iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** I **Month of publication:** January 2018

DOI: http://doi.org/10.22214/ijraset.2018.1216

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

A Survey on Easy OCR Techniques used to build Systems for Visually Impaired People

Aishwarya Karnawat[1], Komal More[2], Tejashree Rade[3], Bhagyashree Rane[4], Manoj Mulik[5]
[1, 2, 3, 4, 5]*Department of Computer Engineering, ABMSP's Anantrao Pawar College of Engineering and Research, Pune, Maharashtra, India.*

Abstract*: *Recently, Optical Character Recognition is usually being employed in most of the easy approach. OCR techniques facilitate you to convert pictures, PDF files and numerous different kinds of documents into editable and \searchable information. Acceptable character recog-nition from all variations of pictures could be an up to date drawback. Various tools, algorithms and implement-tations are obtainable to discover characters from pictures. This survey paper presents a quick summary on favoured OCR techniques like matrix matching, feature extraction; neural network primarily based OCR and discusses OCR software system Tesseract.

Keywords*: OCR, text extraction, text recognition, image processing, character recognition*

I. INTRODUCTION

Optical Character Recognition is that the method of changing pictures, scanned documents, picture of a document etc. into digitally written text. It's conversion of text within the pictures into American Standard Code for Information Interchange type at or the other machine editable form which might be used for exchange of data [1]. OCR includes a wide selection of application from automatic range plate recognition to pen computing. It's a field of analysis in computing and pattern recognition. OCR doesn't cope with quality and sharpness of characters [1]. To overcome the restrictions of OCR, OCV comes into image [1]. This paper discusses the techniques used for OCR in several of the systems designed for the visually impaired folks. The foremost steps in OCR method are:

A. Pre-processing

during this step reduction of noise, correction of orientation, entropy changes, image binarisation etc. are disbursed.

B. Exactly, it's the method of assignment

Same label to pixels sharing sure characteristics. In this, the image is divided into multiple segments. Feature Extraction
It's the foremost vital step within the OCR method for economical recognition of character.

C. Post-processing

Error correction, descriptive linguistics correction, spell check etc. are exhausted this step.

II. HISTORY

1) 1870s – C. R. Carey invented retina scanner which used patterns of photocells for image transmission [2].
2) 1912 – Edmund Fournier d'Albe developed Optophone [6]. Optophone is a handheld scanner. It scans a printed page, and then produces tone which represents a specific character. The character is interpreted by the blind person by understanding the tone.
3) 1916 – John B. Flowers gets patent on "One-Eyed Machine Stenographer". The machine could read and type a script. It superimposed all the letters to identify any character [3].
4) 1921 – Ciro Codelupi invents "Reading machine for the blind", which transformed luminous sensations into tactile sensations [4].
5) 1929 – Gustav Tauschek invented a OCR device- "Reading Machine". It had a photo-sensor which pointed light on words which matched a content template in the memory [5].
6) 1931 – Emanuel Goldberg patents Statistical machine which read characters and converted them into standard telegraph code. It was later acquired by IBM [2].
7) 1954 – First OCR reading machine was installed at Reader's Digest in 1954. It converted typewritten sales reports into punch cards. These punch cards were then given as input to the computer [2].

8) 1962 – "Optacon", the first portable reading device for the blind was developed by John Linvill [7]. He later got patent for it in 1966. Optacon helps blind people to read material that isn't in braille format.

9) 1968 – ATF and Adrian Frutiger introduced OCR-A and OCR-B [10]. These are fonts to facilitate the functions of OCR operations.

*10)*Early 1970's – Recognition Equipment, Inc. developed a system to read credit card receipts from gasoline purchases [9]. The receipt would be imprinted with customer's account number at the time of transaction. It was typically embossed in OCR-A font. The equipment provided by Recognition Equipment, Inc. was a high speed system which was kept at the processing centre. The equipment would read account numbers on the receipt at speeds of 45 to 55 feet per second [9].

*11)*Late 1970's – Recognition Equipment, Inc. developed handheld OCR reader. The product replaced punch-hole price tags with price tags having OCR strings [9].

*12)*1978 – "Kurzweil Reading machine" which is based on OCR was brought in the market by Kurzweil Computer Products. The machine could to read any printed documents, let it be books, magazines etc., out loud so the blind person could read anything he liked. Later in 1980 Kurzweil Computer Products Company was sold to Xerox. Xerox renamed the company as Scansoft [8].

*13)*1984 – First passport scanner was developed by Caere Corporation for the U.S. State Department [9]. Actually, the use of OCR for passport scanning was introduced in 1983 as a part of an international convention [9].

*14)*2000 – OCR technology is made online service known as WebOCR. It is available in cloud computing environment and also in mobile applications. It now has real time applications like translation of foreign-language signs on a smart-phone [9].

*15)*2005 – HP publishes "Tesseract", a free cross-platform OCR engine.2008 – Adobe Acrobat develops OCR support for PDF files.

*16)*2015 – Google Drive provides OCR tools to scan files in over 200 languages for free [11].

III. RELATED WORK

OCR is sensitive to image orientation, resolution, noise, obstacles and geometric deformations. Therefore, image pre-processing is highly required. Pre-processing techniques used in most of the OCR systems are discussed below:

A. Pre-Processing

The input image is pre-processed to facilitate easier detection of text regions. As proposed in [3], the image is converted to the YUV color space (luminance + chrominance), and only the luminance(Y) channel is used for further processing. The conversion is done using the MATLAB function 'rgb2ycbcr' which takes the input RGB image and converts it into the corresponding YUV image. The individual channels can be extracted from this new image. The Y channel refers to brightness or intensity of the image whereas the U and the V channels refer to the actual color information [12]. Since text present in an image has more contrast with its background, by using only the Y channel, the image can be converted to a grayscale image with only the brightness / contrast information present.

Fig. YUV channels for test image (1)

B. De-Skew

If the document is not scanned properly, because of which text is not properly aligned, then it will require few clockwise or counter clockwise tilting so the text is perfectly horizontal or vertical. In early stages of pre-processing adjusting orientation of text areas is a common problem. Techniques to detect skew can be classified as: analysis of projection profile, Hough trend form, connected components, clustering and correlation between line techniques. There are twenty five various techniques as investigated by Hull and Taylor, 1998, in their survey for skew detection. Whereas, there are many new technologies invented for skew detection with varying accuracy. Therefore, choice of technique can be done on the basis of application or type of image used.

Some simple steps for auto-rotation and aligning image as per text are given in:

Convert the image into binary.

Store every first pixel occurred while scanning the horizontal lines from left to right in the array.

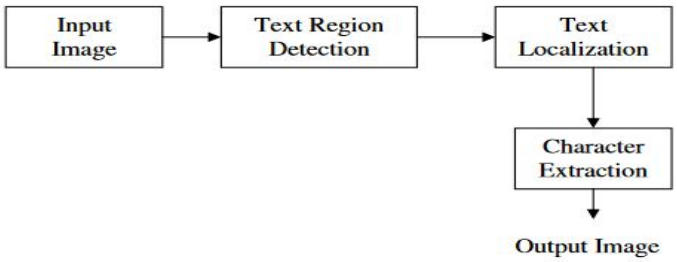


Fig. Basic Block diagram for edge based text extraction.

C. Detection

A Gaussian pyramid is created by successively filtering the input image with a Gaussian kernel of measurement 3x3 and down-sampling the image in every direction by half. Down sampling refers to the process whereby an picture is resized to a lower resolution from its original resolution. A Gaussian filter of size 3x3 will be used as shown in following Fig. Each level in the pyramid corresponds to the input image at a different resolution. A sample Gaussian pyramid with four levels of resolution is shown in following Figure. These images are next convoluted with directional filters at different orientation kernels for edge detection in the horizontal (00), vertical (900) and diagonal (450, 1350) directions.

D. Localization

The process of localization involves further enhancing the text regions by eliminating non-text regions. One of the properties of text is that usually all characters appear close to each other in the image, thus forming a cluster. By using a morphological dilation operation, these possible text pixels can be clustered together, eliminating pixels that are far from the candidate text regions. Dilation is an operation which expands or enhances the region of interest, using a structural element of the required shape and/or size. The process of dilation is carried out using a very large structuring element in order to enhance the regions which lie close to each other. In this algorithm, a structuring element of size [7x7] has been used.

The resultant image after dilation may consist of some non-text regions or noise which needs to be eliminated. An area based filtering is carried out to eliminate noise blobs present in the image. According to [1], only those regions in the final image are retained which have an area greater than or equal to 1/20 of the maximum area region.

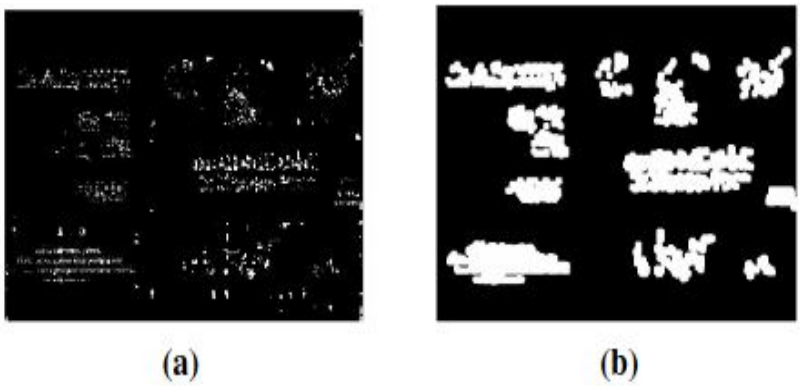


Fig. (a) Before dilation (b) After dilation

E. Character Extraction

The common OCR systems available require the input image to be such that the characters can be easily parsed and recognized. The text and background should be monochrome and background-to-text contrast should be high. Thus this process generates an output image with white text against a black background. A sample test image and its resultant output image from the edge based text detection algorithm are shown in Figures 9(a) and 9(b) below.

(a) (b)

Fig. (a) Original image (b) Result

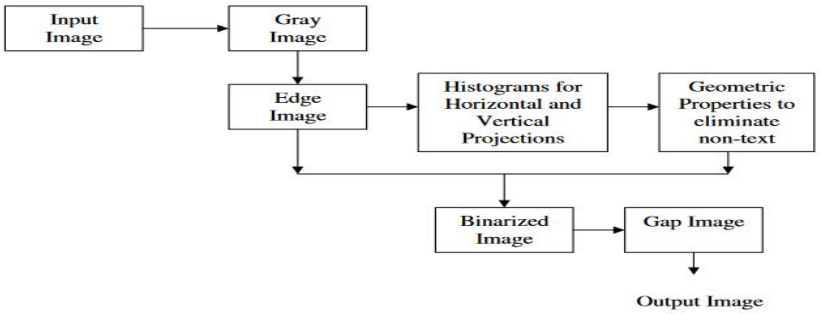


Fig. Basic Block diagram for Connected Component based text extraction.

F. Feature Extraction

This technique defines every character by the presence or absence of key options, as well as height, width, density, loops, lines, stems and alternative character traits. Feature extraction may be an excellent approach for OCR of magazines, optical maser print and prime quality pictures.

G. Segmentation

Segmentation may be a method that determines the constituents of a picture, it's necessary to find the regions of the document wherever information are written and distinguish them from figures and graphics. For example, once playing automatic mail-sorting, the address should be settled and separated from alternative print on the envelope like stamps and company logos, before recognition.

H. Matrix Matching

Matrix Matching converts every character into a pattern among a matrix, and so compares the pattern with associate degree index of glorious characters. Its recognition is strongest on monotype and uniform single column pages.

I. Neural Networks

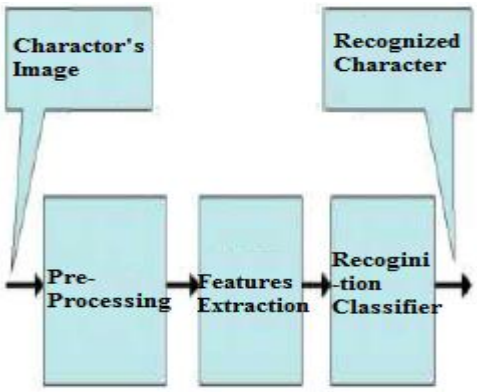
This strategy simulates the method the human neural system works; it samples the pixels in every image and matches them to a glorious index of character pixel patterns. The power to acknowledge the characters through abstraction is nice for mounted documents and broken text. Neural networks are ideal for specific sorts of issues, like process exchange information or finding trends in graphical patterns. All told these approaches Neural Networks are economical than others.

J. Post Processing

They are two types of post processing,

1) Grouping

2) Error-detection and correction.



IV. CONCLUSIONS

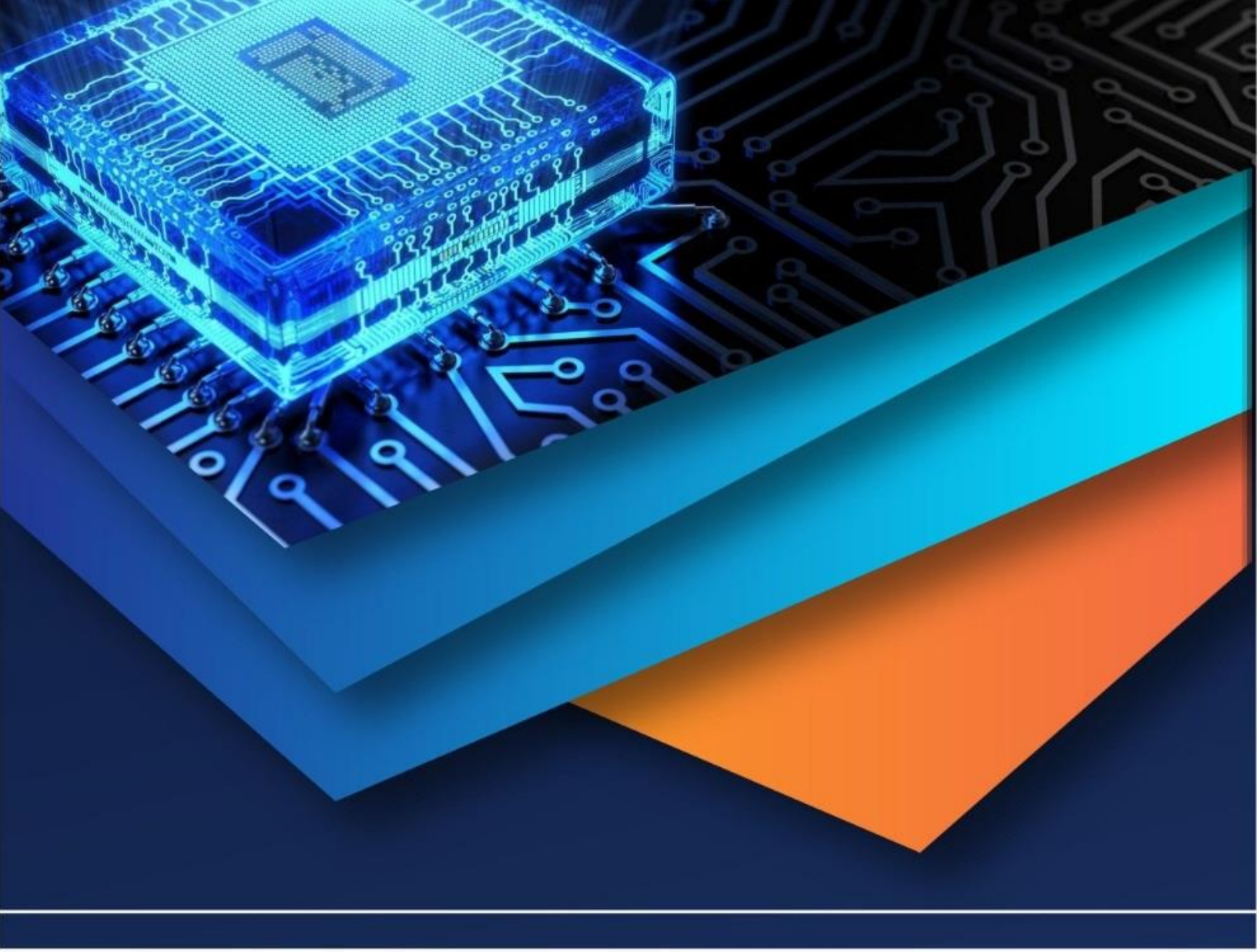
In this paper we have a tendency to survey an out sized variety of methods of optical character recognition. We analysed the benefits and downsides of varied OCR ways. We have a tendency to additionally plan a changed back propagation methodology. It's wide employed in neural network. The planned methodology computes error rate with efficiency, it ends up in increasing the accuracy of neural network. Our planned neural network based methodology is providing 100 percent accuracy in OCR.

V. ACKNOWLEDGMENT

I would prefer to give thanks the researchers likewise publishers for creating their resources available. Additionally, I appreciative to commentator for their vital recommendations moreover I'm conjointly grateful to reviewer for their valuable suggestions and also thank the college authorities for providing the required infrastructure and support.

REFERENCES

[1] Hiral Modi, M. C. Parikh, "A Review on Optical Character Recognition Techniques", International Journal of Computer Applications (0975 – 8887), Volume 160 – No 6, February 2017.
[2] H. F. Schantz, "The history of OCR: optical character recognition", Recognition Technologies Users Association.
[3] La macchina che legge e che scrive, in La scienza per tutti, Year XXIII, n-11, Milano, Casa Editrice Sozogno, 1-June 1916, p.166 (italian).
[4] Macchina per leggere pei ciechi, in La scienza per tutti, Year XXVIII, n-2, Milano, Casa Editrice Sozogno, 15-January 1921, p.20 (italian).
[5] history-computer.com/ModernComputer/Basis/OCR.html
[6] Ankit Sharma, Dipti R Chaudhary, "Character Recognition using Neural Network", International Journal of Engineering Trends and Technology (IJETT), Volume4Issue4, April 2013.
[7] www.afb.org/afbpress/pubnew.asp?DocID=aw040204
[8] www.kurzweiltech.com
[9] Microscan Systems, Inc., "Understanding Optical Character Recognition", Retrieved 18 September 2016.
[10] Frutinger,Adrian, "Type. Sign. Symbol.", ABC Verlag, Zurich, 1980. p. 50.
[11] Dmitriy Genzel, Ashok Popat, Dhyanesh Narayanan, "Paper to Digital in 200+ languages", Google Research Blog, May 06, 2015.
[12] Wojciech Bieniecki, Szymon Grabowski, Wojciech Rozenberg, "Image Pre-processing for Improving OCR Accuracy", MEMSTECH'2007, May 23-26, 2007, Lviv-Polyana, UKRAINE.
[13] Yasser Alginahi, "Preprocessing Techniques in Character Recognition",www.intechopen.com.
[14] Leila kabbai, Anissa Sghaier, Ali Douik and Mohsen Machhout, "FPGA Implementation of filtered images using 2D Gaussian filter", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 7, No. 7, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY