Demonstrated Study of Pattern and optical Character Recognition

Anamika Jain[1], Vikram Khandewal[2]
[1]Assistant professor in department of computer science , [2]Assistant professor in department of computer science
[2]anamika.jain@poornima.org, [1]vikram.khandelwal@poornima.org

Abstract

Optical Character Recognition, prominently alluded to as the OCR, is a dynamic territory of research in this data age. For the pattern recognition, a study related to old information for generating new information is used in documents system. To solve this many handwritten papers are transformed into computer vision systems. OCR has proposed in which computer has enabled to understand the handwritten image script and convert into recognized characters range. It will help to readers , for better understanding of character recognition techniques.

1. Introduction

OCR is the most popular application of pattern recognition where characters are identified. The pattern recognition has three on going steps: observation, segmentation and classification. It is the one of the old techniques to convert out the machine printed character from single character drawn. There are two types for text character identification: one is online where the text characters are converted into images and other one is off line hand writing experts.

OCR is the composition of signal processing, pattern recognition and artificial intelligence with machine vision. It is referred as offline vision character pattern recognition.
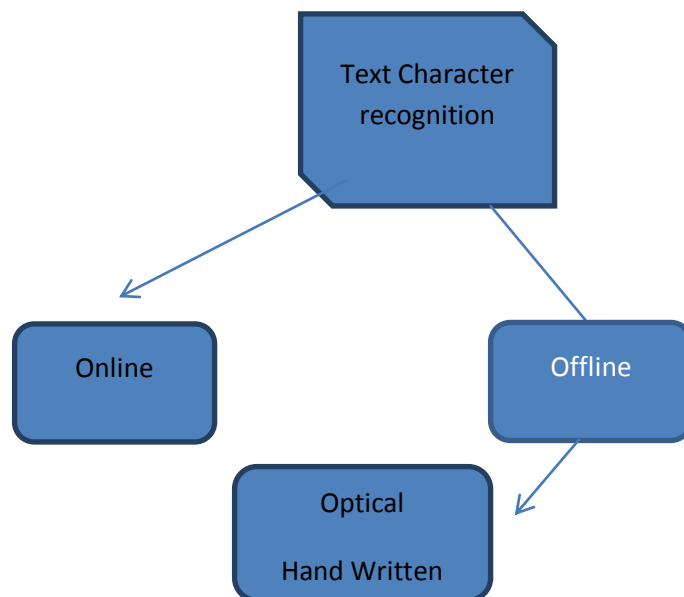

Figure 1: General character recognition

OCR problem was identified as the computer is old. In 1954 , a prototype machine is used and 1967 many companies finally used it.
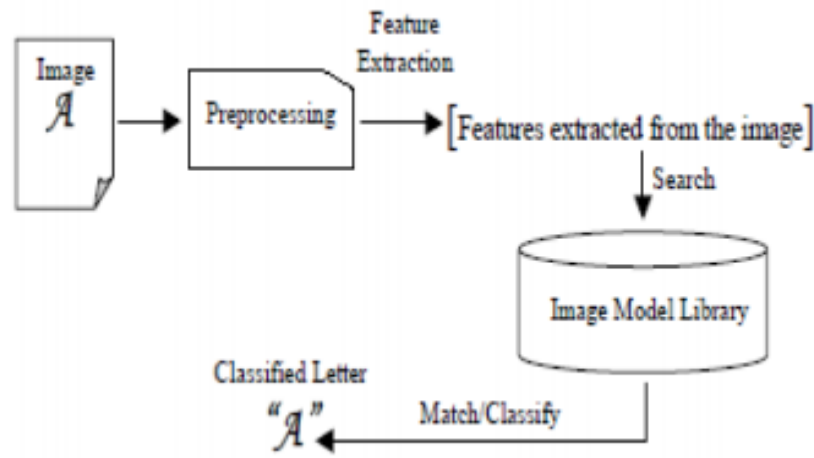
Design of Pattern Recognition System:

Figure 2 Basic View of Character Recognition

1. Sensing:

Sensing is the first input process which is taken by the some digital camera and many devices. The studies of the sensors are also it.

1. Grouping and Segmentation:

Think about a PRS for the mechanized isolation of containers. The information might be through a camcorder that is centered around a transport line that is conveying the bottles. The jugs may have distinctive hues, shapes, sizes, surface and so forth. The jugs may likewise be astonishingly over the transport line. The primary assignment in acknowledgment is to distinguish the jug, the 'limit' that characterizes the container. This undertaking is called division or segmentation.

2. Extraction of features:

There are two major areas related to feature, feature classification and feature extraction. Feature is responsible for time and space of the pattern . So basically it is responsible for complexity of the pattern. An examination must be made as to which of these highlights seem to be "recognizing", ie., which feature(s) separate one kind of container from alternate sorts.

3. Stratification or Classification:

The characterization utilizes the highlights extricated in the past advance to choose the gathering to which the element has a place with. This is normally as some capacity of the feature(s) A few times, numerous classifiers could be utilized and the just arrangement can be

utilized to settle on an official conclusion. Some prominent classifiers are the 1-Nearest Neighbor (or the 1-NN) classifier, the n-Nearest Neighbor (or the n-NN) classifier, the Neural Net Classifier. This stage additionally may utilize the Artificial Intelligence strategies for learning. More on this in later parts. Note: The neural net classifier has "learning" consolidated into itself.

4. Post Production Process:

The contribution to this stage is the classifier code that was presented in the past segment. The part of this stage is to decide. Furthermore, once there is choice making, there will undoubtedly be mistake. In this way, the point of this stage would now be able to be characterized to be the way toward settling on choices with least normal mistake.

2. OCR Techniques:

1. Scanning & Digitization:

The Document should first be checked and spared in a standard organization in order to be process able by any standard picture preparing device.

2. Preprocessing:

The filtered picture is subjected to certain picture preparing calculations, for example, editing and arrangement. The inalienable flaws of the picture are to be evacuated before the following stage with the goal that the preparing is straightforward later. The different procedures required here are : commotion separating, De skewing, page division and textual style acknowledgment.

3. Segmentation:

Division with regards to character acknowledgment can be characterized as the way toward separating from the preprocessed picture the littlest conceivable character units reasonable for acknowledgment; i.e. to fragment the basic units of the archive.

4. Feature extraction:

The way toward portraying the sub pictures generated as the yield of the division strategy in light of certain particular parameters. There are numerous past methodologies like: layout coordinating, zoning.

5. Classification:

To setting every one of the acknowledgment units into one of a pre-decided arrangement of classes, in view of the qualities of that unit as assessed by the component extraction method. In this manner, each character picture is mapped to a printed portrayal

6. Post Processing:

Here, the yield of the characterization organize is changed over into ASCII or ISCII or other standard coding plans so words, sentences and sections are recon-structed from the yields of the grouping stage. A very much organized word reference may likewise be utilized to determine ambiguities in acknowledgment.

3. Conclusion:

This is detailed discussion about handwritten character recognize and include various concepts involved, and boost further advances in the area. The accurate recognition is directly depending on the nature of the material to be read and by its quality. Current research is not directly concern to cursive handwriting and to recognize the child handwriting which require high supervised system . From various studies we have seen that selection of relevant feature extraction and classification technique plays an important role in performance of character recognition rate. This review establishes a complete system that converts scanned images of handwritten characters to text documents.

References:

[1.] MAJIDA ALI ABED HAMID ALI ABED ALASADI Simplifying Handwritten Characters Recognition Using a Particle Swarm Optimization Approach EUROPEAN ACADEMIC RESEARCH, VOL. I, ISSUE 5/ AUGUST 2013 ISSN 2286-4822

[2.] [2] Mohammed Z. Khedher, Gheith A. Abandah, and Ahmed M. Al-Khawaldeh Optimizing Feature Selection for Recognizing Handwritten Arabic Characters PROCEEDINGS OF WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY VOLUME 4 FEBRUARY 2005 ISSN 1307-6884

[3.] [3] Ivan Dervisevic Machine Learning Methods for Optical Character Recognition December 18, 2006 Bhatia, International Journal of Advanced Research in Computer Science and Software Engineering 4(5), May - 2014, pp. 1219-1223 © 2014, IJARCSSE All Rights Reserved Page | 1223

[4.] [4] Diego J. Romero, Leticia M. Seijas, Ana M. Ruedin Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks JCS&T Vol. 7 No. 1 April 2007

[5.] [5] Chirag I Patel, Ripal Patel, Palak Patel Handwritten Character Recognition using Neural Network International Journal of Scientific & Engineering Research Volume 2, Issue 5, May-2011