Scope of Handwriting Recognition in Indic Scripts

Sukhdeep Singh[1]

[1]Assistant Professor in Computer Science, Mata Sahib Kaur Girls College, Talwandi Bhai, Ferozepur, Punjab, India.

ABSTRACT: T Handwriting recognition is the process of understanding handwritten text by a computer or other computing devices. The handwriting recognition system comprises various phases like data collection, preprocessing of data, segmentation of long strokes, feature extraction, classification and post-processing etc. Handwriting recognition has a variety of applications as electronic form filling, signature verification, automated music symbol notation reader, handwritten ancient document reading, write and send SMS in mother tongue and alternative to realistic keyboards etc. The present study explains the scope and applications of handwriting recognition in Indic scripts.

Key Words: Indic, offline handwriting recognition, online handwriting recognition

Introduction

It is the era of technology and technological advancements are being made at a rapid rate. As the technology is advancing, the ways of communicating between human beings and computers are also changing. The speech and handwriting recognition are among novel methods of exchanging ideas or information between computers and its users. Handwriting recognition can be both offline and online. When text is first written on the paper and then fed into a computer, it is called offline handwriting recognition. But on the other hand, when the computer identifies the handwritten text while writing on the screen of a digital device, it is called online handwriting recognition. The progress of handwriting recognition work can be observed from premier journals/conference of document analysis and pattern recognition as PAMI (Pattern Analysis and Machine Intelligence), PR (Pattern Recognition), PRL (Pattern Recognition Letters), IJDAR (International Journal on Document Analysis and Recognition), IWFHR (International Workshop on Frontiers in Handwriting Recognition), ICFHR (International Conference on Frontiers in Handwriting Recognition) and ICDAR (International Conference on Document Analysis and Recognition) etc.

Even though a great research work has been carried out for offline and online handwriting recognition in last decades. However, in contrast to Latin and CJK (Chinese, Japanese, Korean) scripts, the available studies or work for Indic scripts is very limited [1],[2],[3],[4],[5]. The Indic scripts are also called the Brahmi scripts. The people living around the Indus river use Brahmi or Indic scripts. Although, the mainresearch work for handwriting recognition in Indic scripts is done in last two decades, but the major work done in this period is for smaller units as characters or strokes. The work done for Indic scripts' handwriting recognition for larger units as words or sentences is very limited.The Gurmukhi, Devanagari, Tamil, Bangla, Kannada, Malayalam, Oriya, Telugu and Gujarati are the major Indic scripts. These all are not isolated scripts. So the work done for one script will also be useful for other Indic scripts. For example the presence of headline is common for Devanagari and Gurmukhi script. The major Indic scripts' online handwriting recognition face the challenge of dealing stroke and symbol order variations. Thus, there is great scope for carrying out handwriting recognition study for a class of scripts instead of individual scripts.

Nowadays, online and offline handwriting recognition of Brahmi scripts is getting the attention of researchers in twenty first century. The research in this area has received additional drive because of the latest planned financial assistance by the Government of India in the direction of technological development of Indian scripts and languages [6]. The standard and benchmarked datasets of online and offline handwritten characters of some Indic scripts have become available, it will further give momentum to Indic scripts' handwriting recognition research. These contain freely available character datasets of online handwriting of Bangla, Devanagari, Tamil and Telugu scripts. These benchmarked datasets of Indic scripts have been evaluated on point float and direction code histogram based features where Nearest Neighbour (NN), Multilayer Perceptron (MLP), Hidden Markov Model (HMM) and Support Vector Machines (SVM) etc. classifiers have been used for validation of these datasets [7], [8], [9], [10].

The present study has presented the major scope of handwriting recognition work of Indic scripts especially. So this study will also motivate the future readers to work in the area of handwriting recognition. Handwriting recognition has a variety of applications as electronic form filling, signature verification, automated music symbol notation reader, handwritten ancient document reading, write and send SMS in mother tongue and alternative to realistic keyboards etc. The present study explains the scope and applications of handwriting recognition in Indic scripts.

Scope of Handwriting Recognition

a) Electronic form filling
One of the applications of online handwriting recognition is electronic form filling.Internationally, the expenditure for entry of data from handwritten forms, notes and records is trillions of dollars. If we look at 2010 census of our country, more than fifty thousand enumerators were employed to collect data using handwritten forms, where they took six months to do this job. Further, it took two more years to feed this data into servers. So, in such applications, an immediate and direct conversion of handwritten data to typed data will result in reducing the huge cost and it will also increase the productivity. In this way, all government application forms can be completed and filled using handwriting recognition and the data will be directly entered to structured databases. It will be only possible, when the handwriting recognition is standardized, perfected and it is available to all computing devices. Further the writer dependent and independent systems could be developed as per the requirement.

b) Writing electronic applications in one's own handwriting and nativescript
There are a number of native language/script speakers/writers who want to exchange information with the computer system. These writers know their native script only, but don't know typing in their script. One of the problems of these writers is to write electronic applications. In such situations, it becomes difficult for native script writers to communicate with computing devices. One of the options for such writers is speech recognition. But speech recognition has certain limitations/problems. In such scenarios, writing electronic applicationsin one's own handwriting and native script is the best solution. It can be supported using online handwriting recognition.

c) Automated music symbol notation reader
One of the applications of handwriting recognition is the development of automated music symbol notation reader. In this way, a composer can write his composition using all the notations directly. Then it is converted to the standard format to display/print for a book or for his symphony group. Further, the Vedic Sanskrit is identical,with the symbols for *udata, anudata, svaritha, deerghasvaritha and plutha*, and the engine which is capable of recognizing handwritten*Grantha* or Devanagari with all such symbolswill be in great use.

d) Alternative to hardware and software keyboards
Online handwriting recognition is also an alternative to keyboards. In this way, an input to a computing device can be made using handwriting recognition. Although QWERTY keyboards are commonly used for all small alphabet sets of European languages, it is also a mapped for other languages. But when the character set becomes very large for Chinese, Japanese and Indic scripts, it becomes very difficult use the QWERTY interface. One the option in such scenarios is the software keyboards, but online handwriting is the most suitable solution. Using online handwriting recognition, the users also feel the natural interface to input data. All keyboards have finite number of keys, so it limits the number of symbols that can be entered from the keyboard.But online handwriting avoids such limitations and permits input of symbols (without limits), sketches and drawings. It facilitates writers to think freely while inputting the data.

e) Putting in the mathematical equations by simple handwriting
The handwriting recognition is assumed to achieve the goal of automatic conversion of online handwritten mathematical equations to typed form or student notes. So it will not be wrong to imagine the happiness that would be experienced by a chemist, mathematician or an academician, if they can place the mathematical equations by simple online handwriting on the digital surface and make his/her book, plenary talk and journal papers.

f) Writing and sending SMS in his/her mother tongue
Some native script users want to send SMS in their mother tongue handwriting and this can be achieved with mature and developed handwriting recognition technology. Thus there is the possibility that the human beings write and send SMS in their mother tongue. To make it possible,

the State and Union Governments must force the mobile manufacturing and service provider companies to facilitate and support the Unicode fonts to display and transmit of Brahmi scripts.

g) Biometrics and forensics

The handwriting recognition is highly applicable in writer identification, where it is used in forensics and biometrics. So one of the applications of handwriting recognition is solving the ancient manuscript disputes, where the actual writer of the manuscript is identified based on certain features of writers' handwriting. In this way, it assists to avoid false claims of handwriting or manuscript.

h) Digitization of palm leaf manuscripts

As the offline handwriting recognition deals with understanding the handwritten text which is already written on papers or other writable surfaces or materials. So one of the applications of offline handwriting recognition is the digitization of ancient palm leaf manuscripts. This problem is not solved yet. It is not even solved for Latin script completely.

i) Automatic conversion of prescription to typed form

It is one of the challenging tasks to understand the doctor's handwriting. This problem can be solved by using offline and online handwriting. In case of automatic conversion of prescription to typed form, online handwriting will be employed.

Concluding Remarks

In the present study, we have given an overview of the scope of offline and online handwriting recognition in context of Indic scripts especially. Though, a lot of work has been done for Chinese, Japanese, Korean and Latin scripts, but there is great need to work in the direction of handwriting recognition in Indic scripts. Yet, the progress and expansion of Indian language technologies such as online and offline handwritingrecognition has a long way to go. So, in order to survival of Indic scripts in the technological and internet era, a great work is to be done yet. The present work is a motivation for future readers and researchers to work and carrying out research in the area of online and offline handwriting recognition in Indic scripts.

References

1. S. Jaeger, C. L. Liu, M. Nakagawa, "The state of art in Japanese online handwriting recognition compared to techniques in Western handwriting recognition", International Journal on Document Analysis and Recognition, vol. 6, pp. 75-88, 2003.
2. N. Arica, F. T. Yarman-Vural, "An overview of character recognition focused on off-line handwriting", IEEE Trans. Systems Man and Cybernetics-Part C : Applications and Rev. , vol. 31, no. 2, pp. 216-233, 2001.
3. R. Plamondon, S. N. Srihari, "On-line and offline handwriting recognition: a comprehensive sur-vey", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63-84, 2000.
4. C. C. Tappert, C. Y. Suen, T. Wakahara, "The state of art in on-line handwriting recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 12, no. 8, pp. 787-800, 1990.
5. C. C. Tappert, C. Y. Suen, T. Wakahara, "On-Line handwriting recognition: a survey", Proc. of ICPR, vol. 2, pp. 1123-1132, 1988.
6. T. Mondal, U. Bhattacharya, S. K. Parui, K. Das and D. Mandalapu, "On-line Handwriting Recognition of Indian Scripts - The First Benchmark", 12thInternational Conference on Frontiers in Handwriting Recognition, Kolkata, 2010, pp. 200-205.
7. N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression",The American Statistician,46 (3), 1992, 175–185.
8. T. Mondal, U. Bhattacharya, S. K. Parui, and K. Das, "On-line handwriting recognition of Indian scripts - the first benchmark",In: Proceeding of the 12thInternational Conference on Frontiers in Handwriting Recognition,2010, 200–205.
9. L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", Proceedings of IEEE,77, 1989, 257–286.
10. C. Cortes and V. Vapnik,"Support-vector networks", Machine Learning,20 (3), 1995, 273–297.