International Journal of Theoretical & Applied Sciences, 10(1): 106-109(2018)

ISSN No. (Print): 0975-1718
ISSN No. (Online): 2249-3247

Orientation, Scale and Location Invariant Character Recognition System using Neural Networks

Abhisehek Anand[1] and Pankaj Bhambri[2]
[1]Research Scholar, I.K.G. Punjab Technical University, Kapurthala (Punjab), India
[2]Assistant Professor, Department of Information Technology,
Guru Nanak Dev Engineering College, Ludhiana (Punjab), India

(Corresponding author: Abhisehek Anand)
(Received 03 January 2018, Accepted 15 March, 2018)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Textual character recognition systems often suffers from the drawback of the character's size ad their orientation. The zoom-in or out effect due to camera position introduce the size issues and that needs to be taken care off as the character's features vary when size varies. The same case is with the orientation of the character due to rotation of the text paper. Therefore, the size and orientation are the two prime issues that need to be normalized so as to maintain the character features set invariable irrespective of size and rotation. The size invariance can be achieved by using the mean radius of the character measured around the centre of mass of the character and orientation can be normalized by using the orthogonal set of coordinate system.

I. INTRODUCTION

Text extraction in an image is an important task when dealing with large number of images for textual information extraction. The images possess different text or same text in different font's style either in size or format. Therefore, the character identification task has to take care of different aspect regarding the character's shape.

Pattern recognition in image segmentation is been a popular topic of research within engineering. It contains a wide range of image processing problems of great practical use, which range from voice recognition and the classification of documented image, to fault detection in engineering equipments and medical diagnosis. Various research has been done towards pattern recognition in recent years. Some developments are fuzzy logic, neural networks and genetic algorithm. Neural network come up as a powerful tool for pattern recognition and image segmentation. But straight forward neural approach is not adequate for complex patterns such as documented images recognition. In order to solve this problem we need some sophisticated neural network classifiers like multiple multilayer perceptron (MLP) classifier, structure adaptive self-organising map (SOM), and hidden Markov model(HMM).Among above classifiers, the multiplayer perceptron and self-organising map have been most popular and powerful tool for pattern recognition. Their main strength is the capability to learn in various cycle of training and represent implicit knowledge.

The text identification task using image processing techniques requires certain set of features that should be unique and normalized in all respect so as to uniquely distinguish a character from other one. The features set discussed in the presented work are derived using the statistical centre of mass of the character and then dividing the character into four quadrants. Besides feature extraction, one primary task is to extract the character from its image i.e. segmentation of the image with respect to character present in an image. A threshold is used to get the color of the character and then the image is thresholded so that the characters are extracted. Once the characters are segmented and collected in individual frames, then the feature extraction algorithm is applied to extract the features.

The proposed system consists of two parts: in one part the features set is presented to the neural network for training of the neural network system. Once the system is trained for all the characters, the same may be used for application character recognition using the adjusted or trained weights. The weights are basically the neural network intelligence gained during the training part.

II. RELATED WORKS

In page segmentation technique and their layout analysis, the complete analysis is mainly categorised in to three techniques: top-down, bottom-up and hybrid techniques [1]. In page segmentation and block identification designed method we use many types of bottom-up Approaches [5], [11]. Yuan, Tan [2] which invokes textual blocks as input information using edge information from gray scale document images. Its aimis to extract from the newspaper images which are heavy noise infected and separate the required textual region from graphical region.

The White Tiles Approach [3] describes a very noble technique to page segmentation and its various classification. In above method, the classification of documented images are done after the collection and total area estimation of white tiles of each region all together. George Nagy, Mukkai Krishnamoorthy [4] have proposed two methods without using optical character recognition for characterizing the spatial structure and labeling various logical components. Characterizing spatial structure of digitized technical documents.

Projection profile method [6], [13] is applicable for segmenting the text or documented images, which is only applicable for Devanagari Script (Hindi document).

The major drawback of above method occurs when we comes up with random shaped images and non-rectangular shaped text blocks. This non-rectangular shape text blocks could lead to the loss of some textual script. However this drawback can be controlled by applying algorithms applicable for Roman script. Chi-Hwa Liu, Kuo-Chin Fan, Yuan-Kai Wang [15] have came up with an idea and implemented a document analysis system which is featured based. This system of document analysis utilizes the main domain knowledge to segment and separate heavy noise infected newspaper images which are mixture of text/graphics/image documents. The above method is only relevant where blocks contains only mixture of text or documented images, i.e. a set of document which is highly noise infected but contains only text or image portion. The above method is best for documented image segmentation and identification only, as is not suitable for extraction of images.

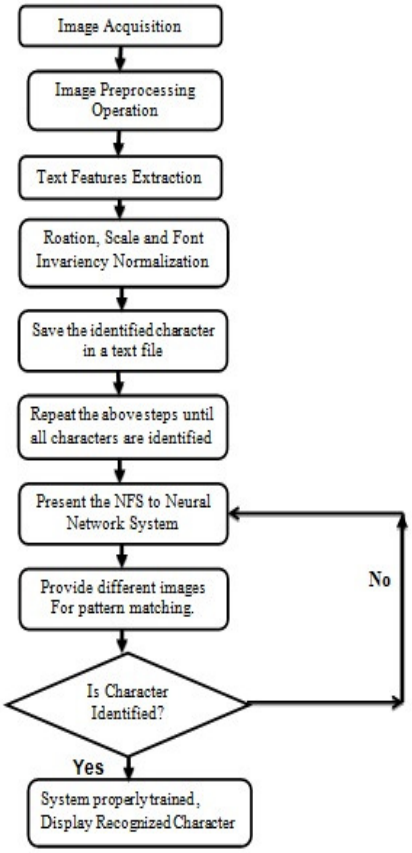
The Constrained Run-Length Algorithm (CRLA) [14] is a famous algorithm for page segmentation techniques. With the help of Manhattan layouts the above algorithm is extremely economical for segmenting documents however not suitable for complex layout pages, e.g. Complex combination of graphics and images which are embedded in a text paragraph. Its main disadvantage is the access of solely native information of documented images throughout the smearing stage, which can cause incorrect intermingling of text and graphics. Laing-Shen Wang, Kuo-Chin Fan, Yuan-Kai Wang [17] introduces an analysis system which is intelligent for the analysis of document to achieve the segmentation of documented images and identification goal. This document analysis system separated into two modules: first is block segmentation and secondly block identification. These two modules helps in achieving two features, which is multi resolution features and connectivity histogram.

I. ALGORITHM

The present work deals with the geometric patterns to recognise and separate using automatic visual inspection. Following are the steps followed in order to achieve the objectives:

The image is acquired using the digital or web camera and stored in jpeg format i.e. colored or RGB image. The jpeg image is then converted to gray scale image (8-bit color format) and exposed to thresholding operation in order to get the binary image. The binary image so obtained is now segmented for textual area identification. Once the text area is localized. The focus is shifted to text/character segmentation. The textual area is extracted based on text density in a particular region. The text characters are segmented using the local neighbourhood connectivity of pixel suing the bwarea open command in matlab. Further, it is important to note the distance between the characters and word. Normally, the distance between two characters of a word is less than the distance between two words. This helps in identifying the characters and text or words in the image.

Now by applying orthogonal transformation the character is made rotational independent. After this, computation of statistical features like, intercepts and maximum radii is done from the image. Maximum radii of the image block is calculated in each quadrant and intercepts is also calculated on each axis. Many other parameters likes Perimeter, standard deviation, figure aspect and area of the image block is also computed from the image. The set of data of various statistical parameters of each image pattern is collected and saved.



Then the obtained set of data is normalised to mean radius and further saved and stored for classification and categorization purposes. The feature set is normalized w.r.t. size, rotation and location so that if the camera or object is varied, the features remain constant for the repeated object. The NFS (normalized feature set) is presented to the neural network in order to train the system. Once the system undergoes complete training, it can be used for validation purposes. The validated system is now ready to be used in application. The identified characters are stored in a text file, from there the same may used in application.

The text file is the result file that can be used for word processor.

Repeat the above steps until all characters are identified. The above steps are repeated till all characters are identified and stored in text file.

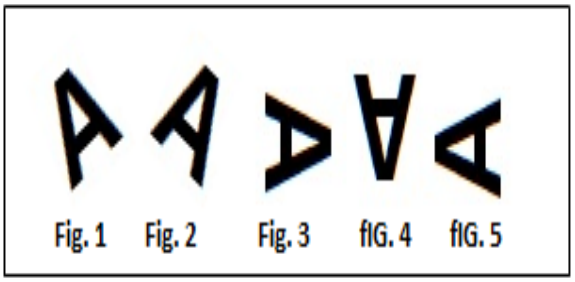
Present the text file as result for further processing. The text file so obtained is now ready to be used for word processing or we can say that is the result file.

III. RESULTS AND DISCUSSION

The feature table for the characters A – Z and numerals 0 – 9 is given in table-1. A back propagation neural network is trained using the extracted features as input neurons and alphabets A-Z and numerals 0-9 as output. Following back-propagation neural network is designed for training of the network:

Input Neurons = 20 (Features)
Learning Rate (η) = 0.75
No. of Hidden Layer = 1
No. of Hidden Layer Neurons = 2
No. of Outputs = 36 (A-Z + 0-9)
Target MSE = 0.00001
Training Set = 180

The training set includes five different font size for each A-Z and 0-9 thereby making (26x5 + 10x5) 180 training sets. After completion of the NN training, the NN was tested by giving different input characters and found to identify with accuracy above 97%,

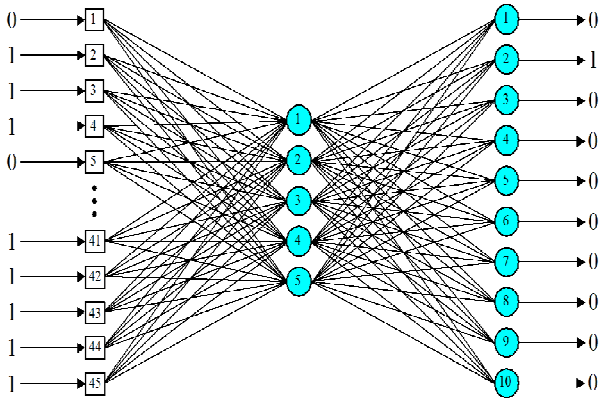


The table 1 in result section shows the first 20 radii in descending order along with mean radius, perimeter, standard deviation and Euler no. for the character 'A' rotated at different angles as shown in fig. 1 to 5. It can be observed from the table that the radii, mean radius, perimeter, std. Dev. And Euler no. are very close to each other, thereby proving the efficiency of the propose algorithm. Same effect is observed with same character in different font style.

By dividing the features by maximum value the character features are stored and then normalized between 0 and 1 for inputting them to a neural network classifier.

The normalized feature set is presented to the neural network in order to train the system.The normalized features will act as input to the neurons of neural network. Here the count of neurons as input are same as the count of the normalized feature vector for the each input character.

Sr. No.	Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5
R1	1.5944281	1.5462864	1.5555151	1.5545055	1.5595873
R2	1.5452849	1.5462864	1.5555151	1.5545055	1.5191996
R3	1.5345495	1.4972632	1.5555151	1.5545055	1.5191996
R4	1.4994348	1.4771985	1.5152329	1.5182238	1.4818075
R5	1.465986	1.4517276	1.5152329	1.5142494	1.4818075
R6	1.4407085	1.4517276	1.4779384	1.5142494	1.4797594
R7	1.4407085	1.4375263	1.4779384	1.4769791	1.4797594
R8	1.4266149	1.4100175	1.4758956	1.4749377	1.4413445
R9	1.399315	1.4100175	1.4758956	1.4749377	1.4413445
R10	1.399315	1.3927161	1.437581	1.4408365	1.4040391
R11	1.399315	1.3846579	1.437581	1.436648	1.4040391
R12	1.3620642	1.3656705	1.400373	1.436648	1.4040391
R13	1.3553045	1.3519448	1.400373	1.3994641	1.4040391
R14	1.345784	1.3519448	1.400373	1.3994641	1.3679339
R15	1.341683	1.3491829	1.400373	1.3994641	1.3679339
R16	1.3389421	1.3450294	1.3643621	1.3634765	1.3679339
R17	1.3293043	1.3394714	1.3643621	1.3634765	1.3679339
R18	1.3293043	1.3394714	1.3643621	1.3634765	1.3331264
R19	1.3293043	1.3394714	1.3643621	1.3634765	1.3331264
R20	1.3293043	1.315588	1.3296455	1.3287825	1.3331264
Perim.	220.898	225.74	237.994	237.839	235.498
MeanR	1.58	1.546	1.539	1.539	1.512
SD	0.052	0.49	0.051	0.052	0.048
Euler No.	0	0	0	0	0



Once the system completes its training cycle and is trained, it is ready for the use and can be used for validation purposes. Here back propagation algorithm is used to train the system.

A complete system of neural network undergoes training for various shapes, size and style of the same character blocks. Together with the proper training of the neural network proper adjustment of weights is done. After complete training of the neural network for each of the character in different style, size and fonts, the output is saved and stored which makes system ready for application in field. Before, putting the system into application, the system is tested repeatedly on the same test images.

IV. CONCLUSION

The neural network trained using the presented feature set for text identification or classification shows a fair distinction between characters. This proves the degree of uniqueness of the feature set used to characterize a character. There are few characters that has hair line difference in features due to inherent shape of the characters like 'c' and 'e', and on some part this introduces errors in classification if noisy images are presented to the system.

REFERENCES

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

[2]. Q. Yuan, C.L. Tan, (2001). "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.

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

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

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

[6]. Jean Duong, Hubert Emptoz, (2002). "Features for Printed Document Image Analysis", *IEEE Transaction on Document Image Analysis,* Vol. **02**, No. 17 November 2002.

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

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

[9]. V Chew Lim Tan, Zheng Zhang, (2000). "Text block segmentation using pyramid structure", Proceedings of SPIE, 4307, 297.

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

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

[12]. T. Perroud, K. Sobottka, and H. Bunke, (2001). "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.

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

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

[15]. Kuo-Chin Fan, Chi-Hwa Liu, Yuan-Kai Wang, (1994). "Segmentation and classification of mixed text/graphics/image documents", Pattern Recognition Letters, Elsevier, pp. 1201-1209, December 1994.