Character Recognition: A Neural Network Approach

R. C. Tripathi[a]* and Vijay Kumar[b]
[a]*Department of IT, Institute of Management Studies, Noida, U.P., India.*
[b]*Department of CSE, Meerut Institute of Engineering & Technology, Meerut, U.P, India*
*ramesh_c_tripathi@yahoo.co.in

Abstract

OCR is the acronym for Optical Character Recognition. This technology allows a machine to automatically recognize characters through as optical mechanism. Human Beings "recognize" many objects in this manner; our eyes are the "optical mechanism." But while the brain "sees" the input, the ability to comprehend these signals varies in each person according to many factors. In same manner "characters" which are nothing but the images made by the different combinations of lines and curves are also optically recognized by our brain. By reviewing these variables, the challenges faced by the technologist developing an OCR system. Character recognition techniques help in recognizing the characters written on paper documents and converting it in digital form. So Character recognition is gaining interest and importance in the modern world. While the area of character recognition is vast we focus on the fundamentals of character recognition, available techniques and emphasis on more recently used technique, neural networks. The paper throws light on, one of the application of Neural Network (NN) i.e. Character Recognition.

Keywords - OCR, NN, CRS, Knowledge-base.

Introduction

Recognizing characters, letters or digits for human beings is not a big task. It can even be done by small child, but doing the same with machine is a difficult task. Machine simulation of human functions has been a very challenging research area since the advent of digital computers. It has been estimated that over 250 billion $ per year worldwide is spent on keying information from paper documents and this is for keying only one percent of the available document, most of the cost is human labour. The ultimate goal of designing a character recognition system with an accuracy rate of 100 % is quite illusionary because even human beings are not able to recognize every hand written text without any doubt. For example, many people cannot even read their own note.

Categorization of CRS

Character recognition system can be classified based upon two major criteria.

Classification According to Data Acquiring Process

Online CRS

Recognizing handwriting recorded with a digitizer as a time sequence of pen co ordinates is known as

online character reorganization. It captures the temporal or dynamic information of writing. This information consist of pen strokes (i.e. the writing from pen down to pen up), the order of pen strokes the direction of writing and the speed of writing within each stroke.

Offline CRS

It is also known as "optical character recognition" because the image of writing is converted into a bit pattern by an optically digitized device such as optical scanner or camera. The bit pattern data is shown by matrix of pixels. These matrixes can be very large so in order to reach the complexity and to insert much data in recognition most scanners are designed to have x-y resolution of typically 100 – 1600 dots per inch.

Categorization According to Text Type

Printed CRS

Printed text includes all the printed materials such as book, newspaper, magazine and documents which are the output of typewriters, printers or plotters. The recognition rate is very much dependent on the age of the documents, quality of the paper and ink which may result in significant data acquisition noise (Sarhan *et al.*, 2007).

Hand written CRS

Hand written character recognition, based on the form of written communication can be divided into two categories: Cursive script and Hand printed characters. It is the most difficult part of character recognition area because depending on the style of the writer and the speed of the writing some character may vary in size and shape (Araokar, 2005).

Model for Character Recognition

For any character recognition system there are four major stages, as shown in the figure below:

- Pre-processing
- Segmentation
- Feature extraction
- Training and recognition / classification (Araokar, 2005).

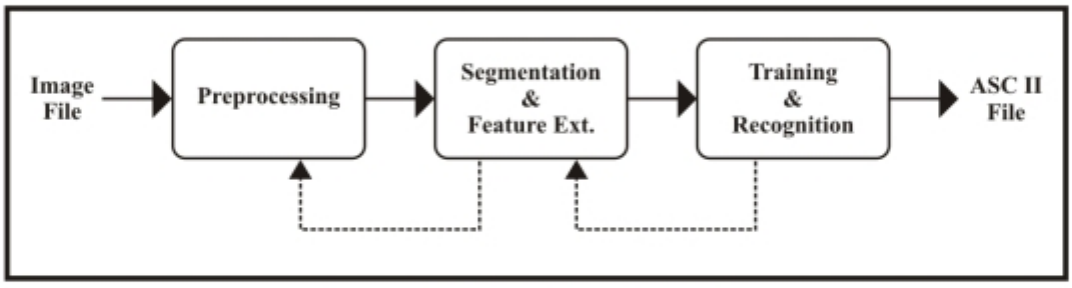

Figure 1: Character Recognition Model

Pre-processing

Pre-processing is done prior to the application of segmentation and feature extraction algorithms. It aims to produce clean document images that are easy for the character recognition system to operate accurately. The major objective of pre-processing are as listed below:

Noise reduction

The noise which is introduced by the optical scanning devices or the writing instruments causes disconnected line segments, bumps and gap in lines, filled loop etc. There are many techniques to reduce noise like filtering and morphological operations. Normalization: Normalization methods aim to remove all type of variations during the writing and obtain standardized data. The basis methods for normalization are skew normalization, slant normalization and size normalization.

Compression

In addition to classical loss less image compression techniques the character can be further compressed by thresh holding and thinning algorithms.

Segmentation

The pre-processing stage yields a clean document in the sense that maximum shape information with minimal noise on normalize image is obtained. The next stage is segmenting the document into its sub components and extracting the relevant features to feed the training and recognition stages (www.xilinx.com, 2009).There are two types of segmentations as listed below.

External segmentation

It is isolation of various writing units such as paragraph, sentences or word prior to recognition. It decomposed the page layout into its logical parts. This is done in two ways i.e. Structural analysis and Functional analysis.

Internal segmentation

It is an operation that seeks to decompose an image of sequence of characters into sub images of individual symbols. Character Segmentation strategies are divided into following categories.

Segmentation by dissection method

It identifies the segments based on character like properties. This process of cutting of image into meaning full components is given a special name, dissection. Dissection is an intelligent process that analyses an image without using any specific class of shape information. Available method based on dissection of an image is connected component analysis.

Recognition based segmentation

It searches the image for components that match pre defined classed. Segmentation is performed by use of recognition confidence including syntactic or semantic correctness of the overall result.

Feature Extraction

During or after the segmentation procedure the feature set which is used in the training stage is extracted. A good feature set plays one of the most important roles in a recognition system. A good feature set should represent characteristic of a class that help distinguish it from other classes. The features can be classified as below:

Global transformation and series expansion features

These features are invariant to global deformation and rotations. One way to represent a signal is by linear combination of a series of simpler well defined functions. Statistical features: These features are derived from the statistical distribution of points. They provide high speed and low complexity and may also be used for reducing the dimension of feature set.

Geometrical and topological features

These features may represent global and local properties of character and have high tolerance to distortion (Hussain *et al*., 1994).

Training and Recognition Techniques

As in many areas of image analysis, character recognition systems extensively use the methodologies of pattern recognition which assigns unknown samples into pre defined classed. Character recognition finds its roots from pattern recognition.

Pattern recognition is a study of how machines can observe the environment, learn to distinguish patterns of interest from their backgrounds and make sound and reasonable decisions.

Recognition which assigns unknown samples into pre defined classed. Character recognition finds its roots from pattern recognition.

Pattern matching

Here the memory representation is a holistic unanalyzed entity (a template). An input pattern is compared to the stored representation. The identity is determined by the selection of template with the greatest amount of overlap. The stored representation is the description of past inputs in terms of list of attributes or features. Inputs are broken down into a small list of constituent features. Identity is determined by selecting the feature list most similar to the input, but this technique faces some severe problems. This technique is intolerant to deviations. Large number of template is required and cannot support similarity difference judgments (Jain *et al*., 2000; www.recognetics.com, 2008).

Statistical approach

Here each pattern is represented in terms of -d- features or measurements and is viewed as point in a -d-dimensional space. The goal is to choose those features that allow pattern vectors belonging to different categories to occupy compact and disjoint regions in -d-dimensional features. Given a set of training patterns form each class the objective is to establish decision boundaries in the feature space with separate patterns belonging to different classes. One way to do this is by clustering analysis. The cluster of characters, features which represent distinct classed are analyzed by way of clustering methods. Clustering can be performed either by an agglomerative or a divisive algorithm. The agglomerative algorithm operates step by step merging small clusters into large ones by a distance criterion. The divisive method splits the characters under the certain rules for identifying the underlying characters.

Artificial neural networks

A neural network is defined a computing architecture that consist of massively parallel interconnection of simple neural process. Because of its parallel nature it can perform computation at a higher rate compared to the classical techniques. A neural network contains many nodes. The output from one node is feed to another one in the network and the final decision depends on the complex interaction of all nodes (Sarhan *et al*., 2007).

Neural Networks

The paper discusses neural networks methodology in context with character recognition. Classical methods of pattern recognition are not considered to be so successful for recognition of characters due to following reasons:

- The same character differs in size, shape and style from person to person and from time to time with the same person.

- Like an image visual characters are subject to spoilage due to noise.
- There is no specific rule that defines the appearance of visual character.

Training the network on different set of noisy data forced to learn how to deal with noise. The back-propagation method is used for learning in neural network (Demuth *et al*., 2006).

In contrast to limitations of classical methods of pattern recognition neural network gains more success because of its humanoid nature (i.e. functioning like human) (Leiva *et al*., 2007). Some of its characteristics are as stated below:

- It is adaptive to minor changes and errors in visual pattern.
- It learns from prior experience.

The recognition of optical characters is a problem of relatively amenable complexity when compared with challenges such as human faces. Artificial intelligence has a considerable success due to humanoid qualities such as adapting from changes and learning from prior experience.

Image digitization: The process of digitization is important for neural networks. In this process the input image is sampled into binary window which forms the input to the recognition system. The sample of this process is shown in the figure below:

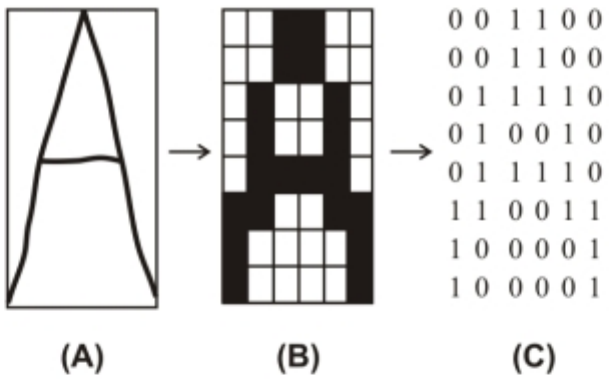

Figure 2: Image Digitization Process

In the figure the alphabet A is digitized into 6X8=48 cells, each having a single color black or white.

This is done in order to make the computer understand the form. And in the process of digitization cell with color black is further assigned value +1 and the cell with color white is assigned a value 0, to give it binary structure. And this creates a binary image matrix I (Kahan *et al*., 1987). This makes input image invariant of actual dimensions architecture of Neural Network studied:

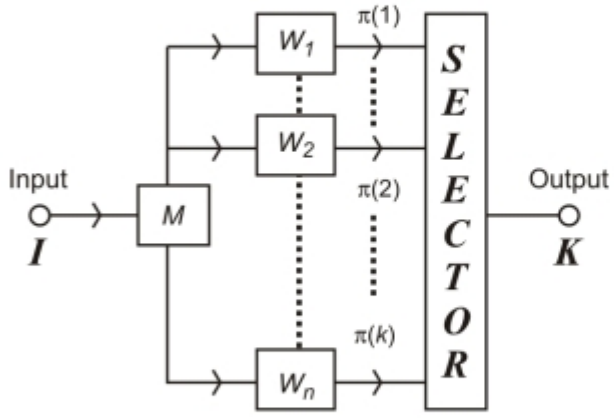

Figure 3: Architecture of the Neural Network

The architecture of the neural network which formed the basis for this study is as shown in the figure above. Here the input to network is pattern I. The block M provides the input matrix M to the weighted blocks Wk for each K. There are total of n weighted blocks for n characters to be learned (Hussain *et al*., 1994; Luger, 2005).

Learning Process

In this process a character is assigned to the network and is a given a label. Several other patterns of the same character are taught under the same label, due to which system learns several variations of a single pattern and gets adaptive to it. During this training process the input assigned is matrix *M* defined as below:

$$If I(i,j) = 1 \; Then \; M(i,j) = 1$$

Else:

$$If I(i,j) = 0 \; Then \; M(i,j) = -1$$

In this method of learning, each character to be taught; processes corresponding weight matrix. For the *Kth* character to be taught its weight matrix is denoted by *Wk*. As the learning of the character progresses it is the weight of the character to be updated. At the commencement of teaching (supervised training), this matrix is initialized to zero. Whenever a character is to be taught to the network, an input pattern representing that character is submitted to the network.

The network is then instructed to identify this pattern as, say, the k^{th} character in a knowledge base of characters.

That means that the pattern is assigned a label k. In accordance with this, the weight matrix *Wk* is updated in the following manner:

```
For all i= 1 to x
{
For allj=1 to y
{
Wk (i,j)=Wk(i,j)+M(i,j)
}}
```

Here *x* and *y* are the dimensions of the matrix *Wk* (and *M*).

The figure below shows the digitization of three input patterns representing *S* that were presented to the system for it to learn.

(a) (b) (c)

Figure 4: Digitization of Three Input Patterns (S)

We can see that the patterns slightly differ from each other, just as handwriting differs from person to person (or time to time) and like printed characters differ from machine to machine (Nafiz, 1998).

The figure below gives the weight matrix, say, *WS* corresponding to the alphabet *S*. The matrix is has been updated thrice to learn the alphabet *S*. This matrix is specific to the alphabet *S* alone. Other characters shall each have a corresponding weight matrix.

1	3	3	3	3	1
3	3	-3	-3	-1	-1

3	-1	-3	-3	-3	-3
3	3	1	-1	-1	-1

-1	3	3	3	3	3

-3	-3	-3	-3	-3	-3

3	-3	-3	-1	1	3
3	3	3	3	3	1

Figure 5: Weight Matrix for the Alphabet S

1. The matrix-elements with higher (positive) values are the ones which stand for the most commonly occurring image-pixels.

2. The elements with lesser or negative values stand for pixels which appear less frequently in the images. Neural networks learn through such updating of their weights. Each time, the weights are adjusted in such a manner as to give an output closer to the desired output than before.

Some statistics that form the base for concluding recognition are:

Candidate Score

This statistic is a product of corresponding elements of the weight matrix Wk of the kth learnt pattern and an input pattern I as its candidate.

Ideal Weight-Model Score

This statistic simply gives the sum total of all the positive elements of the weight matrix of a learnt pattern. It may be formulated as follows (with ideal weight model score for K initialized to 0 each time).

Recognition Quotient

This statistic gives a measure of how well the recognition system identifies an input pattern as a matching candidate for one of its many learnt patterns.

The neural system has some advantages:

- The method is highly adaptive; recognition is tolerant to minor errors and changes in patterns.

Applications of Character Recognition

- Personal organizer, Personal communicator, Notebook.
- Data acquisition devices for order entries, inspection, inventories, survey etc.
- Large scale data processing such as postal address reading, cheque sorting.
- Shorthand transcription.
- Reading aids for visually handicapped.

Conclusion

The basic idea of using extracted features to train a Neural Network seems to work, although the success rate is not impressive, it could have been worse. There are several possible changes that could improve the performance.

In this paper the neural network approach explained here shows the learning ability and adaptability of neural networks. Despite the computational complexity involved, artificial neural networks offer several advantages in pattern recognition and classification in the sense of emulating adaptive human intelligence to a small extent.

References

Araokar, S. 2005. Visual character recognition using artificial neural networks. *MGM's College of Engineering and Technology*.

Arica, N. 1998. An Overview Of Character Recognition Focused On Off-line Handwriting. *A thesis submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University*.

Cognimem, 2008. CogniMem_1K: Neural network chip for high performance pattern recognition, datasheet, Version 1.2.1, www.recognetics.com.

Demuth, H., Beale, M., Hagan, M. 2006. Neural network toolbox for use with MATLAB. *Neural Network Toolbox, IEE Savoy Place, London*.

Hussain, B., Kabuka, M.R. 1994. A novel feature recognition neural network and its application to character recognition. *IEEE Transactions of Pattern Recognition and Machine Intelligence*,16(1),98-106.

Jain, A. K., Duin, R. P.W., Mao, J. 2000. Statistical Pattern Recognition. *IEEE transaction on Pattern Analysis and Machine Intelligence*, 22(1), 4-37.

Kahan, S.T., Pavlidis, T., Baird. 1987. On recognition of Printed characters of any font and size. *IEEE Transactions of pattern recognition and machine intelligence*, PAM1-9, 274-285.

Leiva, L., Vázquez, M., Acosta, N., Sutter, G. 2007. Herramienta de Generación de Arquitecturas Hardware para Reconocimiento de Patrones en Imágenes. *JCRA Jornadas de Computación Reconfigurable y Aplicaciones, Zaragoza, España*.

Luger, G. F. 2005. Artificial Intelligence, Structures and Strategies for Complex Problem Solving. *Addison-Wesley*.

Sarhan, N.S., Laheeb, A. 2007. Recognition of printed Assyrian character based on Neocognitron Artificial Neural Network. *International Arab Journal of Information technology*, 4(1), 82-89.

Xilinx, 2009. Inc. Xilinx Synthesis Technology (XST) User Guide. UG627 (v 11.1.0) www.xilinx.com.

