

Recognition for Handwritten English Letters: A Review

Nisha Sharma, Tushar Patnaik, Bhupendra Kumar

Abstract -- Character recognition is one of the most interesting and challenging research areas in the field of Image processing. English character recognition has been extensively studied in the last half century. Nowadays different methodologies are in widespread use for character recognition. Document verification, digital library, reading bank deposit slips, reading postal addresses, extracting information from cheques, data entry, applications for credit cards, health insurance, loans, tax forms etc. are application areas of digital document processing. This paper gives an overview of research work carried out for recognition of hand written English letters. In Hand written text there is no constraint on the writing style. Hand written letters are difficult to recognize due to diverse human handwriting style, variation in angle, size and shape of letters. Various approaches of hand written character recognition are discussed here along with their performance.

Index Terms— Offline Hand written Character Recognition, Pre-Processing, Feature Extraction, Classification, Post Processing.

I. INTRODUCTION

Optical Character Recognition (OCR) is one of the most fascinating and challenging areas of pattern recognition with various practical applications. It can contribute immensely to the advancement of an automation process and can improve the interface between man and machine. It is the mechanism to convert machine printed, hand printed or hand written document file into editable text format. Typically, there are two different categories of handwriting character recognition: off-line and on-line. In Online character recognition handwriting is captured using a special pen in conjunction with electronic surface. In Offline character recognition Input has been scanned from a surface such as sheet of paper and stored digitally. Offline character recognition include recognition of machine printed, hand printed and handwritten characters. The most difficult problem in the field of OCR is the recognition of unconstrained cursive handwriting. Place during the eighties. During nineties, a fresh interest developed with the rise of new needs. The existing tools for modeling are not yet sufficient with respect to performance due to many variations of human handwriting. The similarities in distinct character shapes, the overlaps, and interconnection of the neighboring characters further complicate the problem. This paper discusses various methodologies for recognition of hand written letters. Hand written character recognition system has major five stages, namely: Pre-processing, Segmentation, Feature Extraction, Classification and Post-processing as shown in Figure 1.

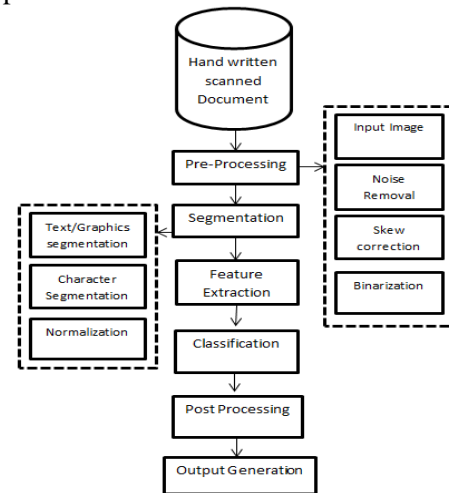


Fig 1. Major Steps of an OCR System

Section II of this paper discusses the need of pre-processing before the input is forwarded to input engine. Section III gives an overview of segmentation and various techniques which have been used in handwritten character segmentation. Section IV discusses the various methods of uniquely extracting the features of letters along with their performance. Section V discusses the classification of the characters along with post processing. Section VI gives the comparative study of various techniques applied and their result. Section VII and Section VIII discuss the future scope and the conclusion of overall paper respectively.

II. PRE-PROCESSING

Pre-Processing Can Be Defined As Cleaning The Document Image And Making It Appropriate For Input To The OCR Engine. Major Steps Under Pre-Processing Are:

- Noise removal
- Skew detection/correction
- Binarization

The Noise introduced by the optical scanning devices in the input leads to poor system performance. These imperfections must be removed prior to character recognition. Noise can be introduced in an image during image acquisition and transmission. Noise can be of different types as Gaussian noise, Gamma noise, Rayleigh noise, Exponential noise, Uniform noise, Salt and pepper noise, Periodic noise etc. Noise can be removed using Ideal filters, Butterworth filters and Gaussian filters. There is a possibility of rotation of image while scanning. Skew detection and correction is used to align the paper document with the coordinate system of scanner. Various skew detection techniques are projection profiles, connected components, Hough transform, clustering etc. In Binarization, colour or grey-scale image is converted into

binary image with the help of thresholding. Binary image can be achieved using Adaptive thresholding, Global thresholding, variable thresholding, Ostu's method etc. Morphological operations are also used in pre-processing. **Dilation and Erosion** are the morphological operations that increase or decrease the image size. Erosion makes an object smaller by eroding away the pixels from its edges. Every object pixel that is touching background pixels is changed into background pixel. However, dilation makes an object larger by adding pixels around its edges. Every pixel that is touching an object pixel is changed into object pixel. Other morphological operations are opening and closing.

III. SEGMENTATION

Segmentation is needed since handwritten characters frequently interfere with one another. Common ways in which characters can interfere include: overlapping, touching, connected, and intersecting pairs etc. In order to separate text from graphs, images, line, text/graphics segmentation is required. The output should be an image consisting of text only. Character segmentation will separate each character from another. It is one of the main steps especially in cursive scripts where characters are connected together. The isolated characters obtained as a result of character segmentation are normalized to specific size for better accuracy. Features are extracted from the characters with the same size in order to provide data uniformity. **Christopher E. Dunn and P. S. P. Wang [5] used a series of region finding, grouping, and splitting algorithms.** Region finding will identify all the disjoint regions. The pixels are originally labeled On/Off where "on" signifies the data areas. Image is examined pixel by pixel until "on" value is found. Once found it is labeled with new region number and its neighbors are searched for additional "on" value. Search proceeds until no "on" value is found. The result is that all disjoint regions will be identified and all pixels in any region will be labeled with a unique number. Grouping deals with the characters which have separate parts or which are broken. **A smallest bounding box is calculated** that completely encloses another region. If for any two regions the bounding box of one region completely encloses another region, then the enclosed region is relabeled to the value of the enclosing region. Thus, the resulting region is composed of two disjoint sub-regions. This is helpful for connecting regions that have been separated due to noise. Splitting [5] deals with touching characters. **Anshul Mehta [2] used Heuristic segmentation algorithm which scans the hand written words to identify the valid segmentation points between characters.** segmentation is based on locating the arcs between letters, common in handwritten cursive script. For this a histogram of vertical pixel density is examined which may indicate the location of possible segmentation points in the word. Other character segmentation approaches [4] are Thinning based method, Contour Fitting method, Robust Statistical technique, Hypothesis Verification, Shape Feature Vector method etc.

IV. FEATURE EXTRACTION

Feature extraction is finding the set of parameters that define the shape of a character precisely and uniquely. Feature extraction [3] methods are classified into three major groups as:

- Statistical features.
- Global transformation and series expansion.
- Geometric and topological features.

Statistical features represent the image as statistical distribution of points. Various methods which use statistical features are Zoning, Crossings and Distances, Projections etc. In global transformation and series expansion various techniques are Fourier transform, Gabor transform, Fourier Descriptor, wavelets, moments, Karhunen-Loeve expansion etc. In Geometric and topological features, the structural features like loops, curves, lines, T-point, cross, opening to the right, opening to the left etc. are used. The various categories are coding (freeman chain code), extracting and counting topological structures, graphs and trees. Geometric features are used along with fuzzy logic to recognize characters [7]. Adnan Amin [6] and Puttipong Mahasukhon [7] used structural information to extract features from a character like Breakpoints, Inflection Point, Cusp Point, Straight Line, Curve, Open or Close Loop etc. Breakpoint divides a path into sub paths. It has two possible conditions- Inflection Point (change in curvature) and Cusp Point (sharp change in direction). Straight line has two points in sequence in a path. Open curve is as in letter "S". Closed curve is as present in "a". These segments are given as input to neural network classifier. Anshul Mehta [2] used Fourier descriptor for extracting unique feature from a character. Initially boundary is detected, then discrete Fourier coefficient $a[k]$ and $b[k]$ are calculated for $0 < k < L-1$. Where L is the total number of boundary points. Fourier descriptor [8] can be used with one new technique known as Border Transition technique (BTT). In it each character is partitioned into four equal quadrants. The scanning and calculation of black-to-white transition take place in both vertical and horizontal directions in each quadrant. The average transition of each direction (horizontal and vertical) in each of the four quadrants of the box surrounding the character will be calculated. Rafael M. O [1] used nine modified feature extraction techniques on a single database. Structural characteristics consist in extracting histograms and profiles and combining then into a single feature vector. In modified edge map an $M \times N$ image is thinned and scaled into a 25×25 matrix. The Sobel operators are used to extract four distinct edge maps: horizontal, vertical and two diagonals. These four maps and the original image are divided into 25 sub-images of 5×5 pixels each. The features are obtained calculating the percentage of black pixels in each sub-image (25 features per image). The features are combined to form a single feature vector containing 125 (25×5) features. Image Projections consists of extracting the radial and diagonal projections. To extract the radial projections, the image must

first be divided into four quadrants: top, bottom, right and left. Radial projections are obtained by grouping pixels by its radial distance to the center of the image in each quadrant separately. The diagonal projection is computed simply by grouping pixels by the two diagonal lines. The values of each projection are normalized to a range [0-1] through the division by the maximum value. The normalized features are Concatenated in a single vector containing 128 features. In Multi Zoning an $M \times N$ character image is divided into several sub-images and the percentage of black pixels in each Sub-image is used as feature. It is a statistical approach as features are calculated based on the number of pixels used to represent an image. Other feature extraction algorithms used are Concavities Measurement, MAT-based Gradient Directional features, Gradient Directional features, Median Gradient features, Camastra 34D features[1].

V. CLASSIFICATION AND POST PROCESSING

The classification is the process of identifying each character and assigning to it the correct character class. The classification techniques [9] can be categorized as:

- Classical techniques.
- Soft computing techniques.

The various classical techniques are Template matching, Statistical techniques, Structural techniques. Whereas the various soft computing techniques are Neural networks, Fuzzy logic, Evolutionary computing techniques. Adnan Amin and W. H. Wilson [6] used Neural network for classification of characters with three layers namely Input layer, Output layer and Hidden layer. The geometric features extracted like dot, line, curve or loops are given as input to the input layer. Each component of the segmented representation is classified as a dot, line, curve, or loop. In each case, the characteristics of the component are determined: if a line, what are its orientation and its size relative to the character frame - short, medium or long. One input neuron is used to encode each of these possible choices (short/medium/long) and each of four possible orientations for a line. One input neuron is used to encode the characteristics of each component extracted by geometric feature extraction technique. Neuron has two modes of operations as training mode and testing mode. In the training mode, the neuron can be trained to fire (or not), for particular input patterns. In the testing mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not. Anshul Mehta, Manisha Srivastava [2] used three networks for the recognition of 26 lower case and 26 upper case letters as Multilayer Perception (MLP) [2,8], Radial Basis Function (RBF) and Support Vector Machine (SVM). Multilayer perception is a feed forward neural network with one or more layers between input and output layer. Radial basis function (RBF) networks typically have three layers: an input layer, a hidden layer with a non-linear RBF activation function and a linear output layer. Post-processing mainly consists of two tasks – output string

generation and error detection/correction. Output string generation will reassemble the strings which have been separated in the process of segmentation whereas error detection/correction will correct errors with the help of dictionary.

VI. COMPARISION TABLE

The major steps of an OCR engine are feature extraction and classification. The various feature extraction techniques in combination with various classification techniques along with their result which have been used by the researchers are discussed in Table I:

Table I

Author	Feature extraction Method	Classification Method	Result
Anshul Gupta , Manisha Srivastava , Chitrakleha Mahanta[2]	Fourier descriptor with magnitude	SVM classifier	86.66%
Anshul Gupta , Manisha Srivastava , Chitrakleha Mahanta[2]	Fourier descriptor with phase	SVM classifier	98.74%
Puttipong Mahasukhon, Hossein Mousavinezhad, Jeong-Young Song[7]	Geometric features	Fuzzy theory	90%
Rafael M. O. Cruz, George D. C. Cavalcanti and Tsang Ing Ren[1]	Multi Zoning	MLP network	89.67%
Anshul Gupta , Manisha Srivastava , Chitrakleha Mahanta[2]	Fourier descriptor with magnitude and phase	SVM classifier	98.04%
Yuk Ying Chung, Man to Wong [8]	Fourier descriptor and topological properties	MLP with back propagation	96%

VII. CONCLUSION

The major approaches used in the field of handwritten character recognition during the last decade have been reviewed in this paper. Different pre-processing, segmentation, feature extraction, classification techniques are also discussed. Though, various methods for treating the problem of hand written English letters have developed in last two decades, still a lot of research is needed so that a viable software solution can be made available. The existing OCR for handwritten has very low accuracy. We need an efficient solution to solve this problem so that overall performance can be increased.

VIII. FUTURE RESEARCH

From 1950's OCR is an active area of research. Many techniques for recognition of Offline English Handwritten Characters have been suggested. But still an efficient OCR for the recognition of hand written letters does not exist. Few steps have been taken for Hand written and Hand printed (which is a constrained hand writing) English letter recognition. Various challenges are identified which may provide more lively interest to the researchers. These challenges are: difficulty to identify the diverse human writing styles, different angles of letters, different shapes and size of letter, pure input quality, low accuracy rate in recognition etc. Hence, a lot of research work is to be done to solve these problems.

REFERENCES

- [1] Rafael M. O. Cruz, George D. C. Cavalcanti and Tsang Ing Ren "An Ensemble Classifier For Offline Cursive Character Recognition Using Multiple Feature Extraction Techniques" IEEE 2010.
- [2] Anshul Mehta, Manisha Srivastava, Chitralkha Mahanta "Offline handwritten character recognition using neural network" IEEE 2011 International conference on computer applications and Industrial Electronics.
- [3] Oivind Due Trier, Anil K Jain and Torfinn Text "feature extraction methods for character recognition- a survey" 1996
- [4] Jayashree R Prashad, Dr. U V kulkarni "Trends in handwriting recognition" IEEE 2010, Tthird international conference on emerging trends in engineering and technology.
- [5] Christopher E. Dunn and P. S. P. Wang "Character Segmentation techniques for handwritten text-A Survey" IEEE 1992.
- [6] Adnan Amin and W. H. Wilson "Hand-Printed Character Recognition System Using Artificial Neural Networks" IEEE 1993.
- [7] Puttipong Mahasukhon, Hossein Mousavinezhad, Jeong-Young Song "Hand-Printed English Character Recognition based on Fuzzy Theory" IEEE 2012
- [8] Yuk Ying Chung, Man to Wong "handwritten character recognition by Fourier descriptors and neural network" IEEE 1997, Speech and Image Technologies for computing and telecommunication.
- [9] Shabana Mehfuz, Gauri katiyar "Intelligent Systems for Off-Line Handwritten Character Recognition: A Review" International Journal of Emerging Technology and Advanced Engineering 2012.
- [10] Nafiz Arica and Fatos T. Yarman-Vural "An Overview of Character Recognition Focused on Off-Line Handwriting." IEEE 2001.
- [11] Sameer singh, Adnan Amin "Neural network recognition and analysis of hand printed letters. IEEE 1998.

AUTHOR'S PROFILE



Ms. Nisha Sharma received her B.Tech. Degree from Punjab Technical University. Currently she is pursuing her M.Tech in computer science from Centre for development of advance computing (CDAC), Noida. Her interest area is Image processing, Algorithms, and Database Management Systems.



Scripts".

Mr. Tushar Patnaik (Sr. Lecturer/Sr. Project Engineer) joined CDAC in 1998. He has eleven years of teaching experience. His interest areas are Computer Graphics, Multimedia and Database Management System and Pattern Recognition. At present he is leading the consortium based project "Development of Robust Document Image Understanding System for Documents in Indian



Mr. Bhupendra Kumar (Senior Technical Officer) joined CDAC in 2005, he received his M.Tech. degree from IIT Allahabad with the specialization in wireless communication and computing. His interest areas are Advanced Image processing, pattern recognition, computer network, wireless network, MANETs. Currently he is involved in project "Development of Robust Document Image Understanding System for Documents in Indian Scripts".