International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161
©2015 INPRESSCO®, All Rights Reserved
Available at http://inpressco.com/category/ijcet

Review Article

A Review Paper on Malayalam Text to Braille Transliteration

Salah C.†* and A. Ranjith Ram†

†Advanced Communication & Signal Processing Laboratory, Department of Electronics & Communication Engineering, Government College of Engineering Kannur,Kannur, Kerala 670 563, India

Accepted 30 June 2015, Available online 02 July 2015, **Vol.5, No.4 (Aug 2015)**

Abstract

Braille is a tactic writing system that has been used by blind. This paper presents a brief survey of character recognition and transliteration works on Braille system so as to design a system that will transliterate Malayalam scanned books and magazines in to Braille. The proposed system deals with a printed Malayalam text to Braille transliteration system which includes a Malayalam optical character recognition system (OCR) and Malayalam to Braille mapping.

Keywords: *Braille system, Bharathi Braille script, Malayalam Braille,Optical character Recognition, Support vector machine*

1. Introduction

Visually impaired people are an integral part of the society. In this era of technology, the knowledge resources are at the finger tips, but in the world of blind people, they have only two sources of knowledge – audio and Braille script. Braille is a tactile writing system. Initially the Braille was developed by Charles Barbier for soldiers. Later Louis Braille modifies the Charles Barbiers method and developed today's Braille system. Each Braille character made up of 6 dot positions, arranged in a rectangle containing 2 columns of 3 dots each called a Braille cell. A dot may be raised at any of the 6 position to form total 64 permutations. Positions being universally numbered 1 to 3 from top to bottom, on left and 4 to 6 from top to bottom on right as shown in fig 1. Since Braille became one of the most important ways for the blind to learn and obtain information, transliterating normal text into Braille became a necessity. However, manual transliteration is time consuming and prone to have errors; hence systems to perform automatic transliteration have been conceived. Out of the 37 million blind people across the globe, over 15 million are from India. But most of the available knowledge resources for blind persons in Braille script are in English, Chinese and Arabic and are not available in Indian languages. The proposed system deals with conversion of printed Malayalam text into Braille. The main part of the system is a Malayalam OCR system.

Fig.1 A 6-dot Braille cell with dot position numbered universally

*Corresponding author: **Salah C**

Braille for Indian language is called Bharathi Braille. The beauty of Bharathi Braille is that it is based on phonetics. Bharati Braille is the adaptation of the six dot system for the languages of India. The history of Bharati Braille dates back to the period prior to India's independence. At a conference held at Beirut in 1951, a body of world scholars examined the possibility of a phonetically derived system of six dots that could be used for most of the languages of India, pakistan and Sri Lanka. Bharati Braille was a result of this exercise. IIT Madras has worked a lot in the field of Bharathi Braille under the guidance of Professor Kalyana Krishnan. Bharathi Braille follows grade-1 Braille in English that is Bharati Braille assigns the cells to the basic sounds of the Indian languages (these are called aksharas) in a manner where vowels and consonants that find direct equivalents in English are given the same representation as in English. Bharati Braille retains all the basic conventions relating to the representation of numerals, punctuation and special symbols just as in Standard English Braille [Manzeet Singh *et.al*., 2010]. Syllable is always represented through its consonants and vowels by explicitly writing consonants and vowels one after the other. Ligatures (conjuncts and chillu) in print are handled with viarama followed by consonants. Malayalam Braille is one of the Bharati Braille alphabets, and it largely conforms to the letter values of the other Bharati alphabets. Malayalam Braille alphabets are shown in fig 2.

The rest of the paper is organized as follows. In Section II, we present the current state of the art related to Braille transliteration. In Section III, we discuss the various approaches of Malayalam OCR. In Section IV we discuss the proposed system and in Section we discuss its applications. Finally Section VI summarizes this paper with some concluding remarks.

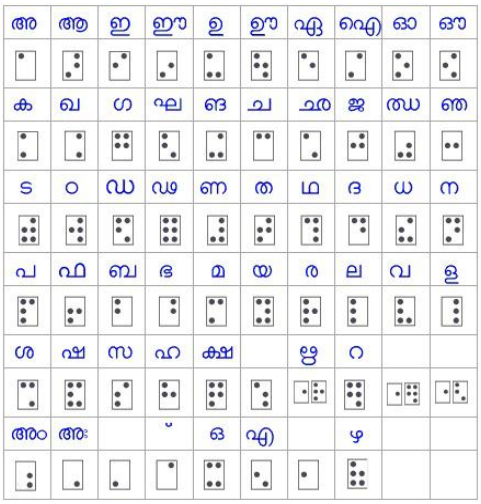

Fig.2 Malayalam Braille alphabets

2. Current state of the art

The automatic text to Braille transliteration is not a new research field. It started as early as mid-1960s. The transliteration process can be done at software level or at hardware level. Most researches and products are at software level. Duxbury Braille Translation (DBT) Software is a Window based software that automates the process of conversion from regular print to Braille and vice versa. The system has only an English interface [AbdulMalik S *et.al*.,2008]. Duxbury Braille Board (DBB) is another window-based software system It can translate from English text to Braille system. The user can write directly in Braille using the normal keyboard. When the user presses a letter on the keyboard, its corresponding Braille cell will be displayed on screen. The system supports only English language. WinBraille (Index Braille, 2007), Braille maker are also good Window-based software system for braille transliteration. Most of these systems for text to Braille transliterations are based on English.

P. Blenkorn [P. Blenkorn *et.al*, 2001] have worked on the problem of converting Word-Processed Documents into formatted Braille document. The problem has been addressed in context to the users of the word processor want to produced Braille document. The translator will be integrated with the word processor. It makes easy to use as translation is possible with the help of menu item in MS Word. They have use DLL written in C, to translate Text to Braille. Braille Out system produces reliable contracted Braille from a wide variety of documents. The layout is good and speed is acceptable.

M. Y. Hassan [M. Y. Hassan *et.al*.,2011] presented a paper on conversion of English characters in to Braille using Neural network. A feed forward artificial neural network was designed and tested to convert English characters into grade I literary Braille code.

Vandana [Vandana *et.al*.,2012] proposed a method for implementing Gurmukhi to Braille. The input Gurumukhi is applied by direct inputting. For the system single characters gives 98% accuracy, numerals gives 100%, words gives 95% and vowels gives 96% accuracy.

Manzeet Singh [Manzeet Singh *et.al*.,2010] presented a paper on Automated Conversion of English and Hindi Text to Braille Representation. For the transliteration they used a look-up table. When any input Hindi or English text is entered that text is first broken in to corresponding letters and that are matched against the look-up table. If any match is found corresponding Braille cell was displayed.

Xuan Zhang [Xuan Zhang *et.al*.,2006] have addressed the problem of translating text to Braille based on FPGA in 2006. To achieve fast translation FPGA with a big programmable resource has been utilized, and an algorithm, proposed by P. Blenkhorn [P. Blenkorn *et.al*, 2001], has been revised to perform fast translation. B. L. Shivakumar [B. L. Shivakumar *et.al*.,2013] have addressed the problem of converting English text to Braille in 2013. They have used one to one matching technique. They have used Visual Basic 6.0 as front end tool and MS Access 2002 as back end tool. So they have considered as a Client/Server Architecture. Dasgupta [Dasgupta *et.al.,*2012] have presented an automatic Dzongkha text to Braille forward transliteration system. Dzongkha is the national language of Bhutan. The system is aimed at providing low cost efficient access mechanisms for blind people. It also addresses the problem of scarcity of having automatic Braille transliteration systems in language slime Dzongkha. The present system can be configured to take Dzongkha text document as input and based on some transliteration rules it generates the corresponding Braille output.

Bindu Philip [Bindu Philip *et.al*.,2009] proposed a system of Smart OCR for the Visually Challenged. The paper addresses the integration of a complete Malayalam Text Read-out system for the visually challenged. The system was developed using C++ on Windows XP platform.

3. Approaches of Malayalam OCR

Optical Character Recognition is a process that can convert text, present in digital image, to editable text [Sukhpreet Singh *et.al*.,2013]. It allows a machine to recognize characters through optical mechanisms. The output of the OCR should ideally be same as input in formatting. The process involves some pre-processing of the image file and then acquisition of important knowledge about written text. Various approaches used for the design of OCR systems are Matrix Matching which converts each character into a pattern within a matrix, and then compares the pattern with an index of known characters, Fuzzy logic which is a multi-valued logic that allows intermediate values to be defined between conventional evaluations like yes/no, true/false, black/ white etc., Feature Extraction in which each character by the presence or absence of key features, including height, width, density, loops,

lines, stems and other character traits, Structural Analysis which identifies characters by examining their sub features shape of the image, sub-vertical and horizontal histograms and Neural Networks which simulates the way the human neural system works. It samples the pixels in each image and matches them to a known index of character pixel patterns.

Malayalam character recognition is a complex task since it consists of large number of character set and high similarity between characters. Reasonable amount of Malayalam OCR works is done in the last decades. Renju john [Renju john *et.al*.,2007] presented a paper on 1D wavelet transform of projection profile for hand written Malayalam characters. Bindu Philip [Bindu Philip *et.al*.,2009] uses a support vector machine (SVM) classifier for Malayalam character recognition for the use of Braille transliteration of Malayalam printed text. M. S. Rajasree. [Rajasree *et.al*.,2009] proposed a printed Malayalam OCR using back propagation neural network. Jomy John [Jomy John *et.al*.,2011] proposed a Malayalam OCR system using wavelet transform and SVM classifier.

Recently support vector machine based OCR has received much attention. The SVM is a new type of classifier based on novel statistical learning techniques. The SVM classifier in its basic form implements two-class classifications. The principle of an SVM is to map the input data onto a higher dimensional feature space nonlinearly related to the input space and determine a separating hyperplane with maximum margin between the two classes in the feature space [Bindu Philip *et.al*.,2009]. SVM can classify only a fixed length data vectors. It is a generalized linear classifier with maximum-margin fitting function. The fitting function provides regularization which helps the classifier generalized better.

4. Proposed system

The proposed system will be an SVM based printed Malayalam text to braille transliteration system. The block schematic of proposed system is shown in fig 3. The main part of the system is a Malayalam OCR system. The system first segment each Malayalam characters. Some features from each of the character are then extracted. SVM is used to classify and to identify the characters. SVM constructs a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class, since in general the larger the margin the lower the generalization error of the classifier. The fundamental idea of SVM classifier is to find a seperating hyper plane between two classes so that there is a minimal distance with respect to the training vectors. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Malayalam OCR system has large number of classes and high dimensional feature vectors. Variability of character is also high at each occurrence. SVM are well suited in such problems. SVM is trained with the set of feature. Once training is over, the SVM is used to classify new sets of characters. The second part of the system is a Braille transliterator which finds the corresponding Braille cell for each recognized characters.

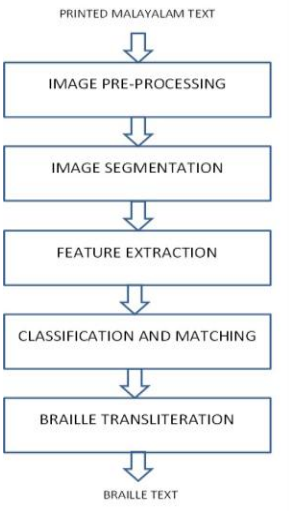

Fig.3 Block diagram of Malayalam to braille transliteration system

5. Applications

The printed Malayalam text to Braille transliteration system finds interesting applications in libraries, offices where instructions and notices are to be read and also in assisted filling of application forms. There are very limited knowledge resources for blind persons available in Braille script, they can only get their academic books. Novels and news sources are hardly available in Braille. With this system along with a Braille embosser we can provide large number of Malayalam novels and literature in Braille. By this system it is possible that all the study material whether that is an e-book, scanned papers or other available to the blind people who can understand the Braille lipi. As this is a small attempt to enhance the functionality of the Braille lipi. Further a lot of work progress can be made on the way to make more study resources available to blind people. More material will be available to them, for studying. The system will be used for reproducing most textbooks and publications.

Conclusions

As in this technological era, our little efforts can help us to do something for the blind people. Their disabilities have made them to have less access to computers, study resources and high quality educational software than the people with clear vision. The Braille system is one of the most used systems for non-visual communications as it is based on the touching sense, through which the visually impaired people can explore the world of knowledge. The paper describes that reasonable amount of work is done in various languages to Braille conversion. But still more and accurate work is required in Malayalam text to Braille conversion. The printed Malayalam text to braille transliteration system using SVM is a fast, low cost and accurate solution. With these efforts they can see the world from letter they read. This will help them to learn new things.

References

Bindu Philip and R. D. Sudhaker Samuel (2009), Human Machine Interface A Smart OCR for the Visually Challenged*, International Journal of Recent Trends in Engineering*, 2(3)

Bindu Philip and R. D. Sudhaker Samuel (2009), Preferred Computational Approaches for the Recognition of different Classes of Printed Malayalam Characters using Hierarchical SVM Classifiers, *International Journal of Computer Applications*, 1(16

Bindu Philip and R. D. Sudhaker Samuel(2009), A Novel Bilingual OCR System based on Column- Stochastic Features and SVM Classifier for the Specially Enabled, *Second International Conference on EmergingTrends in Engineering and Technology, ICETET*,

AbdulMalik S. and Al-Salman(2008), A Bi-directional Bi-Lingual Translation Braille-Text System, *J. King Saud University, Computer and Information Science*, Vol. 20

S.Padmavathi,Manojna K. S. S, Sphoorthy Reddy .S and Meenakshy.D(2013), Conversion of Braille to text in English, Hindi and Tamil languages ,*International Journal of Computer Science, Engineering and Applications (IJCSEA)*, 3

Vandana, Nidhi Bhalla and Rupinderdeep Kaur(2012), Architecture of Gurmukhi to Braille conversion system, *International Journal of Computer Science and Information Technology and Security (IJCSITS)*, 2

Mohammed Y. Hassan and Ahmed G. Mohammed (2011), Conversion of English characters into Braille Using neural network , *IJCCCE*, 11

Xuan Zhang, Cesar Ortega-Sanchez, and Iain Murray (2006), Text-to-Braille translator in a chip, *4th International Conference on Electrical and Computer Engineering ICECE*.

Vandana, Rupinderdeep Kaur and Nidhi Bhalla(2012), Implementation of

Gurmukhi to Braille, *International Journal of Computer Science And Technology,* 3(2)

Manzeet Singh and Parteek Bhatia (2010), Automated conversion of English and Hindi text to Braille representation, *International Journal of Computer Applications,* 4(6)

Manzeet Singh and Parteek Bhatia (2010), enabling the disabled with translation of source text to Braille, *proceedings of national conference on advanced computing, communication engineering and technology*.

Xuan Zhang, Cesar Ortega-Sanchez and Iain Murray (2006), Hardware-Based Text-To-Braille Translator, *4th International Conference on Electrical and Computer Engineering ICECE*

Mukul Bandodkar and Virat Chourasia(2014), Low Cost Real-Time Communication Braille Hand-Glove for Visually Impaired Using Slot Sensors and Vibration Motors, *International Journal of Electrical, Robotics, Electronics and Communications Engineering*, 8(6)

Sukhpreet Singh(2013), Optical Character Recognition Techniques: A Survey,*4th International journal of Emerging Trends in Computing and Information Sciences*, 4(6)

TirthankarDasgupta, Manjira Sinha and Anupam Basu (2012), Forward Transliteration of Dzongkha Text to Braille, *Proceedings of the Second Workshop on Advances in Text Input Methods (WTIM 2)*, pp. 97–10

Al-Salman AbdulMali, AlOhali Yosef, AlKanhal Mohammed and Al- Rajih Abdullah(2007), An Arabic Optical Braille Recognition System, *Proceedings of the First International Conference in Information and Communication Technology*, pp. 81–

M.Rahmath Riswana, Rani Thottungal and V.Kandasamy(2013), Multi Character Identification for Visually Impaired Using Braille System,*Proceedings of International Journal on Recent and Innovation Trends in Computing and Communication*, 1(3), pp. 186–191.

R. John, G. Raju and D. Guru(2007), 1D wavelet transform of projection profile for isolated handwritten Malayalam character recognition, *Conference on Computational Intelligence and Multimedia Applications*, 2 pp. 481–485,

M. Rahiman and M. Rajasree(2009), Printed malayalam character recognition using back-propagation neural networks, *Advance Computing Conference, IEEE International*, pp. 197–201

J. John, K. Pramod, and K. Balakrishnan(2012), Unconstrained handwritten malayalam character recognition using wavelet transform and support vector machine classifier, *International Conference on Communication Technology and System Design*, 30, pp. 598–6

B. L. Shivakumar and M. R. Thipathi(2013),English to Braille Conversion Tool using Client Server Architecture Model,*International Journal of Advanced Research in Computer Science and Software Engineering*, 3(8), pp. 295–299.

P. Blenkhorn and G. Evans(2001),Automated Braille Production from Word-Processed Documents, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*9(1), pp.81–85.

Bharati Braille URL: http: //acharya.iitm.ac.in\disabilities \bhbrl.php.

Acharya-Multilingual Computing for Literacy and Education [Online] Available www:acharya. gen.in\ Introductionto Braille.htm.