A Survey of OCR Applications

Amarjot Singh, Ketan Bacchuwar, and Akshay Bhasin

Abstract—**Optical Character Recognition or OCR is the electronic translation of handwritten, typewritten or printed text into machine translated images. It is widely used to recognize and search text from electronic documents or to publish the text on a website. The paper presents a survey of applications of OCR in different fields and further presents the experimentation for three important applications such as Captcha, Institutional Repository and Optical Music Character Recognition. We make use of an enhanced image segmentation algorithm based on histogram equalization using genetic algorithms for optical character recognition. The paper will act as a good literature survey for researchers starting to work in the field of optical character recognition.**

Index Terms— **Genetic algorithm, bimodal images, Captcha, institutional repositories and digital libraries, optical music recognition, optical character recognition.**

I. Introduction

Highlight in 1950's [1], applied throughout the spectrum of industries resulting into revolutionizing the document management process. Optical Character Recognition or OCR has enabled scanned documents to become more than just image files, turning into fully searchable documents with text content recognized by computers. Optical Character Recognition extracts the relevant information and automatically enters it into electronic database instead of the conventional way of manually retyping the text.

Optical Character Recognition is a vast field with a number of varied applications such as invoice imaging[2], legal industry [2], banking, health care industry [2] etc. OCR is also widely used in many other fields like Captcha [3], Institutional repositories and digital libraries[4], Optical Music Recognition [5] without any human correction or human effort, Automatic number plate recognition [6] and Handwritten Recognition [7] .

The paper presents a survey of Optical Character Recognition applications and further focuses on three important applications of Optical Character Recognition, Captcha, institutional repositories and digital libraries and Optical Music Recognition.

Captcha is a challenge-response test most often placed within web forms to determine whether the user is human. The purpose of Captcha is to block automated scripts that post spam content everywhere they can. Institutional repositories and digital libraries helps to open up the outputs of an

institution and give it visibility and more impact on worldwide level whereas optical music recognition has many applications like processing of different classes of music, large scale digitization of musical data and also it can be used for diversity in musical notation.

The primary step involved in the implementation of optical character recognition is object segmentation. According to the literature survey, a number of methods have been used for segmentation. In the bimodal images used for the applications, the histogram shows a sharp and deep valley between two peaks representing the objects and background respectively. The histogram can be used to select the threshold representing the bottom of this valley. Khankasikam et al [8], [9] proposed the valley sharpening techniques which restricts the histogram to the pixels with large absolute values of derivatives where as S. Wantanable et. Al [10], [11] proposed the histogram difference method, which selects threshold at the gray level with the maximum amount of difference. This method uses the information concerning neighboring pixels or edges in the original picture to modify the histogram so as to make it useful for thresholding. Another method includes directly dealing with the grey level histogram by parametric techniques. By a sum of Gaussian distributions histogram is approximated in the least sense square and a statistical decision procedures are applied [10]. These methods are tedious and they involve high computational power. In Di gesu [12], [10] the idea of using both intensities and spatial information has been considered to take into account local information used in human perception. A number of new methodologies and strategies have been proposed over the past few years to find global as well as local solutions in nonlinear multimodal function optimization [12], [13], [14]. In addition attempts have also been made to use Fuzzy Logic [15], Artificial Neural Network [16] for optical character recognition. With the help of crowding multiple peaks can be maintained in multimodal optimization problem. Crowding method is extremely reliable in detecting the peaks on bimodal histogram. Further, GA can be applied to discover the valley bottom between these peaks which can be used as threshold for extracting information from the background. Genetic algorithms offer a particularly interesting approach as they are quite effective for rapid global search of large, non-liner and poorly understood spaces. Moreover genetic algorithms are every effective in solving large scale problems [17].

This paper makes use of an OCR system which makes use of histogram equalization to extract images. The histogram used by the mentioned algorithm is bimodal in nature hence it can be divided into two classes [18]. Genetic algorithm is further used to select the threshold from the histogram for

Manuscript received April 7, 2012; revised May 14, 2012.

A. Singh and K. Bacchuwar are with Electrical Department of National Institute of Technology, Warangal, India (email: amarjotsingh@ieee.org).

A. Bhasin is with Electronics and Communication Department of Guru Nanak Dev University, Jalandhar, Punjab, India.

extracting the object from the background. The abilities of the system are tested on images with noise, blur and change in illumination.

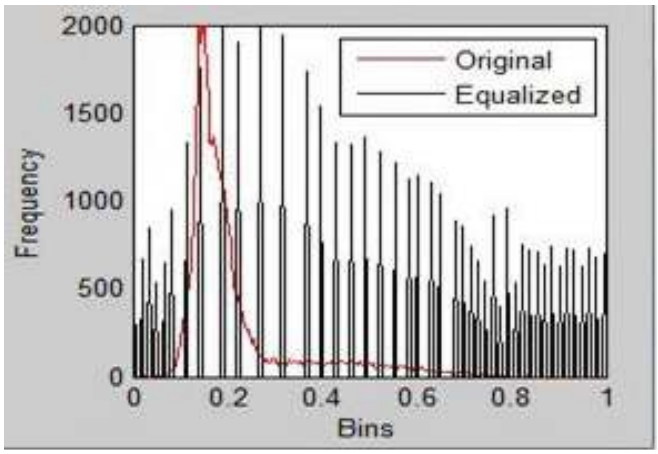

Fig. 1. Normal and Equalized histogram of the resultant enhanced image

The paper is divided into five sections. Second section explains the application of OCR in various fields while the Third section explains the proposed and standard algorithms/methodology used in paper to enhance and extract the characters followed by applications. The fourth section explains the results while the last section gives a brief summary of the paper.

II. Applications

Optical character recognition has been applied to a number of applications. Some of them have been explained below.

A. Invoice Imaging

Invoice imaging [2] is widely used in many businesses applications to keep track of financial records and prevent a backlog of payments from piling up. In government agencies and independent organizations, OCR simplifies data collection and analysis, among other processes. As the technology continues to develop, more and more applications are found for OCR technology, including increased use of handwriting recognition. Furthermore, other technologies related to OCR, such as barcode recognition, are used daily in retail and other industries.

B. Legal Industry

Legal industry [2] is also one of the beneficiaries of the OCR technology. OCR is used to digitize documents, and directly entered to computer database. Legal professionals can further search documents required from huge databases by simply typing a few keywords.

C. Banking

Another important application of OCR is in banking [2], where it is used to process cheques without human involvement. A cheque can be inserted into a machine where the system scans the amount to be issued and the correct amount of money is transferred. This technology has nearly been perfected for printed checks, and is fairly accurate for handwritten checks as well reducing the waiting time in banks.

D. Healthcare

Healthcare [2] has also seen an increase in the use of OCR technology to process paperwork. Healthcare professionals always have to deal with large volumes of forms for each patient, including insurance forms as well as general health forms. To keep up with all of this information, it is useful to input relevant data into an electronic database that can be accessed as necessary. Form processing tools, powered by OCR, are able to extract information from forms and put it into databases, so that every patient's data is promptly recorded.

E. Captcha

A CAPTCHA [3] is a program that can generate and grade tests that human can pass but current computers programmers' cannot. Hacking is a serious threat to internet usage. Now a day's most of the human activities like economic transactions, admission for education, registrations, travel bookings etc are carried out through internet and all this requires a password which is misused by hackers. They create programs to like dictionary attacks and automatic false enrolments which lead to waste of memory and resources of website. Dictionary attack is attack against password authenticated systems where a hacker writes a program to repeatedly try different passwords like from a dictionary of most common passwords. In CAPTCHA, an image consisting of series of letters of number is generated which is obscured by image distortion techniques, size and font variation, distracting backgrounds, random segments, highlights and noise in the image. This system can be used to remove this noise and segment the image to make the image tractable for the OCR (Optical Character Recognition) systems.

F. Institutional Repositories and Digital Libraries

Institutional repositories [4] are digital collections of the outputs created within a university or research institution. It is an online locale of intellectual data of an institution, especially a research institution where it is collected, preserved and aired. It helps to open up the outputs of an institution and give it visibility and more impact on worldwide level. Enables and encourages interdisciplinary approaches to research and facilitates the development and sharing of digital teaching materials and aids. It is basically a collection of peer reviewed journal articles, conference proceedings, research data, monographs, books, theses and dissertations and presentations. Their first role is to provide the Open Access literature. Practical implementation of this includes setting up a system which consists of scanner which scans the documents. This scanned document is then fed as an input to an Optical Character Recognition system where information is acquired and retained in digitized form.

G. Optical Music Recognition

Automated learning system extract information from images and is part of major researches. Optical music recognition (OMR) [5] born in 1950's is a developed field and initially was aimed towards recognizing printed sheets which can be edited into playable form with the help of electronic and electrochemical methods. An OMR system has many applications like processing of different classes of music,

large scale digitization of musical data and also it can be used for diversity in musical notation. Image enhancement and segmentation is the basic step and hence the paper focuses on it.

H. Automatic Number Recognition

Automatic number plate recognition [6] is used as a mass surveillance technique making use of optical character recognition on images to identify vehicle registration plates. ANPR has also been made to store the images captured by the cameras including the numbers captured from license plate. ANPR technology own to plate variation from place to place as it is a region specific technology. They are used by various police forces and as a method of electronic toll collection on pay-per-use roads and cataloging the movements of traffic or individuals.

I. Handwriting Recognition

Handwriting recognition [7] is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices. The image of the written text may be sensed "off line" from a piece of paper by optical scanning (optical character recognition) or intelligent word recognition. Alternatively, the movements of the pen tip may be sensed "on line", for example by a pen-based computer screen surface.

III. Algorithm

The paper makes use of an improved algorithm described below.

A. Standard Algorithm

The algorithm is a histogram based approach used to extract information from bimodal images. Genetic algorithm is applied on the histogram of the bimodal image to extract useful information from the background. The histogram is a plot or graph of the frequency of occurrence of each gray level in image across gray scale values. Let a random population of N size is initialized and the element acquire value between 0 to 255. The operations are carried out over randomly chosen two parents. The appropriate values of crossover probability and mutation probability are fixed. The winner of each tournament (the one with the best fitness) is selected. After computing fitness values of the off-spring, tournament selection strategy is used to allow off-springs to compete with the parents. A competition among the randomly chosen individuals is carried out and the fittest between the both is selected. Genetic algorithms use this methodology for selection from the population. Two parents and two off-springs compete to give two best individuals as a result. This method is termed as crowding network. The resulting elements are located in their respective classes. The method basically replaces the older elements in the population by the fittest elements in the resulting generation which helps to reduce replacement error. The repetition performed on all the elements results into convergence. This converged value is the gray value corresponding to the minima between two peaks and this gray value is used as threshold value and then image is segmented.

B. Proposed Algorithm

Proposed method is divided in two steps [18]. Firstly, the image histogram is equalized which results into the redistribution of the intensities as shown in the Fig. 1. For reference input normalized histogram is shown in Fig. 1, it shows the redistribution of pixel intensities. It results into an increase in the intensities of the pixels which are low in grayscale while performs a decrement for high intensity valued pixels. The intensities with lower intensity value are upgraded to higher values and vice versa. The changes in the intensity values results into the enhancement of the bimodal images. With the enhancement in the histogram, the intensity values increase as well as decrease for every pixel due to redistribution of the histogram. This leads to an enhancement in the gray values of the class which is having a gray scale value near to the upper hand (129 to 225). The pixels belonging to upper band move towards upper band. In order to extract the threshold genetic algorithm is applied to equalized histogram. Due to increase in the pixel intensities the selected threshold will be higher as compared to the previous methodology and higher threshold will be required to extract lower class pixels from higher class pixels. The pixel values corresponding to upper class can be easily separated from the pixels of the pixels of lower class pixels from higher class pixels. On applying the threshold the pixels values corresponding to the upper class can be easily separated from the pixels of the lower band as the histogram equalization already moved the pixels which were lower in value but belonged to the upper band towards the higher value. This method is highly efficient and can be effectively used to extract the information from the image with different conditions. The technique used doesn't oblige any valley sharpening techniques.

IV. Results

The results obtained from the simulation enable us to explain the capabilities of the methodology applied to bimodal images. The robustness and efficiency of the method is also further tested on two different images on an Intel Core 2 Duo 2.20 GHz machine. The result obtained after processing of the image enable us to scrutinize the method applied. The paper uses the images whose histogram displays a bimodal character. The coding for the algorithm is done in MATLAB (Version 7.10.0.499) on an Intel Core 2 Duo 2.20 GHz machine. The paper uses bimodal images of a music script which is used to extract useful musical signs and encryptions using the algorithms mentioned above.

The music encryption, Institutional repository and Captcha were given as inputs to the system and the algorithms were one by one applied to obtain the signatures for all the three applications. In order to segment the images, the two peaks in the histogram are obtained using crowding. The parameters of the algorithm are selected as (i) number of population elements " N " is 20 (ii) crossover probability $c\ P$ =0.9 (iii) mutation probability $m\ P$ decreases trailing an exponential rate with starting value 0.05. The threshold value is obtained using genetic algorithm. According to the proposed algorithm, the histogram of the input image is equalized histogram as shown in the Fig. 1. The equalization makes the image much brighter as compared to the original image hence the pixels which are suppose to be in the upper band or

segmented move to a higher intensity. This method can also be effectively used to deal with illumination changes and motion blur [18]. Image segmentation is further carried out on the enhanced images using the GA based approach as described in the algorithm. The parameters used for the standard GA algorithms are similarly used here. The results obtained for the proposed algorithm are show in Fig. 2 for Optical music recognition, in Fig. 3 for Institutional repository and in Fig. 4 Captcha.

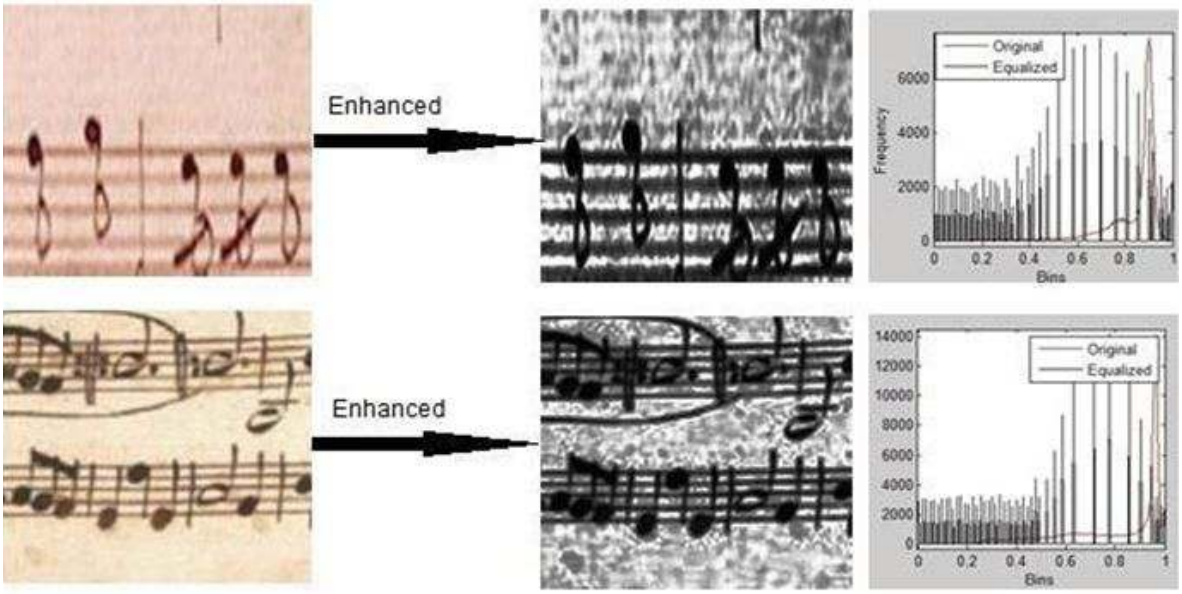

Fig. 2. Segmentation results for optical Music recognition

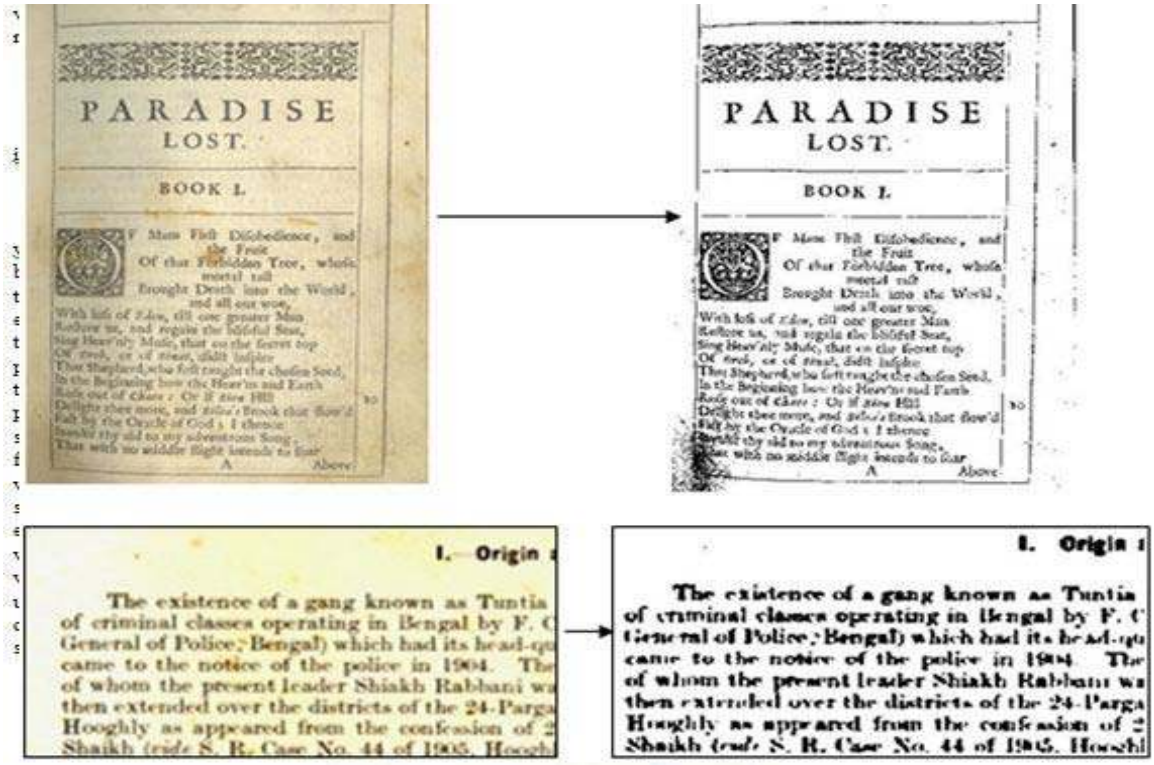

Fig. 3. Segmentation results for Institutional Repository

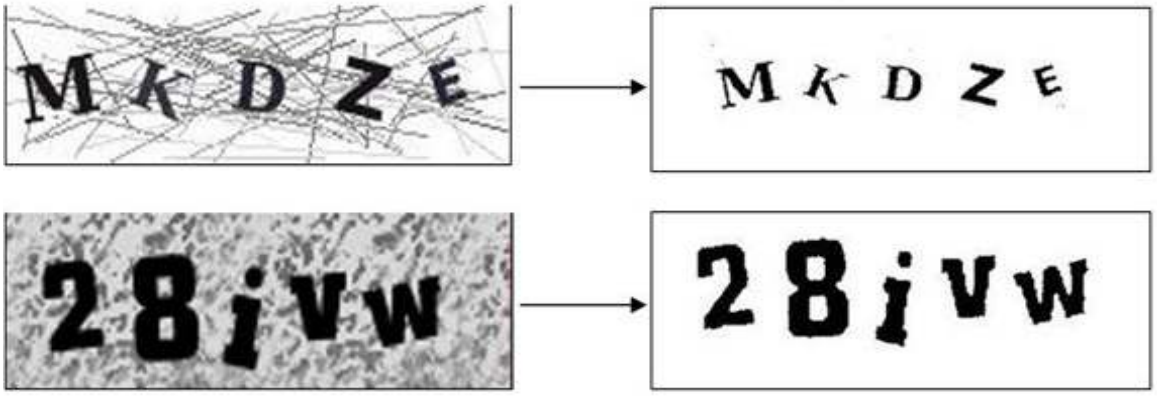

Fig. 4. Segmentation results for Captcha

V. Conclusion

The paper presents a brief survey of the applications in various fields along with experimentation into few selected fields. The proposed method is extremely efficient to extract all kinds of bimodal images including blur and illumination. The paper will act as a good literature survey for researchers starting to work in the field of optical character recognition.

References

[1] K. Bachuwar, A. Singh, G. Bansa, S. Tiwari, "An Experimental Evaluation of Preprocessing Parameters for GA Based OCR Segmentation" in *3rd International Conference on Computational Intelligence and Industrial Applications (PACIIA 2010),* 2010, proceedings, Vol. 2, pp. 417 -420.

[2] M.D. Ganis, C.L. Wilson, J.L. Blue, "Neural network-based systems for handprint OCR applications" *in IEEE Transactions on Image Processing*, 1998, Vol: 7 , Issue: 8, p.p. 1097 - 1112

[3] R. Gossweiler, M. Kamvar, S. Baluja, "What's Up CAPTCHA? A CAPTCHA Based On Image Orientation", in *WWW*, 2009.

[4] B. Joanna," Building an institutional repository at Loughborough University: some experiences, program: *Electronic library and information systems,* 2009.

[5] A. Singh, K. Bacchuwar, A. Choubey, S. Karanam, D. Kumar, "An OMR Based Automatic Music Player", in *3rd International Conference on Computer Research and Development (ICCRD 2011) in, (IEEE Xplore)*, 2011, Vol. 1, pp. 174-178.

[6] S.L. Chang, T. Taiwan , L.S. Chen, Y.C. Chung, S.W. Chen, " Automatic license plate recognition" in *IEEE Transactions on Intelligent Transportation Systems,* 2004, Vol: 5 , Issue: 1, p.p. 42 - 53

[7] R Plamondon, S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey" *IEEE transaction on pattern Analysis and machine Intelligence*, 2000, 22(1), 63-84

[8] K. S. Fu, J.K. Mui, "A Survey on Image segmentation". *Pattern Recognition*, 1981, 13, 3–16

[9] M.S. Kalas, "An Artificial Neural Network for Detection of Biological Early Brain Cancer", *International Journal of Computer Applications,* 2010, 1(6), 17–23.

[10] R. Storn, K. Price, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", *Journal of Global Optimization*, 1997, 341–359

[11] IP, H.H.-S., S.-L. Chan, "Hypertext-Assisted Video Indexing and Content- based Retrieval", *ACM* 0-89791-866-5, 1997, PP. 232–23.3

[12] N.R. Pal, S.K. Pal, "A Review on Image Segmentation Techniques", *Pattern Recognition*, 1993, Vol 26(9), 1277–1294

[13] Weszka, J.S., Nagel, R.N., Rosenfeld, A. 1974, "A Threshold selection technique*", IEEE Trans. Computer* C-23, 1322–1326

[14] Kao, Y.-T., E. Zahara, "A hybrid genetic algorithm and particle swarm optimization for multimodal functions", *Applied Soft Computing,* 2008, Vol. 8(2), 849–857

[15] J.M. Fonseca, "Optical character recognition using automatically generated Fuzzy classifiers" in *Fuzzy Systems and Knowledge Discovery (FSKD)*, 2011, Vol: 1, pp: 448 – 452.

[16] Mani, N., "An optical character recognition using artificial neural network", in *IEEE International Conference on Systems, Man, and Cybernetics*, 1996, Vol.3, pp: 2244 – 2247.

[17] Yang, J., Honavar, V., "Feature subset selection using genetic algorithms" *Proc. IEEE Intelligent Systems*, 1998, Vol. 13, no. 2, pp. 44-49.

[18] A. Singh, K. Bacchuwar, A. Choubey, S. Karanam, "A Novel GA Based OCR Enhancement and Segmentation Methodology for Marathi Script in Bimodal Framework"in *Springer Verlag,* 2011.