IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105100 |

Handwritten Annunacer Using Matlab

Mr. Deepak V. Kamble, Sanjana .A. Jagtap, Mansi .N.Kamble, Rohan .M.Kole, Sahil .B . Patil, Avdhut.V.Sargar

Guide, Department of Electronic and Telecommunication, Sanjay Ghodawat Polytechnic, Atigre, India

Department of Electronic and Telecommunication, Sanjay Ghodawat Polytechnic, Atigre, India

ABSTRACT—If handwriting is scanned and then understood by the computer, it is called offline handwriting recognition. In case, the handwriting is recognized while writing through touch pad using stylus pen, it is called online handwriting recognition. From the classifier perspective, character recognition systems are classified into two main categories i.e. segmentation free (global) and segmentation based (analytic).

I. INTRODUCTION

- **Optical Character Recognition**

Optical Character Recognition deals with the problem of recognizing optically processed Characters. Optical recognition is performed off-line after the writing or printing has been completed, as opposed to on-line recognition where the computer recognizes the characters as they are drawn. Both hand printed and printed characters may be recognized, but the performance is directly dependent upon the quality of the input documents. Character recognition is a fundamental, but most challenging in the field of pattern recognition with large number of useful applications. It has been an intense field of research since the early days of computer science due to it being a natural way of interactions between computers and humans. More precisely Character recognition is the process of detecting and recognizing characters from the input image and converts it into ASCII or other equivalent machine editable form . The technique by which a computer system can recognize characters and other symbols written by hand in natural handwriting is called handwriting recognition system. Handwriting recognition is classified into offline handwriting recognition and online handwriting recognition . If handwriting is scanned and then understood by the computer, it is called offline handwriting recognition. In case, the handwriting is recognized while writing through touch pad using stylus pen, it is called online handwriting recognition. From the classifier perspective, character recognition systems are classified into two main categories i.e. segmentation free (global) and segmentation based (analytic).

II. LITERATURE SURVEY

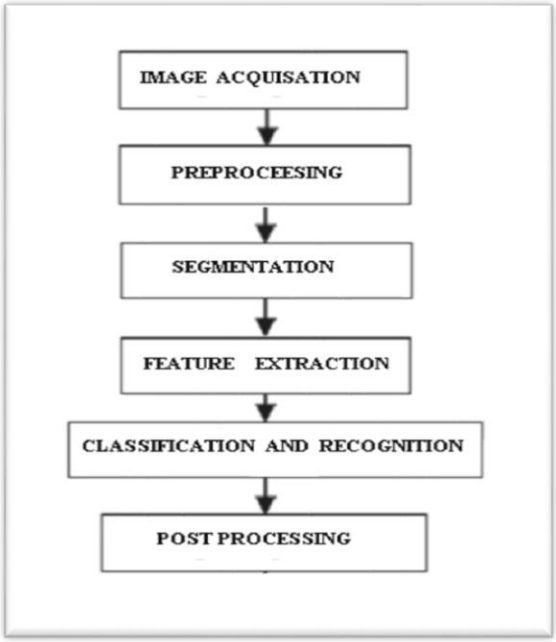
K. Gaurav, Bhatia P. K. Et al, this paper deals with the various pre-processing techniques involved in the character recognition with different kind of images ranges from a simple handwritten form based documents and documents containing colored and complex background and varied intensities. In this, different preprocessing techniques like skew detection and correction, image enhancement techniques of contrast stretching, binarization, noise removal techniques, normalization and segmentation, morphological processing techniques are discussed. It was concluded that using a single technique for preprocessing, we can't completely process the image. However, even after applying all the said techniques might not possible to achieve the full accuracy in a preprocessing system. Salvador España-Boquera et al in this paper hybrid Hidden Markov Model (HMM) model is proposed for recognizing unconstrained offline handwritten texts

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105100 |

III. BLOCK DIAGRAM

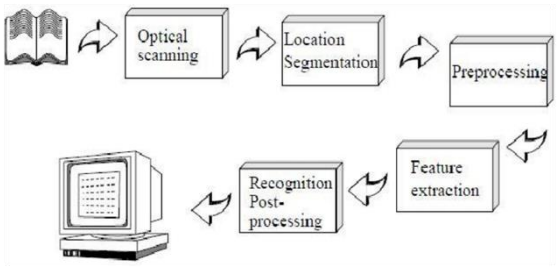


The above block diagram shows the dataflow pattern according to which a character is recognised optically using the optical character recognition in MATLAB. The sensor used is basically a CMOS sensor which can be found in a VGA camera followed by a the image acquisition into the MATLAB system where the image is acquired into the system by the method of using the various relay systems .This is then followed by the various pre-processing techniques such as edge detection thresholding conversion to grayscale from rgb etc. The next step is feature extraction which is used to extract the features of the images using the characteristic loci method and then finally the image is recognised by the system via classification and comparing with the database. and can be given as output to the user.

- COMPONENTS OF AN OCR SYSTEM

A typical OCR system consists of several components. In figure 3 a common setup isillustrated. The first step in the process is to digitize the analog document using an optical scanner. When the regions containing text are located, each symbol isextracted through a segmentation process.

components



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105100 |

IV. METHODOLOGY

PREPROCESSING DATA

- LOCATION AND SEGMENTATION

Segmentation is a process that determines the constituents of an image. It is necessary to locate the regions of the document where data have been printed and distinguish them from figures and graphics. For instance, when performing automatic mailsorting, the address must be located and separated from other print on the envelope like stamps and company logos, prior to recognition. Applied to text, segmentation is the isolation of characters or words. The majority of optical character recognition algorithms segment the words into isolated characters which are recognized individually. Usually this segmentation is performed by isolating each connected component that is each connected black area. This technique is easy to implement, but problems occur if characters touch or if characters are fragmented and consist of several parts. The main problems in segmentation may be divided into four groups:

V. ADVANTAGES

I. It is easy to use due to its command line interface and file oriented structure.

II. MATLAB is platform independent and hence it can be installed on different Operating Systems such as Windows, Vista, Linux and Macintosh

III. .MATLAB has huge built-in library of functions for many predefined tasks. This makes job more comfortable and saves time. These functions are available as part of various toolkits which include signal processing, image processing,

VI. FUTURE DEVELOPMENT

- MATLAB is inexpensive software.
- Implement and test your algorithms easily.
- Develop the computational codes easily.
- Debug easily.
- Use a large database of built in algorithms.
- Process still images and create simulation videos easily.
- Symbolic computation can be easily done.
- Call external libraries.

Perks of Using ICR Software

VII. SOFTWARE USED

MATLAB
SYNTAX
VARIABLES
VECTORS/MATRICES
SEMICOLONS
STRUCTURES

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105100 |

VIII. CONCLUSION

, By using the methods of characterization loci and SVM toolbox we were able to achieve an accuracy in the range of 60-70 % on untrained data for online character recognition which dramatically increased over 85% when the same data was trained into the database. The coding used for noise reduction was also successful and capable of removing stray marks on the sheet of paper being used as well as the other noises that came while taking the picture were also removed automatically. Today's uses of OCR are still somewhat limited to the scanning of the written word into useable computer text. The uses include word processing, mail delivery system scanning, ticket reading, and other such tasks. In the future, the uses of OCR have been speculated to be far more advanced. Some of these advancements are already in the workings. Most believe that the use for reading the written text will diminish as electronic data interchange (EDI) is used more andmore, while paper documents are phased out.

REFERENCES

1. WWW.GOOGLE.COM
2. WWW.MATHWORKS.COM
3. MATLAB HELP
4. MATHWORKS FILE SHARE
5. WIKIPEDIA
6. IEEE Paper on "Neural Network based Handwritten Character Recognition system without feature extraction" by J.Pradeep, E.Srinivasan, S.Himavathi
7. IEEE Paper on "Recognition of Persian handwritten digits using Characterization
8. Loci and Mixture of Experts" by Mohammad R. Moradian ,Alireza Esmkhani ,Farzad M. Jafarlou
9. IEEE Paper on "Character Recognition Using Neural Networks" by Rokus Arnold, Poth Miklos

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details