



Handwriting Recognition System – A Survey

Authors

Pooja Yadav & Ms. Neha Popli

ITM University, Gurgaon

Email: pooja.yadav113@gmail.com

nehapopli@itmindia.edu

Corresponding Author

Pooja Yadav

ITM University, Gurgaon

Email: pooja.yadav113@gmail.com

Abstract:

Handwriting Recognition System has been studied in the last few decades. Many approaches are presented to recognize the hand written documents or paper. These approaches focus on how we recognize our hand written words and lining documents. Today there is no. of application area of handwriting recognition system. So an overview of hand writing recognition system and their evolution is presented by available technique with their superiorities and limitations are reviewed. So current status of handwriting recognition is focusing on off-line and online handwriting recognition system. This overview represent as an update for the state of art in the hand writing recognition field.

Keywords – Handwriting recognition, offline and online handwriting recognition, segmentation, feature extraction and training.

INTRODUCTION

Handwriting Detection is a technique or ability of a Computer to receive and interpret intelligible Handwritten Input from source such as paper documents, touch screen, photo graphs etc. When we write something and give it as an input to the system then how we recognize those words and hand written documents. This process is done by Handwriting recognition system. So we can say that handwriting recognition system can be used as umbrella term. A machine simulation of human writing is an application area of handwriting recognition system. Handwriting recognition system is a way of interface between human and machine. There are no. of way for write a word like a word can be write in cursively , upper case letter and lower case letter. So handwriting recognition techniques have the ability to recognize cursive word etc.

From few decades handwriting recognition system is most growing research area. Because today the needs of handwriting recognition system are increasing day by day. Because writing is a natural way of storing the information and transmitting the information. Handwriting doesn't provide a way of

communication between human to human only it also provides the interface between human and machine. It is way of automatic transformation bulk amount of papers and documents into machine. It also provides web interface to web documents.

Handwriting recognition system is classified into two categories online handwriting recognition system and offline handwriting recognition system.

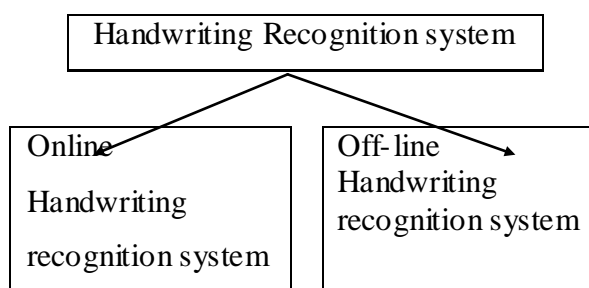


Fig.1 types of Handwriting recognition system.

In online mode the handwriting is provided by an electronic pen or a mouse and writer's pen movement will captured, stroke order will also count.

In offline mode the writing is captured by scanner and complete writing will be present as an image. Pixels of images are captured.

There are number of application area are presented where handwriting recognition system is using. In bank when we fill a cheque then the amount which we fill in the cheque that amount is hand written then handwriting recognition system is used to recognize those words and help to find out the right amount. Handwriting recognition system is used in machine simulation. When we give any hand written paper or documents as an input then handwriting recognition system first convert this documents in ASCII code then it will used as an

input. So several applications like mail sorting, bank processing, documents reading and postal address recognition need offline handwriting recognition system. So offline handwriting approaches is an active area for exploring the newer techniques.

In handwriting recognition system the input can be in two forms one can be in printed form another can be in handwritten words. Sample of these inputs are



Fig2. Hand written word

In hand written mode words can be in cursive form or it can be in lower and upper case letter. Hand recognition system is more difficult to perform in cursive mode because hand written word depends on user. Different types of user write the different types of writing.

PREVIOUS WORK IN HANDWRITING RECOGNITION SYSTEM

Handwriting is the most natural mode of collecting the different types of information and transfers that information through centuries. Now Handwriting recognition system is medium of communication between human and machines. It provides an efficient application for transferring the data into machines.

In 1900-1980 ages - History of Handwriting recognition system can be studied from 1900s when the Russian Scientist was trying to make an aid for visually handicapped. The first Handwriting recognizer introduced in middle of 1940s when Digital computer discovered. Early work on the automatic recognition of handwriting has been focused on machine printed text or hand written symbols. In machines printed recognition system an image compared to a library of images this is known as template matching. In this image processing on binary images were used for feature extraction.

In 1950s commercial handwriting recognizer available in this an electronic tablets capture the x-y co-ordinate of data by pen-tip movement.

In 1980-1990 ages - In 1990s structural techniques were introduced in many systems with statistical methods. In these approaches systems broke the character into set of patterns like lines and curve. So handwriting recognizer was focused on shape of the words.

After 1990s – The real progress is achieved during this period using new techniques and methodologies, Image processing, pattern recognition methods combined with artificial intelligence techniques. Today more powerful computers and more accurate electronic equipments like scanner, electronic tablets are used. Many methodologies such as Neural Network, HMM (Hidden Markov Models) and Fuzzy Set Reasoning are using in handwriting recognition system.

LITERATURE REVIEW

A systematic literature search for hand writing recognition was started during the period 1900-2009 with the keyword neural network.

3.1 Handwriting recognition using word length estimation

–This research (C.Allgrove et al, 1998) describe a method which can be used to estimate the length of hand written word. This method defines a number of components with recognition technique. It aims to directly identify the number of letters in the word as supporting information to aid more handwriting recognition technique. The method of word length estimation has many applications in area of text analysis. The main application of this method in bank cheque processing for recognition of the legal amount field. This method has been tested on two different languages, English and French.

There are three main issues that need to be taken into account when selecting the features that are used in the method of word length estimation for hand writing recognition.

The first issue that needs to be addressed is unconstrained nature of the data. This problem is not unique in word length estimation, but forms the main limiting factor in the performance of hand writing recognition technique.

The second main issues that each distinct class contains many different words. The fig shows the word that make up the class of four letter word in the application in bank cheque processing.

The third issue indicate that the feature set also needs to be able to lend itself to a computationally simple implementation.

3.2 Hand writing recognition using Whiteboard

Notes – This research (Moore, 2002; Reiter and Rigoll et al, 2004) motivated by smart meeting room applications, where not only speech and data are recorded but also hand written notes on a white board also recorded. In order to allow for indexing and browsing of the data collected during various meetings, the automatic recognition of whiteboard notes, i.e. their transcription into ASCII format, is needed.

In this the eBeam interface is used for recording the handwritten notes. It allows us to write on a white board using a normal pen, which sends infrared signals to a triangular receiver mounted in one of the corner of the white board. The X-Y coordinates representing the location of the tip of the pen together with a time stamp for each location. The data is in XML format and the frame rate of the recordings varies from 30 to 70 frames per second.



Fig3. Illustration of the recording

3.3 Hand writing recognition using neural network

– This research motivated by feed forward neural network. A new method called diagonal based feature extraction is introduced for extracting the feature of the handwritten alphabets. Fifty data sets each are containing 26 alphabets written by various people. The proposed recognition system performs quite well yielding higher level of recognition accuracy compared to

the system employing the conventional horizontal and vertical methods of feature extraction. This system is suitable for converting handwritten documents into structural text form and recognizing handwritten names.

There are basic six steps of recognition system which we used in Handwriting recognition using neural network

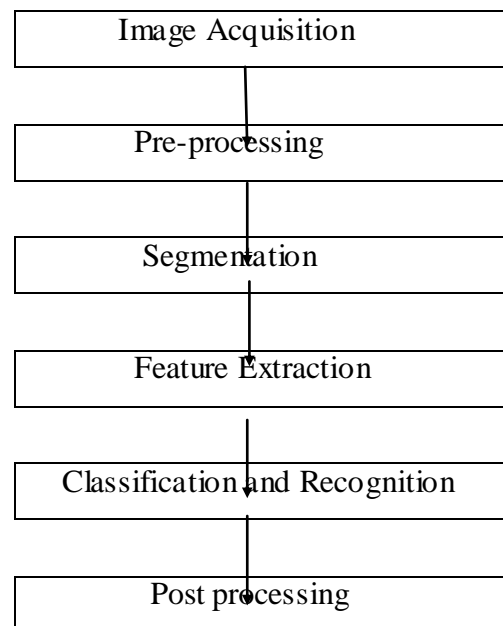


Fig.4 systematic diagram of the proposed recognition system

3.3.1 Image Acquisition – In this the recognition system acquires a scanned image as an input image. The image should have a specific format such as JPEG, JPG and BMT etc. This image is acquired through scanner, digital camera any other suitable digital input device.

3.3.2. Pre-Processing -The main objectives of pre-processing are:

- Noise removing,
- Binarization,
- Edge Detection,
- Dilation and Filling,
- Processed image for feature extraction,
- Normalization of the data,

3.3.3. Segmentation – In the segmentation stage, an image of sequence of character is decomposed into sub image of individual character. Each individual character is uniformly resized into 90*60 pixels for classification and recognition field. Two types of segmentation are

- External segmentation
- Internal segmentation

3.3.4. Feature Extraction Method – This is an important stage as its effective Functioning improves the recognition rate and reduces the misclassification. Feature extraction can be performed by

- Diagonally
- Vertically
- Horizontally

3.3.5. Classification and Recognition – For classification neural network use feed forward back propagation neural network. The network training parameters are

- Input nodes: 54/69,
- Hidden nodes: 100 each,
- Output nodes: 38(26 alphabets, 10 numerals and 2 special symbols),

-Training algorithm: Gradient Decent with momentum and adaptive learning,

- Perform function: Mean Square error,
- Training momentum constant: 0.9,
- Training epochs: 1000000

$$a^i = \log \text{sig} (w^i a^{i-1} + b^i)$$

Where

$$I = 1, 2, 3, 4, \dots, n$$

w^i = weight vector of i^{th} layer

A^i = output of i^{th} layer

B^i = Bias vector of i^{th} layer

3.3.6. Post-Processing – After classification, post-processing takes place and find out that character will belong to which block. The incorporation of context and shape information in all the stages of Handwriting systems is necessary for meaningful improvements in recognition rates. This is done in the post processing stage with a feedback to the early stages of Handwriting. The simplest way of incorporating the context information is the utilization of a dictionary for correcting the minor mistakes of the Handwriting systems.

CONCLUSION

This review paper focuses on various Handwriting techniques available and makes a comparative study between them. We conclude that modern techniques are more beneficial compared to traditional ones. Modern techniques are embedded with features with the help of which we can overcome the drawbacks of the traditional approach. An improved approach in almost all stages of Handwriting research is needed in case of texts which are handwritten under poor conditions or for free style handwriting. Number digit or limited vocabulary form (bank checks, envelopes and forms designed for specific applications) recognition is a popular application area.

REFERENCES

- 1) Bakis, R. 1976. Continuous speech recognition via centisecond acoustic states. In Proc. 91st Meeting of the Acoustic Society in America.
- 2) Dempster, A.P., N.M. Laird and D.B. Rubin. 1977. "Maximum Likelihood from

- Incomplete Data via the EM Algorithm”
Journal of Royal Statistical Society B
39(1):1–38.
- 3) Fink, G. A., M. Wienecke and G. Sagerer. 2001. Video-based on-line handwriting recognition. In Proc. 6th ICDAR. pp. 226–230.
 - 4) Forney, G. D. 1973. The Viterbi algorithm. In Proc. IEEE. Vol. 61 pp. 268–278.
 - 5) G`unter, S. and H. Bunke. 2004. “HMM-based handwritten word recognition: on the optimization of the number of states, training Iterations and Gaussian components.” Pattern Recognition 37:2069–2079.
 - 6) Johansson, S. 1986. The tagged LOB Corpus: User’s Manual. Norwegian Computing Centre for the Humanities, Norway.
 - 7) Marti, U.-V. and H. Bunke. 2001. “Using a statistical language model to improve the performance of an HMM-based cursive handwriting recognition system.” IJPRAI 15:65 – 90.
 - 8) Marti, U.-V. and H. Bunke. 2002. “The IAM-database: an English sentence database for offline handwriting Recognition.” IJDAR 5:39 – 46.
 - 9) Moore, D. 2002. The IDIAP Smart Meeting Room. Technical report IDIAP-Com.
 - 10) BLUMENSTEIN, M., and VERMA, B. (1999): A new segmentation algorithm for handwritten word recognition. Proc. 11) IJCNN International Joint Conference on Neural Networks, Washington DC, USA, 872–882.
 - 12) BUNKE, H., ROTH, M., and SCHUKAT-TALAMAZZINI, E.G. (1995): Offline cursive handwriting recognition using Hidden Markov models. Proc. Pattern Recognition, 28(9): 1399–1413.
 - 13) EASTWOOD, B., JENNINGS, A., and HARVEY, A. (1997): A feature based neural network segmenter for handwritten 14) Words. Proc. ICCIMA International Conference on Computational Intelligence and Multimedia Applications, Gold Coast, Australia, 286–290.
 - 15) C. Y. Suen, C. C. Tappert, T. Wakahara, “The State of the Art in On-Line Handwriting Recognition”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 8, pp.787-808, 1990.
 - 16) Bazzi, R. Schwartz, J. Makhoul, “An Omnifont Open Vocabulary OCR System for English and Arabic”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol.21, no.6, pp.495-504, 1999.