International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 05, May-2015

Issues in Transliterating Malayalam Text to Braille

Salah C.
Department of Electronics & Communication Engineering,
Government College of Engineering Kannur,
Kannur,Kerala,India

A. Ranjith Ram
Department of Electronics & Communication Engineering,
Government College of Engineering Kannur,
Kannur,Kerala,India

***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 and some issues in transliterating Malayalam text to Braille 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. The system includes a Malayalam optical character recognition system (OCR) and a Malayalam to Braille mapping system.**

Keywords-Braille system, Bharathi Braille script, Optical character Recognition, Support vector machine

I. 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 soliders. Later Louis Braille modifies the Charles barbies 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 [23]. 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.

Braille for Indian language is called Bharathi Braille [23],[24]. 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 [23]. Bharathi Braille follows grade-1 Braille in English.

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 issues in transliterating Malayalam to Braille. In Section IV we propose the work that will transliterate printed Malayalam text to Braille. Section V we discuss its applications. Finally Section VI summarizes this paper with some concluding remarks.

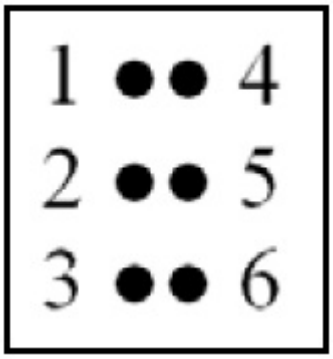


Figure 1. A 6-dot Braille cell with dot position numbered universally

II. 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 [4]. 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), Braillemaker 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 [22] 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

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

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.

In 2011, M. Y. Hassan et. al. [7] presented a paper on conversion of English characters in to Braiolle using Neural network. A feed forward artificial neural network was designed and tested to convert English characters into grade I literary Braille code.

In 2012, Vandana et. al. [6], [9] 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.

In 2010, Manzeet Singh et. al. [10] 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 et. al. [12], [8] 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 [22], has been revised to perform fast translation. B. L. Shivakumar et. al. [21] 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 et. al. [15] 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.

In 2009, Bindu Philip et. al. [1], [2] 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.

III. ISSUES IN TRANSLITERATING MALAYALAM TEXT TO BRAILLE

The six dot system is viewed as another script to write the languages of India. The compromise effected is that a syllable will always be represented through its consonants and vowels by explicitly writing the consonants and vowels one after the other. Ligatures are therefore eliminated. True, this would mean that more cells would be required but considering the fact that one gets uniformity and grade-1 Braille in English also follows this convention, the system is indeed viable.

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 is thus a system for writing syllables using a basic set of 63 shapes, each corresponding to a cell. Here, the most basic approach to writing syllables using generic consonants has been used. A syllable in Indian languages can take a pure vowel (V) form or a basic consonant and a vowel (CV) form or a conjunct with two or more consonants and a vowel (CC...V) form. The cell assignment for a consonant assumes that the consonant has an implied vowel a as part of it. A pure consonant (also known as a generic consonant) has no vowel and so to distinguish a basic consonant from its generic equivalent a special symbol is used in the writing systems. This is known as the halanth and its shape is a language specific ligature added to the shape of the basic consonant. Bharati Braille has set apart one cell for this purpose and this cell placed before the cell for a basic consonant turns it into a generic consonant. The idea here is that one can use this principle to write syllables in the CC...V form simply by concatenating the cells for each generic consonant [2]. Fig 2 shows Bharathi Braille examples which follows all the above rules.

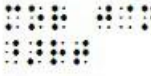
Example	Braille
ಕನ್ನಡವೇ ಸತ್ಯ	
मेरा भारत महान	
தமிழ் மொழி இனியது	
తెలుగు మాట తీపు	
കേരളം മനോഹരമാണ്	
અમારો દેશ ભારત છે	
ਭਾਰਤ ਮਾਡਾ ਦੇਸ਼	
ভারত আমার দেশ	
ମୋର ଜନ୍ମଭୂମି ଭାରତ	

Figure 2. Bharathi Braille examples

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 show in Fig 3. The biggest issue in Braille conversion is that the database is limited in present softwares. Due to which, it is hard to use for these broad number of applications. So, it is primarily required to work on the size of the database. Neither of the Braille conversion software work with scanned document. Software should be a system where we can give the input document as the scanned copy and can get the output as Braille document. Malayalam character recognition is a complex task since it consists of large number of character set and high similarity between characters.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

IV. PROPOSED SYSTEM

The proposed system takes a scanned document image and produces text documents containing the extracted Malayalam characters and their corresponding Braille cells. The system first segment each Malayalam characters. Some features from each of the character are then extracted. SVMis used to classify and to identify the characters. The block schematic of proposed system is shown in fig 4.

The scanned image is first converted into grayscale. Gaussian and/or salt and pepper noise are commonly present in scanned document images. To remove the effect of noise filtering is performed on the image. Global thresholding approach is used for binarizing the filtered image. Skeletonization is performed to remove extra pixels from the character patterns in the binarized image. The scanned document image is first segmented into lines which are further segmented into words and ultimately into characters and sub characters.

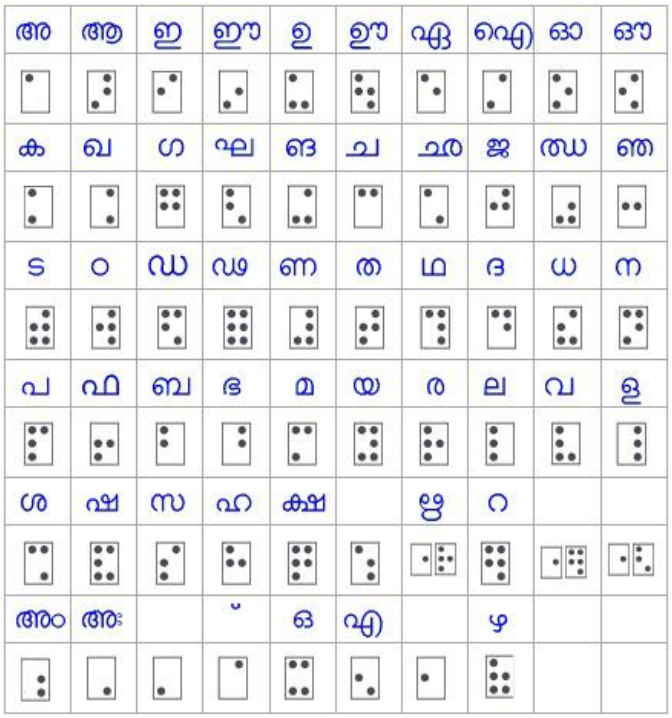


Figure 3. Bharathi Braille examples

A representative set of features is extracted from the normalized character to form a feature vector of dimension 20. The classifier used is Support Vector Machine (SVM). SVM is one of the most popular tools for pattern recognition It is suited for both linear as well as non-linear classification problems. The Support Vector Machine (SVM) classifier in its basic form implements two-class classifications. The objective is to further improve the recognition rate by using support vector machine (SVM) at the segment classification level. The advantage of SVM, is that it takes into account both experimental data and structural behavior for better generalization capability based on the principle of structural risk minimization (SRM). Its formulation approximates SRM principle by maximizing the margin of class separation, the reason for it to be known also as large margin classifier.

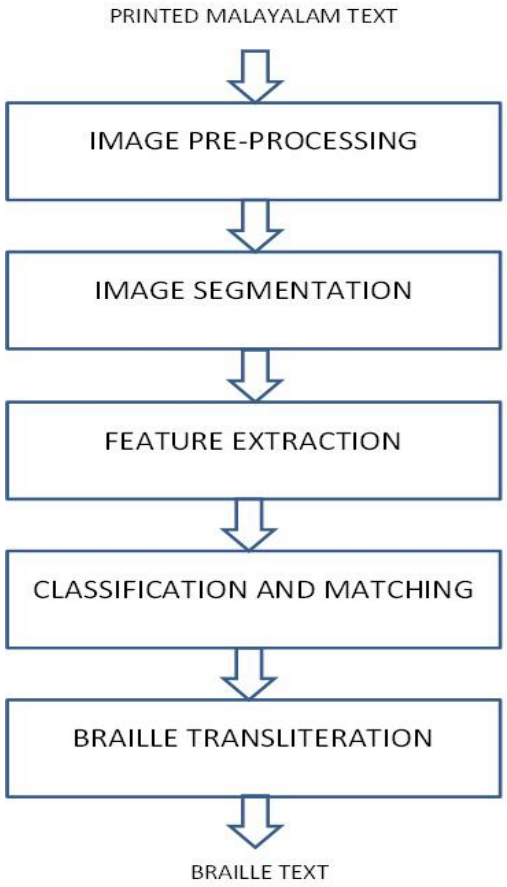


Figure 4. Block diagram of Malayalam to braille transliteration system

The fundamental idea of SVM classifiers is to find a separating hyperplane between two classes (h : wT x + b = 0), so that minimal distance with respect to the training vectors, called margins, is maximum. The optimal solution is obtained when this hyperplane is located in the middle of the distance between the convex envelopes of the two classes. This distance is denoted by d_m and is expressed by,

$$d_m = \frac{2}{\|w\|}$$

The support vectors are situated on the margins of the two classes. If the training vectors membership is defined by

$$u_k = 1 \text{ if } x_k \in w_1$$
$$u_k = -1 \text{ if } x_k \in w_2$$

Then the support vectors can be written in the form generate non-linear separating surfaces. The basic idea is to project the input vectors in higher dimension space where the classes become linearly separable. Fig 5 demonstrates linear classification using SVM.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

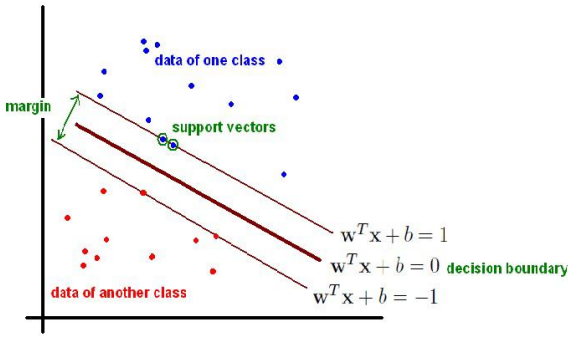


Figure 5. SVM : Linear Classification Example

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 transliteration system which finds the corresponding Braille cell for each recognized characters.

V. 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.

V. CONCLUSION

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. There are some issues like Limited database, Scanning problem, etc. when transliterating Malayalam character to Braille. However Printed Malayalam text to Braille transliteration system is a fast, low cost and accurate system for transliterating scanned Malayalam document to Braille. A promising feature of this work is that the system was able to successfully classify and transliterate as many as 102 Malayalam characters.

REFERENCES

[1] Bindu Philip and R. D. Sudhaker Samuel, "Human Machine Interface A Smart OCR for the Visually Challenged," International Journal of Recent Trends in Engineering, Vol. 2, No. 3, November 2009.

[2] Bindu Philip and R. D. Sudhaker Samuel, "Preferred Computational Approaches for the Recognition of different Classes of Printed Malayalam Characters using Hierarchical SVM Classifiers," International Journal of Computer Applications,Vol. 1, No. 16, 2009.

[3] Bindu Philip and R. D. Sudhaker Samuel, "A Novel Bilingual OCR System based on Column- Stochastic Features and SVM Classifier for the Specially Enabled," Second International Conference on Emerging Trends in Engineering and Technology, ICETET, 2009.

[4] AbdulMalik S. and Al-Salman, "A Bi-directional Bi-Lingual Translation Braille-Text System," J. King Saud University, Computer and Information Science,, Vol. 20 pp. 13-29, 2008.

[5] S.Padmavathi,Manojna K. S. S, Sphoorthy Reddy .S and Meenakshy.D, "Conversion of Braille to text in English, Hindi and Tamil languages ," International Journal of Computer Science, Engineering and Applications (IJCSEA), Vol.3, No.3, June 2013.

[6] Vandana, Nidhi Bhalla and Rupinderdeep Kaur, "Architecture of Gurmukhi to Braille conversion system," International Journal of Computer Science and Information Technology and Security (IJCSITS), Vol. 2, No.2, April 2012.

[7] Mohammed Y. Hassan and Ahmed G. Mohammed, "Conversion of English characters into Braille Using neural network ," IJCCCE, Vol. 11, Issue 2, pp. 30- 37, Jan 2011.

[8] Xuan Zhang, Cesar Ortega-Sanchez, and Iain Murray, "Text to-Braille translator in a chip," 4th International Conference on Electrical and Computer Engineering ICECE, December 2006.

[9] Vandana, Rupinderdeep Kaur and Nidhi Bhalla, "Implementation of Gurmukhi to Braille," International Journal of Computer Science And Technology, Vol. 3, Issue 2, April – June 2012.

[10] Manzeet Singh and Parteek Bhatia, "Automated conversion of English and Hindi text to Braille representation," International Journal of Computer Applications, Vol. 4, issue 6, pp. 25-29, July 2010.

[11] Manzeet Singh and Parteek Bhatia, "Enabling the disabled with translation of source text to Braille," proceedings of national conference on advanced computing, communication engineering and technology, January 2010.

[12] Xuan Zhang, Cesar Ortega-Sanchez and Iain Murray, "Hardware-Based Text-To-Braille Translator," 4th International Conference on Electrical and Computer Engineering ICECE, October 2006.

[13] Mukul Bandodkar and Virat Chourasia, "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, Vol. 8 No. 6, 2014.

[14] Sukhpreet Singh, "Optical Character Recognition Techniques: A Survey," 4th International journal of Emerging Trends in Computing and Information Sciences, Vol. 4, No. 6 June 2013.

[15] TirthankarDasgupta, Manjira Sinha and Anupam Basu, "Forward Transliteration of Dzongkha Text to Braille," Proceedings of the Second Workshop on Advances in Text Input Methods (WTIM 2), pp. 97–106, December 2012.

[16] Al-Salman AbdulMali, AlOhali Yosef, AlKanhal Mohammed and AlRajih Abdullah, "An Arabic Optical Braille Recognition System," Proceedings of the First International Conference in Information and Communication Technology, pp. 81–86, April 2007.

[17] M.Rahmath Riswana, Rani Thottungal and V.Kandasamy, "Multi Character Identification for Visually Impaired Using Braille System," Proceedings of International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 1, Issue. 3, pp. 186–191, March 2013

[18] R. John, G. Raju and D. Guru, "1D wavelet transform of projection profile for isolated handwritten Malayalam character recognition," Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference, Vol .2 pp. 481–485, 2007.

[19] M. Rahiman and M. Rajasree, "Printed malayalam character recognition using back-propagation neural networks," Advance Computing Conference, IEEE International, pp. 197–201, 2009.

[20] J. John, K. Pramod, and K. Balakrishnan, "Unconstrained handwritten malayalam character recognition using wavelet

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

transform and support vector machine classifier," International Conference on Communication Technology and System Design, Vol. 30, no. 0, pp. 598–605, 2012.
[21] B. L. Shivakumar and M. R. Thipathi, "English to Braille Conversion Tool using Client Server Architecture Model," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue. 8, pp. 295–299, August 2013.
[22] P. Blenkhorn and G. Evans,, "Automated Braille Production from Word-Processed Documents," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 9, Issue. 1, pp.81–85, March 2001.
[23] Bharati Braille URL: http:\\acharya.iitm.ac.in\disabilities \bhbrl:php:
[24] Acharya-Multilingual Computing for Literacy and Education [Online] Available: www:\\acharya:gen:in\IntroductiontoBraille:htm:

(This work is licensed under a Creative Commons Attribution 4.0 International License.)