paragraph. Its main drawback is the use of only local information during the smearing stage, which may lead to erroneous linkage of text and graphics. Kuo-Chin Fan, Liang-Sheen Wang, Yuan-Kai Wang [17] proposed an intelligent document analysis system to achieve the document segmentation and identification goal. The proposed system consists of two modules: block segmentation and block identification. Two kinds of features, connectivity histogram and multi resolution features are extracted.

III. Proposed Solution Algorithm

The presented work deals with the automatic visual inspection of the geometric patterns to recognise and classify. The algorithm basically involves extracting the features of each character effectively and then uses these features for neural network training. Following is the flowchart of the algorithm followed in order to achieve the objectives:

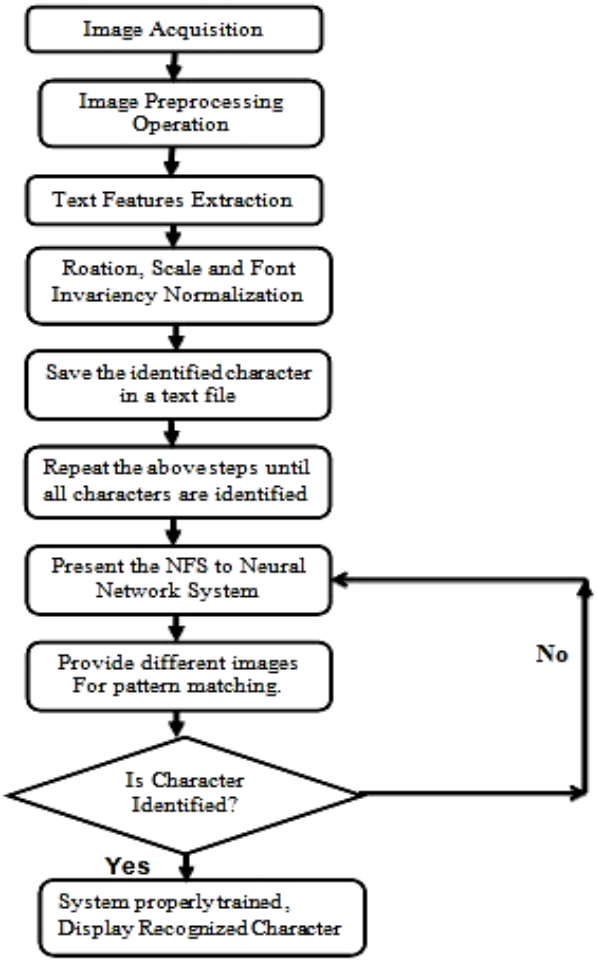

Fig. 1.Flowchart of the algorithm

Image acquisition is the very first step while conversion of text from scene to a textual information. The image of the text matter is either scanned or clicked using the digital camera. The acquired image is in jpeg format i.e. 24-bit colour format. The same is converted to 8-bit colour format i.e. gray color format using the rgb2gray command in

ISSN: 2278 – 1323

matlab. The gray image is now binarized using the Otsu algorithm. This gives an binary image consisting of black color text material on white back ground.

Now the text material is segmented using the labelled image conversion process. The binary text material or image is scanned for the grouped pixel in black using the 8-connectivity in a 3x3 kernel.

The segmented characters are now exposed to feature extraction process. The binary image so obtained is now made rotation invariant by computing the angle of orientation using second order moments as follows:

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

The angle α is given by:

$$\alpha = \text{Tan}^{-1}(\sum(X^2 + Y^{2)}/\sum X.Y)$$

where X2, Y2 and X.Y are the sum of square of x and y coordinates and product of x and y coordinates respectively.

Where (X',Y') is the new co-ordinate when axis is rotated at an angle α , with respect to old (X,Y) Co- ordinate axis[5].

The α oriented image is divided into four quadrants around the centre of gravity (COG). The COG is computed using the first order statistical moments as follows:

$$Gx = (1/N) \sum Xi$$
$$\text{And } Gy = (1/N) \sum Yi$$

Where (Gx,Gy) is the co-ordinate of COG and (Xi,Yi) and 'N' are the co-ordinates of i_{th} pixel and total no. of pixel on the object respectively.Now, maximum and minimum radii and intercepts on axes are computed in each quadrant using the following equation:

$$R = \sqrt{\{ (G_x - X_1)^2 + (G_y - Y_1)^2 \}}$$

Where (G_x,G_y) and (X_1,Y_1) are coordinates of COG and pixel on contour of the pattern.

The perimeter and character area is computed by counting the pixels on boundary and total pixels on the character portion. The perimeter is find out by first extracting the boundary of each character .The boundary of each character is by using the following command in matlab,

```
[B L]=bwboundaries(~BW,'noholes')
```

This command extract the boundary pixel only leaving behind the pixels which are present inside the object and the boundary pixels are added to find the perimeter of each character.

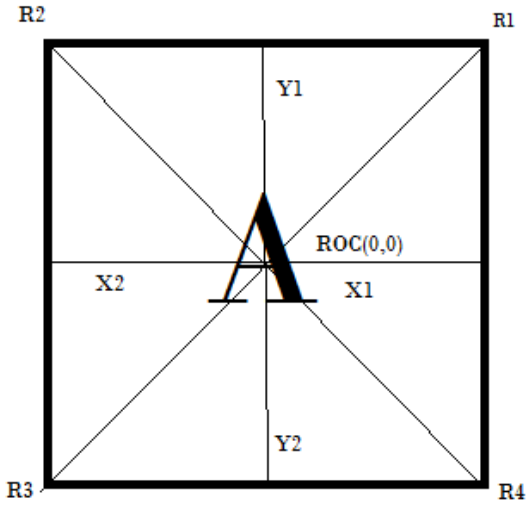

Fig. 2 Segmentation of image into 4 different Quadrant

The extracted features are divided by the mean radius in order to normalize the features size invariant. The features are already location independent as all the features are computed around the centre of gravity of the character under study.

IV. Character Identification

The character features are now normalized between 0 and 1 by dividing the features by the maximum value before inputting them to a neural network classifier. Normalization plays a very vital role in character identification. The features act as input to the neurons, the number of neurons are equal to the number of the normalized feature vector of the each character, once the system is trained, it can be used for validation purposes.

ISSN: 2278 – 1323

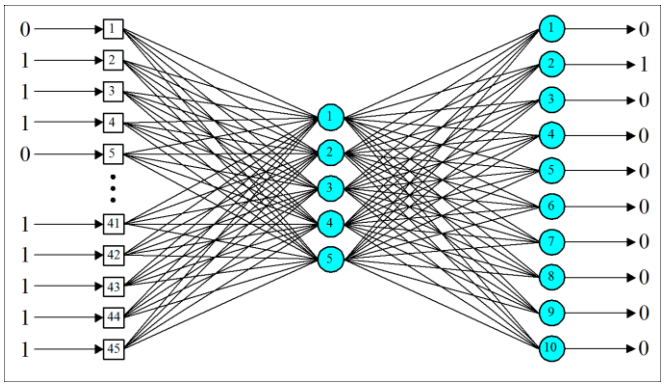

Fig. 3 Extracted feature set act as input to the neural network

. The system is trained by using back propagation algorithm.A neural network is trained for different shapes and style of the each character for proper training of the neural network and proper adjustment of weights. Once the neural network is trained for each of the character in different style or fonts, the system is ready for application in field. Before, putting into application, the same is tested on some test images.

V.Conclusion

The presented algorithm is an approach for extraction of text from the scene and can be extended to application like in library, word processing, and image to text converters. Some existing approaches for the same are studied and found the disadvantage of size, location and rotation invariance. In the proposed approach, invariance's are inserted into the features extraction process so that the characters are normalized in all respect. Once the features are set and are constant for the same object in either form, then the features may be used for identification purposes. The algorithm is developed in matlab version 7.5 and can be converted to any front end language for commercial applications

VI.References:

1) Sung-Bae cho"Neural-Network Classifiers for RecognizingTotallyUnconstrained Handwritten Numerals" IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 1,2013

2) Swapnil Khedekar Vemulapati Ramanaprasad Srirangaraj Setlur, "Text - Image Separation in Devanagari Documents", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03), 2003

3) Q. Yuan, C.L. Tan, "Text Extraction from Gray Scale Document Images Using Edge Information". Proceedings of the International Conference on Document Analysis and Recognition, ICDAR'01, Seattle, USA, pp. 302- 306 September 10- 13, 2001

4) A. Antonacopoulos and R T Ritchings "Segmentation and Classification of Document Images", The Institution of Electrical Engineers, U.K. 1995

5) G. Nagy, S. Seth, "Hierarchical Representation of Optically Scanned Documents". Proceedings of 7th Intl. Conf. on Pattern Recognition, Montreal, Canada, pp. 347-349, 1984.

6) Kyong-Ho Lee, Student Yoon-Chul Choy, and Sung-Bae Cho, "Geometric Structure Analysis of Document Images: A Knowledge-Based Approach", IEEE transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, November 2000.

7) Jean Duong, Hubert Emptoz, "Features for Printed Document Image Analysis", IEEE Transaction on Document Image Analysis Vol. 02, No. 17 November 2002.

8) Nikoluos G. Bourbakis, "A Methodology of Separating Images from Text Using an OCR Approach", Binghamton University, Center for Intelligent Systems, Binghamton, NY 13902.

9) Karl Tombre, Salvatore Tabbone, Loïc Pélissier, Bart Lamiroy, and Philippe, "Text/Graphics Separation Revisited" Proceedings of the 5th International Workshop on Document Analysis Systems, 2002.

10) V Chew Lim Tan, Zheng Zhang, "Text block segmentation using pyramid structure", Proceedings of SPIE, 4307, 297, 2000

11) R. M. K. Sinha, Veena Bansal, "On Devanagari Document Processing", IEEE International Conference on System Man, and Cybernetics, 22-25 October 1995.

ISSN: 2278 – 1323

12) P. Mitchell, H. Yan, "Newspaper document analysis featuring connected line segmentation", Proceedings of International Conference on on Document Analysis and Recognition, ICDAR'01, Seattle, USA, 2001.

13) T. Perroud, K. Sobottka, and H. Bunke, "Text extraction from color documents clustering approaches in three and four Dimensions", Sixth International Conference on Document Analysis and Recognition (ICDAR'01), pp 937-941, September 10-13, 2001.

14) Abdel Wahab Zramdini and Rolf Ingold, "Optical font recognition from projection profiles", Proceedings of Electronic Publishing. Vol. 6(3), pp.249– 260, September, 1993.

15) Dimitrios Drivas and Adnan Amin, "Page Segmentation and Classification Utilizing Bottom-Up Approach", IEEE International Conference on multimedia and expo, Taipei, Taiwan, June 1995.

Author Profile

Nitin khosla is pursuing M.E Fellowship (2012-2015) in Electronics and Communications from Chitkara University, Punjab.He has been teaching in Chitkara University, Himachal Campus since May, 2012. His research interests are in Digital Image Processing using Matlab and Labview,Neural Network.

Prof. (Dr.) Sheifali Gupta is an Associate Professor in the Department of Electronics & Communication Engineering at Chitkara University, Punjab Campus-India. She has got over fourteen years of full-time teaching experience. Dr. Sheifali Gupta specializes in the Digital Image Processing & appli - cation of software Tools in engineering education. Experimental studies in support of the computational work are also part of her research agenda. She has worked with both undergraduate and graduate students throughout his research career and plans to continue to involve students in her research and is eager to participate in senior design projects and guide independent student research. She has published more than 30 research papers and articles in National & International journals and conferences.

ISSN: 2278 – 1323