International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 8 206 – 210

Online Handwriting Recognition using HMM

Ms. Sarika Dadasaheb Lokhande
M.E. Computer Engineering
ARMIET College, Asangaon,
Maharashtra, India.
sarika15.l@gmail.com

***Abstract*—** Basically handwriting recognition can be divided into two parts as Offline handwriting recognition and Online handwriting recognition. Highly accurate output with predefined constraints can be given by Online handwriting recognition system as it is related to size of vocabulary and writer dependency, printed writing style etc. Hidden markov model increases the success rate of online recognition system. Online handwriting recognition gives additional time information which is not present in Offline system.

A Markov process is a random prediction process whose future behavior rely only on its present state, does not depend on the past state. Which means it should satisfy the Markov condition. A Hidden markov model (HMM) is a statistical markov model. In HMM model the system being modeled is assumed to be a markov process with hidden states. Hidden Markov models (HMMs) can be viewed as extensions of discrete-state Markov processes.

Human-machine interaction can be drastically getting improved as On-line handwriting recognition technology contains that capability. As instead of using keyboard any person can write anything by hand with the help of digital pen or any similar equipment would be more natural. HMM build a effective mathematical models for characterizing the variance both in time and signal space presented in speech signal.

Keywords*-Online Handwriting Recognition, Hidden markov model, Markov process*

*****

I. INTRODUCTION

Handwriting recognition classified by its input method into two types as off-line handwriting recognition and on-line handwriting recognition. For off-line recognition, the writing is usually captured using an input device like scanner. Where as in online recognition, a digitizer samples the handwriting to time-sequenced pixels as it is being written.

Hence, additional time information is present in the on-line handwriting signal which is not available in the off-line signal. On-line handwriting recognition contains potential for improving human-machine interaction. As instead of using keyboard any person can write anything by hand with the help of digital pen or any similar equipment would be more natural. HMM build a effective mathematical models for characterizing the variance both in time and signal space presented in speech signal.

Some on-line handwriting recognition systems have gives highly accurate recognition results by applying constraints such as printed small vocabulary size, writer-dependence, and writing style. These constraints can make on-line handwriting recognition considerably easier. Namely, a system performs cursive, large vocabulary, and writer-independent (WI) on-line handwriting recognition, as well as the ability to perform this task in real time.

Hidden Markov Models (HMMs) used with high success rate for statistical modeling of speech for last few years. recently HMM also been applied to on-line handwriting recognition with good success rate.

II. OBJECTIVES

The two main objective of this project is as follows:

- To effectively solve the difficulties in online handwriting recognition.
- To build effective mathematical models for characterizing the variance both in time and signal space presented in speech signal

III. LITERATURE REVIEW

This overview describes the nature of handwritten language, how it is transuded into electronic data, and the basic concepts behind written language recognition algorithms. Both the on-line case (which pertains to the availability of trajectory data during writing) and the off-line case (which pertains to scanned images) are considered. Algorithms for preprocessing, character and word recognition, and performance with practical systems are indicated. Other fields of application, like signature verification, writer authentication and handwriting learning tools are also considered.

Various methods have been proposed by researchers for preserving privacy and integrity in Handwriting recognition. Many Proposals are based on various methos like neural network, OCR, HMM methods.

In 2014, Najiba Tagougui, Houcine Boubaker, Monji Kherallah, Adel M.ALIMI [1] focused on hybrid NN/HMM

model for online Arabic handwriting recognition. The proposed system is based on Hidden Markov Models (HMMs) and Multi Layer Perceptron Neural Networks (MLPNNs)..

Y. Zhang, G. Shi, and J. Yang[2] in 2009 presents a recognition system based on Hidden Markov Model (HMM) for isolated online handwritten mathematical symbols. It give design of continuous left to right HMM for each symbol class and use four online local features, including a new feature: normalized distance to stroke edge.

In 2006 Marcus Liwicki and Horst Bunke [3] present an on-line recognition system for handwritten texts acquired from a whiteboard. This input modality has received relatively little attention in the handwriting recognition community in the past.

Marcus LIWICKI a and Horst BUNKE [4] in 2006 research on Notes written on a whiteboard is a new modality in handwriting recognition research that has received relatively little attention in the past.

Samuel Krasnik [5] in 2004, focused on a system which is able to recognize handwritten words from a lexicon in an on-line setting with high accuracy and is able to deal with highly deformed and highly ambiguous words.

Hiroshi tanaka, Naomi Iwayama and Katsuhiko Akiyama [6] in 2003 done research on online handwriting recognition and its applications. Which gives the recognizer as a supplemented with methods for processing the on-line data and generating the images. The recognizer is supplemented with methods for processing the on-line data and generating the images are the two technologies for character recognition.

Weide Chang [7] in 2003 focused on Improving Hidden Markov Models with a similarity Histogram for typing Pattern Biometrics.it improve the process with a histogram of similarity measured between actual observation and cluster centriod that it resembles.

Rejean plamondon and sargur N. srihari[8] in 2000 focused on Handwriting has contribution to persist as a means of communication and recording information in day to day life even with the new technique .in this paper describes the overview of handwritten lang. ,how it is transduced into electronic data.

Jianying Hu*, Sok Gek Lim, Michael K. Brown [9] in 1999 describe a Hidden Markov Model (HMM) based writer independent handwriting recognition system. A combination of signal normalization preprocessing and the use of invariant features make the system robust with respect to variability among different writers as well as different writing environments.

Han Shu[10] in 1996 introduces New global information-bearing features improved the modeling of individual letters, thus diminishing the error rate of an HMM-based on-line cursive handwriting recognition system. This system also demonstrated the ability to recognize on-line cursive handwriting in real time.

Jianying Hu, Michael K. Brown and William Turin [11] in 1996 done research in Hidden Markov Model (HMM) based recognition of handwriting. Which is now quite common, but the incorporation of HMM's into a complex stochastic language model for handwriting recognition is still in its infancy.

J. R. Bellegarda, D. Nahamoo, K. S. Nathan, and E. J. Bellegarda[12] in 1994 introduces Supervised Hidden Markov Modeling for On-Line Handwriting Recognition which gives the performance of a large alphabet handwriting recognition system based on a probabilistic framework is critically tied to the quality of the prototype distributions that are established in the relevant feature space.

IV. EXISTING SYSTEM

In day-to-day life security is major issue arises. New global information-bearing features improved the modeling of individual letters, thus diminishing the error rate of an on-line handwriting recognition system. This system also demonstrated the ability to recognize on-line handwriting. This recognition system is applied to on-line handwriting recognition. A handwriting recognition system is presented with a handwritten word, usually on a digitizer tablet with a pen device in the form of a sequence of spatial and temporal coordinates. In writer-independent handwriting recognition, the preprocessing is a very important part because each writer has its individual writing style.

The off-line preprocessing steps of (Marti and Bunke, 2001) are supplemented with additional on-line preprocessing operations to reduce the noise of the recording interface. The recorded on-line data usually contain noisy points and gaps within strokes. In the word await a spurious point occurs that leads to the introduction of a large artifact, i.e. two long additional strokes. Furthermore, we observe in the first line that there are many gaps within the word, which are caused by loss of data. In the last few years we have been experimenting with HMM-based methods for on-line handwriting recognition. For some characteristics, the results have been better than those achieved with speech recognition. For example, writer independence is achieved for handwriting recognition by a combination of preprocessing to remove much of the variation in handwriting due to varying personal styles and writing influences.

A. PROBLEM DEFINATION

The on-line handwriting signal contains additional time information which is not presented in the off-line signal. This addresses the problem of on-line handwriting recognition. To improve human-machine interaction On-line handwriting recognition technology is very beneficial.

B. PROBLEM SOLUTION

Hidden Markov Models (HMMs) used with high success rate for statistical modeling of speech for last few years. recently HMM also been applied to on-line handwriting recognition with good success rate.

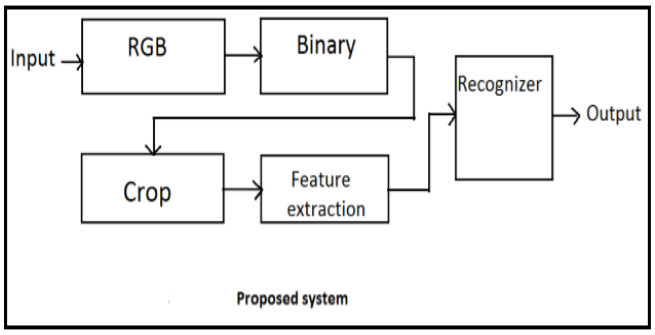
Now a day's various research organizations have concentrated on mathematical approaches, specially for on-line handwriting recognition hidden Markov models (HMMs) is considered. Success of HMMs' in the speech recognition domain has motivated the use of HMMs in the on-line handwriting domain. for characterizing the variance both in time and signal space presented in speech signals, HMMs is a effective statistical models .

C. SCOPE

1. To achieve a solution of complications introduced due to different writing styles of people.
2. To study the different technologies and methods for hand writing recognition.
3. To study the hidden markov model (HMM) for hand writing recognition.
4. For achieving the ultimate goal to mimic and extend the pen and paper approach.
5. Specially used in smart meeting rooms and class rooms.

V. SYSTEM ARCHITECTURE

A. Preprocessing:



The proposed scheme is to preserve the privacy and security of handwritten text. This scheme uses the GUI for creating user interface for easy access. The given selected input is sent to RGB. Then RGB convert original image into gray image.

Now present proposed system architecture including the preprocessing and normalization steps. The segmentation principle used is depicted and finally, we detailed the recognition system used based on HMMs. After cropping the edges image get proceed for Feature extraction.

The given input should be recognized using the system recognizer and to proceed that here used a histogram method. Histogram is a graphical representation of the data values. it converts the data into x-axis and y-axis to proceed it. The RGB used to transform input digit/character into gray image. Gray image is used because MATLAB can easily detect gray image.

The only real reasons for avoiding the obvious choice of Neural Nets are:

1) It has been done before for character recognition.

2) It does not seem to be in the same vein as defined class topics and lectures, and methods.

The general flow of statistic based character recognition algorithms is as follows:

- Compute the relevant statistics for a digitized image.
- Compare the statistics to those from a predefined database.

Suppose that the image of digit is already extracted from background and contain only real contours of digit.

a. A digitized image is, after all, just a collection of numbers.

b. A binary images, every point or pixel is assigned a value of either 0 or 1;

c. A gray level images pixel values range from 0 to 255,

d. A color images pixel values usually consist of three numbers, each in the range of 0 to 255.

B. SYSTEM FLOW

The flowchart is a means of visually presenting the flow of data through an information processing systems, the operations performed within the system and the sequence in which they are performed.

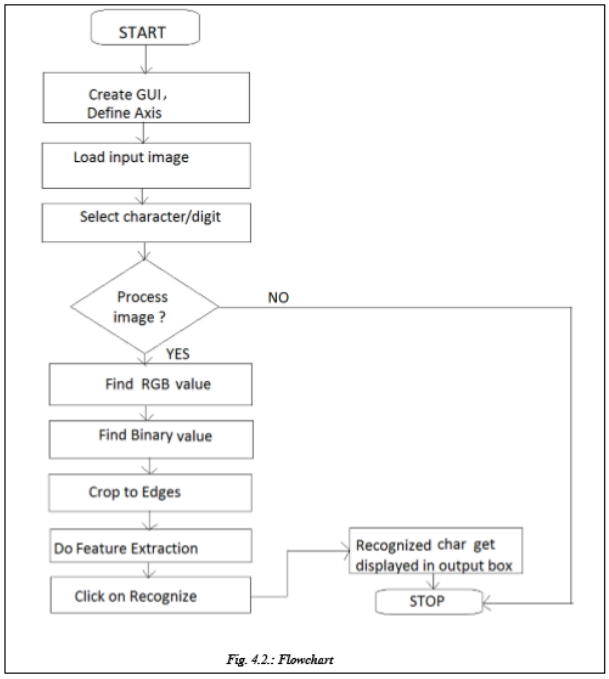


Fig. 4.2.: Flowchart

C. REQUIREMENTS SPECIFICATION

Requirements engineering provides the appropriate mechanism for understanding what the customer wants,

analyzing need, assessing feasibility, negotiating a reasonable solution, specifying the solution unambiguously, validating the specification, and managing the requirements as they are transformed into an operational system.

1. HARDWARE SPECIFICATIONS:
   Hard Disk Drive : min. 80 GB
   Ram : min. 1 GB
2. SOFTWARE SPECIFICATIONS:
   Operating System : Windows XP
   Front End : GUI (Graphical User Interface)
   Back End : MATLAB 7.12.0

VI. ADVANTAGES

1. OneNote for Android updated, adds tablet UI and handwriting recognition
2. Handwriting input lets you translate a written expression, even if you don't know how to type the characters.
3. Google enhanced its Translate service with handwriting recognition technology, allowing users to just scrawl the character on the touch screen.
4. On-line devices are that they capture the temporal or dynamic information of the writing.
5. On-line handwriting recognition means that the machine recognizes the writing while the user is writing. On-line handwriting recognition requires a transducer that captures the writing as it is written. The most common of these devices are electronic tablet or digitizer.

VII. FEATURES

1. Invariant features:

   The concept of invariant features arises frequently in various machine vision tasks. Depending on the specific task, the geometric transformation ranges from simple rigid plane motion to general affine transformation, to perspective mapping.

2. High-level features:

   In on-line handwriting recognition, some attempts have been made on approaches based on high-level features such as loops, crossings and cusps, where each handwritten word is converted into a sequence of primitives based on extraction of high-level features.

3. Segmental features:

   The purpose of introducing character level segmental features, which are features computed over the whole segment representing a character, is to capture shape characteristics of the whole character. This is desirable because, after all, characters are the basic shape units used to construct a word.

VIII. CONCLUSION

It provides privacy and integrity with security for information. The HMMs are able to model both the variance in time and signal space presented in cursively written letters, words, and sentences. Solutions to three basic problems of HMMs enabled us to perform various cursive handwriting recognition tasks: the solution to the training problem for training the HMM model parameters, the solution to the evaluation problem for handwriting recognition of isolated letters, and the solution to the decoding problem for handwriting recognition of sentences. In terms of security it significantly strengthens the security for handwritten text.

As already existing Digit recognition system also gives a good result but it is limited for numerical digits only. It gives around 70% efficiency. In terms of efficiency, the result of this project depends significantly on the input given. In general it gives 80% efficiency. The time delay depends upon the version of mat lab used.

ACKNOWLEDGMENT

This research was supported by Prof. Pankaj Vanwari who provides insight and expertise for the research.

I would like to express my heartfelt thanks to all the teachers and staff members of **Computer Engineering Department of ARMIET College** for their full support.

I would like to thank my Principal for conductive environment in the institution.

I am also grateful to the library staff of **ARMIET** for the numerous books, magazines made available for handy reference and use of internet facility.

I would also like to thank my parents for their constant help and support.

REFERENCES

[1] A Hybrid NN/HMM Modeling Technique for Online Arabic Handwriting Recognition Authors:Najiba Tagougui, Houcine Boubaker, Monji Kherallah, Adel M. ALIMI(Submitted on 2 Jan 2014)

[2] Y. Zhang, G. Shi, and J. Yang, "HMM-based online recognition of handwritten chemical symbols," In Proc. International Conference on Document Analysis and Recognition,pp. 1255–1259, 2009.

[3] Marcus Liwicki and Horst Bunke,"HMM-Based On-Line Recognition of Handwritten Whiteboard Notes" Institute of Computer Science and Applied Mathematics University of Bern, Neubr¨uckstrasse 10, CH-3012 Bern, Switzerland,2006.

[4] Marcus LIWICKI a and Horst BUNKE,"Handwriting Recognition of Whiteboard Notes" Department of Computer Science, University of Bern Neubr¨uckstrasse 10 CH-3012 Bern, Switzerland

[5] "On-line Handwritten Word Recognition" by Samuel Krasnik CS674 Natural Language Processing - Spring 2004.

[6] "online handwriting recognition and its applications" by Hiroshi tanaka, Naomi Iwayama and Katsuhiko Akiyama Manuscript received November 28,2003.

[7] "Improving Hidden Markov Models with a similarity Histogram for Typing Pattern Biometrics" by Weide Chang California state university,Sacramento2003

[8] "Online and offline handwriting recognition" by Rejean plamondon and sargur N. srihari member IEEE VOL.22,NO.1,January 2000.

[9] Writer independent on-line handwriting recognition using an HMM approach Jianying Hu*, Sok Gek Lim, Michael K. Brown Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Murray-Hill, NJ 07974, USA Received 26 August 1998; accepted 12 January 1999

[10] On-Line Handwriting Recognition Using Hidden Markov Models by Han Shu S. B., Electrical Engineering and Computer Science Massachusetts Institute of Technology (1996)

[11] HMM Based On-Line Handwriting Recognition by Jianying Hu, Member, IEEE, Michael K. Brown, Senior Member, I€€€, and William Turin, Senior Member, IEEE VOL. 18, NO. 10, OCTOBER 1996

[12] J. R. Bellegarda, D. Nahamoo, K. S. Nathan, and E. J. Bellegarda, "Supervised Hidden Markov Modeling for On-Line Handwriting Recognition," presented at International Conference on Acoustics, Speech, and Signal Processing, 1994.