*Asian Journal of Convergence in Technology* *Volume1, Issue 5*
*Issn No.:2350-1146, I.F-2.71*

# Classification and Comparative Analysis of Character Recognition Techniques

Nikhil Pai
Department Of Electronics Engineering,
Bharatratna Indira Gandhi College of Engineering,
Solpaur, Mahrashtra,india

Vijaykumar S. Kolkure
Department Of Electronics Engineering,
Bharatratna Indira Gandhi College of Engineering,
Solpaur, Mahrashtra,india

***Abstract*: The popularity of Character Recognition (CR) Technique is increasing day by day. This technique is used for data entry, data safety in Banking and all other several fields, where data safety is a prime issue. So, Character Recognition progressed to a level sufficient to produce technology driven applications. The basic aim of Character Recognition is to scan the documents, processes over it and convert them into editable format.There are different methods of Character Recognition has been implemented so far. This paper presents the classification of Character Recognition Techniques. This study also compares the different methods of Character Recognition.**

***Keywords- CR***

## I. INTRODUCTION

Character recognition (CR) has been extensively studied in the last half century and progressed to a level sufficient to produce technology driven applications. Now, the rapidly growing computational power enables the implementation of the present CR methodologies and creates an increasing demand on many emerging application domains, which require more advanced methodologies. Its aim is to classify scanned, recorded or pictured images of machine printed or handwritten text, numerals, letters, symbols and generate a description in the desired format. Optical Character Recognition (OCR) is one of the most successful applications in automatic pattern recognition. OCR has a very active field for research and development. Since mid 1950's techniques developed were able to recognize high quality printed text documents or neatly written text. CR has been classified based upon the two important aspects: According to the manner in which data has been acquired (On-line and Off-line) and According to the text type (machine printed and handwritten).

*1. Classification based on Mode of Data Acquisition:*

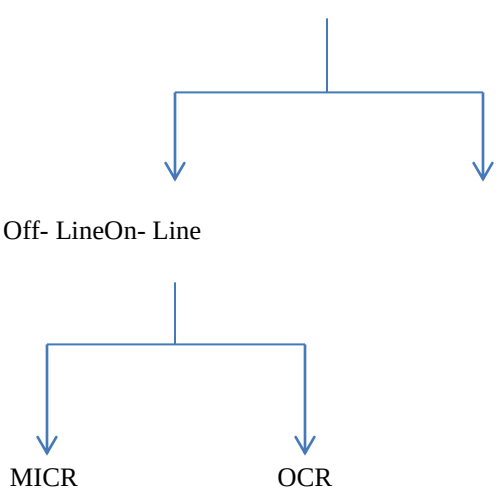


Figure 1.1: Classification based on mode of data acquisition

Basically, CR technology classify in different way. Figure 1.1 shows the classification based on mode of data acquisition. Mainly it gets distinguished in Off-Line method and On- Line method. The Off-Line method further gets divided into Magnetic Ink Character Recognition (MICR)and Optical Character Recognition (OCR).

*1.1Off- line Character Recognition:*

Offline character recognition is known as Optical Character Recognition, because typewritten or handwritten character is scanned by optical scanner or camera and converted into form of a binary or gray scale image. Offline character recognition is a more challenging and difficult task as there is no control over the medium and instrument used. The artifacts of the complex interaction between the instrument medium and subsequent operations such as scanning and Binarization present additional challenges to the algorithm for the offline character recognition. Therefore offline character recognition is considered as a more challenging task then its online counterpart.

Offline character recognition is further classified in two types according to the input provided to the system for recognition of characters.

- Magnetic ink Character Recognition (MICR)
- Optical Character Recognition (OCR).

*1.1.1MICR:*

MICR is a unique technology that relies on recognizing text which has been printed in special fonts with magnetic ink usually containing iron oxide. The system has been in efficient use for a long time in banks around the world to process checks as results give high accuracy rates with relatively low chances of error.

*1.1.2 OCR:*

Optical Character Recognition defines as a process by which we convert printed document or scanned page to ASCII character that a computer can recognize. The document image itself can be either machine printed or handwritten, or the combination of two. Computer system equipped with such an OCR system can improve the speed of input operation and decrease some possible human errors. Recognition of printed characters is itself a challenging problem since there is a variation of the same character due to change of fonts or introduction of different types of noises. Difference in font and sizes makes recognition task difficult if preprocessing, feature extraction and recognition are not robust. There may be noise pixels that are introduced due to scanning of the image. Besides, same font and size may also have bold face character as well as normal one. Thus, width of the stroke is also a factor that affects recognition. Therefore, a good character recognition approach must eliminate the noise after reading binary image data, smooth the image for better recognition, extract features efficiently, and classify patterns.

*1.2. On-Line Character Recognition:*

As the name infers online character recognition is real time recognition of characters. Online systems obtain the position of the pen as a function of time directly from the transducer that captures the writing as it is written. The most common of these devices is the electronic tablet or digitizer. Online recognition systems utilize the digitizers which directly captures writing with the order of the strokes, speed, pen-up and pen-down information. These systems recognize text while the user is writing with an on-line writing device, capturing the temporal or dynamic information of the writing. This information includes the number, duration, and order of each stroke (a stroke is the writing from pen down to pen up). The recognition system functions on the basis of the x and y coordinates generated in a temporal sequence by the pen tip movements as they create recognizable patterns on a special digitizer as the text is written. On-line systems are limited to recognizing handwritten text. Some systems recognize isolated characters, while others recognize cursive words.

*2. Classification based on Mode of Writing:*

CR technology can also classify in another manner. Figure 2.1 shows the classification based on mode of Writing. Mainly it gets divided in Printed Character Recognition method and Handwritten Character Recognition method. The Handwritten Character Recognition method further gets divided into CursiveScript andHand printed Character.

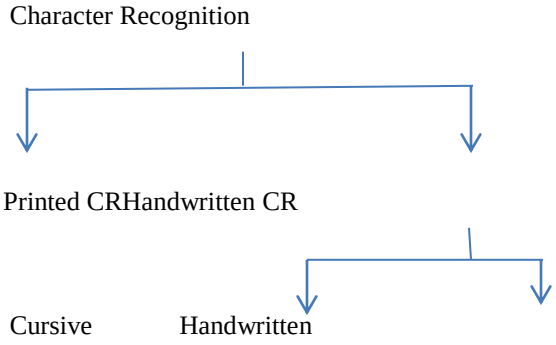


Figure 2.1: Classification based on Mode of Writing

*2.1 Printed Character Recognition*

The printed texts include all the printed materials such as books, newspapers, magazines and documents which are the outputs of typewriters, printers or plotters. The problem of printed character recognition is relatively well understood and solved with little constraints. When the documents are typed on a high quality paper with modern printing technologies, the available systems yield as well as 99% recognition accuracy. Many powerful and efficient techniques are available for the recognition of printed documents and the recognition rates of the commercially available products are very much dependent on the age of the documents quality of paper and ink which may result in significant data acquisition noise. On the basis of the capabilities and complexities printed characters can be further classified as Fixed-font Character Recognition, Multi-font Character Recognition and Omni-font Character Recognition.

*2.2 Handwritten Character Recognition*

Handwriting recognition has been one of the most fascinating and challenging research areas in field of image processing and pattern recognition in the recent years. It contributes immensely to the advancement of an automation process and can improve the interface between man and machine in numerous applications. Several research works have been focusing on new techniques and methods that would reduce the processing time while providing higher recognition accuracy.Handwritten character recognition, based on the form of written communication, can be divided into two categories: cursive script and hand printed characters. In practice, however, it is difficult to draw a clear distinction between them. A combination of these two forms can be seen frequently. Hand written character recognition systems have still limited capabilities even for recognition of the Latin characters and It is also the most difficult part of the OCR area, because depending on the style of the writer and the speed of writing, some characters may vary in size or shape in stroke number and order dynamic variation Human visual system is insensitive to the position orientation and size changes of characters. However, representation of the knowledge of these variations affects the recognition rates a great deal.

## II.COMPARISON BETWEEN OFF-LINE AND ON-LINE METHODS OF CHARACTER RECOGNITION

TableI.Off-line/Online comparison

| **Parameter** | **Off-Line** | **On-line** |
|---|---|---|
| Efficiency | Good | Good |
| Method | Challenging | Easy |
| Recognition Rate | Extremely High | Comparatively less |

## III.CONCLUSION

This paper explains the different methodologies of Character Recognition. This idea classifies the various Character Recognition methods. This paper also discusses available CR methods in detail with the comparative analysis of Off-line and On-line method. Each method having merits and demerits. By discussing all the methods, it can be conclude that, Off-line CR method is having more challenges so as to recognize a character. Since, a handwritten character varies from person to person. Though Off-line method is having challenges but still this method is having comparatively greater recognition rate than On-line method.

## REFERENCES

[1]Yusuf Perwej and AshishChaturvedi, "Neural networks for handwritten english alphabet recognition," *Int. J. of Computer Applications (0975 – 8887)*, vol. 20, no. 7, pp. April 2011

[2] Abdul Rahiman M and Rajasree M S, "An Efficient Character Recognition System for Handwritten Malayalam Characters Based on Intensity Variations," *Int. J. of Computer Theory and Engineering*, vol. 3, no. 3, June 2011.

[3]Anilkumar N. Holambe and Dr.Ravinder.C.Thool, "Printed and Handwritten Character Number Recognition of Devanagari Script using SVM and KNN," *Int. J. of Recent Trends in Engineering and Technology*, vol. 3, no. 2, May 2010.

[4] Louisa Lam and Ching Y. Suen, "An Evaluation of Parallel Thinning Algorithms for Character Recognition,"*IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 9, September 1995.

[5] Nei Kato, Shin'ichiroOmachi, HirotomoAso, Member, and Yoshiaki Nemoto, "A Handwritten Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 21, no. 3, March 1999.