IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Review on Gujarati to Braille Character Recognition

[1]Pof. Namita Patel, [2]Prof. Samarth Modi, [3]Prof. Shachi Joshi

[1] Assistant professor,[2] Assistant professor,[3] Assistant professor,
[1,2,3] Aditya Silver Oak Institute of Technology, Ahmedabad, India

Abstract: This paper describes the classification of a subset of printed or digitized Gujarati characters. Gujarati belongs to the genre of Devanagari scripts from the Indian subcontinent. Very little work is found in the literature for recognition of Indian language scripts. For this paper a subset of similar appearing Gujarati characters was chosen and subjected to classification by different classifiers. The sample and test images for the characters were obtained from digital images available on the Internet and from scanned images of printed Gujarati text. For their classification, the Euclidean Minimum Distance and the k–Nearest Neighbor classifiers were used with regular and invariant moments. The characters were also classified in the binary feature space using Hamming Distance classifier. The pa- per presents the recognition rates for these classifiers. A recognition rate is achieved. The work described in this paper is preliminary; however, since ICDAR'99 is being held in India, we hope that this would be of interest to the participants.

Index Terms **- Gujarati language, Braille language, Text processing, Transliteration, Bhartiya Braille**

I. INTRODUCTION

Now-a-days IT is penetrating in each and every field. This is the era when day-by-day new technologies are emerging, it becomes necessary for researchers to develop tools which increase literature in Braille for visually impaired people. India is a multilingual country. There are various scripts such as Hindi, Bangla, Kannada, Punjabi, Rajasthani, Malayalam, Marathi and Gujarati is also one of it. In Indian languages, there is no notion of uppercase and lowercase. Sometimes, two or more characters are combined to form compound characters. There are huge character sets (vowels and consonants) in the Indian Languages also [1].

Gujarati Script

Gujarati script is derived from Devanagari script [2]. The shapes of Gujarati characters are also very typical. Gujarati is descended from Sanskrit. There are over 50 million people worldwide who use Gujarati for writing and speaking. The earliest known document in Gujarati script is a manuscript dating from 1592, and the script first appeared in print in 1797 advertisement [3]. Gujarati alphabet utilizes overall 75 distinct legitimate and recognized shapes, which mainly includes

59 Characters and 16 diacritics. Fifty-nine characters are divided into 36 consonants (34 Singular and 2 Compound (not lexically though)) means ornamented sounds, 13 vowels (pure sounds), and 10 numerical digits [4]. Sixteen diacritics are divided into 13 vowels and 3 other characters. The alphabet is ordered logically by grouping the vowels and the consonants based on their pronunciations [5].

Gujarati is a phonetic language in western India. Gujarati script is written from left to right, with each character representing a syllable. The vowels are called Swar and consonants are called Vyanjan. Gujarati consists of a set of special modifier symbols called Maatras, corresponding to each vowel, which are attached to consonants to change their sound. Modifiers are placed at the top, at bottom right or at bottom part of consonant. They are attached at different positions for different consonants. They can also occur in different shapes as shown in Fig 1.

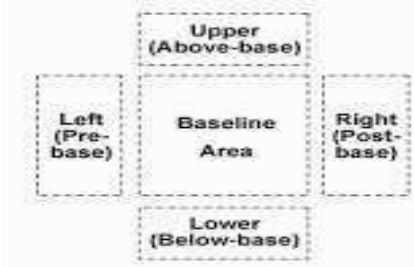

Fig. 1 Gujarati Characters and digits

A character is conjunct, if two half consonants are joined [6]. So, characters in Gujarati can be the combination of consonant, vowels and diacritics. Gujarati language has many characters as shown in Fig. 2.

Fig. 2 Gujarati Characters and digits

Braille Language

Braille is a tactile writing system used by blind and visually impaired. It is traditionally written in embossed form. They can write Braille with the original slate and stylus or type it on Braille writer, such as portable Braille note-taker, or on a computer that prints with a Braille embosser [7].

Braille is named after its creator, Frenchman Louis Braille born in Coup ray village in 1809, who went blind following a childhood accident [7-11]. In 1824, at the age of 15, Braille developed his code for French alphabet as an improvement on night writing [7]. He published his system, which subsequently included musical notation, in 1829 [12].

Standard Braille is an approach to create documents which could be read through touch. This is accomplished through the concept of a Braille cell consisting of raised dots on a thick sheet of paper. The protrusion of the dot is achieved through a process of embossing. A cell consists of six dots arranged in the form of a rectangular grid of two dots horizontally (row) and three dots vertically (column). With six dots arranged this way, one can obtain sixty-four different patterns of dots. A visually Handicapped person is taught Braille by training them in discerning the cells by touch, accomplished through their fingertips. Each arrangement of dots is known as a cell and will consist of at least one raised dot and a maximum of six [13]. The layout of Braille cell is shown in fig. 3 [9-11, 14-15].

Fig. 3 A Braille cell with 6 dots

A printed sheet of Braille normally contains upwards of twenty-five rows of text with forty cells in each row. The physical dimensions of a standard Braille sheet are approximately 11 inches by 11 inches. The dimensions of the Braille cell are also standardized but these may vary slightly depending on the country. The dimension of a Braille cell, as printed on an embosser is shown in fig. 4 [13].

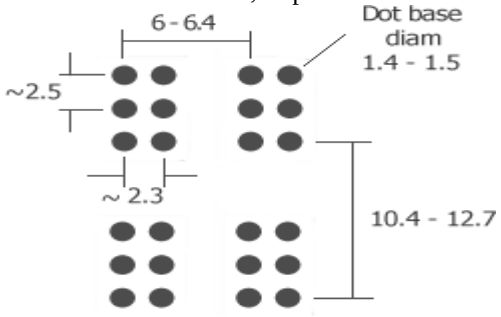

Fig. 4 A Braille cell dimension (in millimeters)

All dots of a Braille page should fall on the intersections of an orthogonal grid. When texts are printed double-sided (recto-verso), the grid of the verso text is shifted so that its dots fall in between the recto dots as shown in fig. 5 [16].

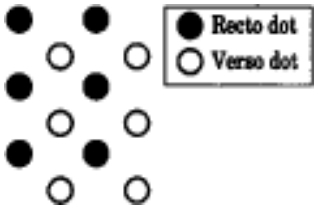

Fig. 5 Positioning of recto and verso Braille.

Bharati Braille or Bhartiya Braille is a largely unified Braille script for writing the languages of India. When India gained independence, eleven Braille scripts were in use, in different parts of the country and for different languages [17]. Gujarati Braille is one of the Bharati Braille. Fig. 6 shows Gujarati alphabet and punctuation as written Braille [17].

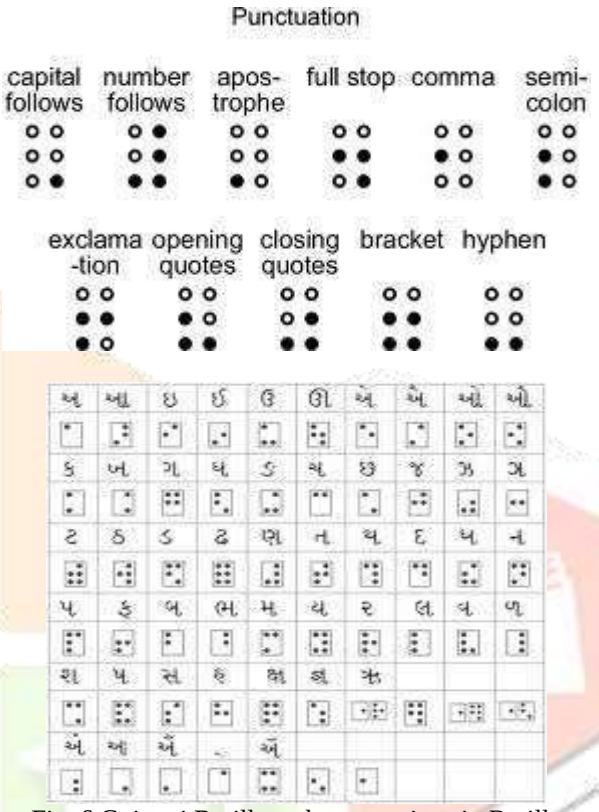

Fig. 6 Gujarati Braille and punctuations in Braille

II. LITERATURE REVIEW

The overview of some of the research work related to the conversion of various types of scripts to Braille is as follows:

P. Blenkhorn et. al. [18] have worked on the problem of converting Word-Processed Documents into formatted Braille documents. The problem has been addressed in context to the users of the word processor who want to produce Braille documents. The translator will be integrated with the word processor. It makes easy to use as translation is possible with the help of menu items in MSWord. They have used 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.

P. Blenkorn [10] has addressed the problem of converting Braille characters into print. The problem has been addressed in the context of Standard English Braille Characters. To address this problem the researcher has presented Finite State System and matching algorithm. Author has also focused on the contracted English Braille. He has also focused on the letter sign conversion. A detailed table is given in the paper which is used by the author to generate the Text. The author has also used decision table concept to clearly state all the rules used in the algorithm. The system is tested on the set of standard English Braille words and on the Braille, files produced by Torch Trust for Blind.

M. Singh et. al. [9] addressed the problem of converting English and Hindi text into Braille characters. The authors have generated the database of the English and Hindi characters along with Braille equivalent. They read a character string, Separate the words on the basis of blank space, Break the corresponding word into corresponding letters, Access look- up table for matching characters of input string with look-up table characters, If characters match, then print corresponding Braille characters as it is and Repeat steps until all the characters of input string are matched with look-up table characters. They have tested their system for some Hindi and English newspapers. They got 100% accuracy for English text to grade I Braille and 99% accuracy for Hindi text.

M. Y. Hassan et. al. [14] have addressed the problem of Conversion of English Characters into Braille. Neural Network is used to recognize the English characters. MATLAB programming is used to design the tool. A scanned page image is converted into 400 elements columns that represent a 20*20 grid for each character. The output of the network is used to generate the corresponding Braille character according to the Braille rules. The generated Braille characters are saved as a matrix of 2*6 elements for each character. NN is designed and trained to recognize 63 characters using a back-propagation training algorithm. Conversion using NN is a fast and accurate method. Dr. B. L. Shivakumar et. al. [19] have addressed the problem of converting English text to Braille. 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 Client/Server Architecture. The database stores all the Braille cell formats. The steps are as follows: Read the input value up to the enter key, Separate the words on the basis of blank space, Break the word into single letters and Access the Braille database based on various rules and conditions. The software will check all 64 (26) combinations of Braille matching value for all input character values. And if the match is not found then it will provide an appropriate error message. It accurately works for pre-defined set of characters in the database.

X. Zhang et. al. [20] have addressed the problem of translating text to Braille based on FPGAS. Compared with most of the commercial methods, their translator is able to carry out the translation in hardware instead of using software. To achieve fast translation FPGA with a big programmable resource has been utilized, and an algorithm, proposed by

P. Blenkhorn, has been revised to perform fast translation. Their system achieves superior throughput compared to Blenkhorn's original algorithm. H. R. Shivakumar et. al. [21] have developed a user-friendly interface, for the rapid and efficient conversion of printed Tamil Books for the use of persons with visual disability. The tool has been developed in Java using Eclipse SWT and runs on Linux, Windows and Mac. It is developed as an open source project.

T. Dasgupta et. al. [22] 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.

O. Khan Durrani et. al. [23] have designed an architecture for transcribing Braille from optically recognized Indian language. The system will help to convert masses of information in different Indian languages into a tactile reading form. The system mainly consists of OCR modules designed in an efficient manner to promote portability and scalability. They have also introduced the importance and necessity of the work, the properties of Braille and Indian scripts. They have also described the OCR work done with respect to Indian languages and the related work to our system. Finally, the System architecture is explained clearly followed by some conclusion and future work. The paper also identifies the needs to be fulfilled to percolate the benefit of the technology developed to the masses.

III. PROBLEMS

As mentioned in Braille language, characters are made up of cells as shown in fig. 2. One cell is made up of 6 dots. So, a total of 64 characters can be formed through it. But in Gujarati there are a total of 75 characters. So, in Gujarati Braille as shown in fig. 4, some characters are identical i.e. same dots are used to form the characters. So, some assumptions are to be considered while writing Gujarati Braille. For example, 0-9 digits are represented in the same way as some of the consonants or vowels i.e. is represented in the same way as અ, ૧, ૨, ૫ is represented in the same way as and so on. So, '#' character which is a digit identifier is used to indicate that the character written is digit. Another example, as known in Gujarati language there are compound characters. And in Braille there is no separate character available to specify half or compound characters. So, again here 4th dot within a cell is used as an identifier which indicates that following character is half character. To write some Gujarati characters in Braille combinations of more than one character is used. For example, શ્ર in Gujarati is written in Braille as a combination of 4th dot, શ, ર (3 characters). Another important consideration is, Braille word building also depends on the pronunciation of the words. It is spelled as it is pronounced. So, to Convert Gujarati text into Braille above mention challenges are to be considered otherwise the meaning of the text may change.

IV. CONCLUSION

The paper describes that a reasonable amount of work is done in various languages to Braille conversion. But still more and accurate work is required in Gujarati text to Braille conversion. So, low cost technology should be developed to assist visually impaired people. If work is carried out in this domain then literature will also increase for visually impaired people. It will be very useful for the visually impaired people and also for associated people who want to learn about the Braille language.

REFERENCES

1. N. B. Jariwala and Dr. B. C. Patel, "Multi-Oriented Gujarati Characters Recognition: A Review" International Journal of Research in Computer Science and Information Technology, Vol. 2, Issue 2(A), March 2014, P.P.1-5.
2. M.J. Baheti, K.V.Kale, M.E. Jadav, "Comparison of Classifiers for Gujarati Numeral Recognition", International Journal of Machine Intelligence, Vol. 3, Issue 3, 2011, P.P.160-163.
3. Omniglot – The Online encyclopedia of writing systems &Languages.
4. Babu Suthar - Gujarati-English Learner's Dictionary.
5. M. Kayasth and Dr. B. C. Patel, "Offline Typed Gujarati Character Recognition", National Journal of System and Information Technology, Vol. 2, Issue 1, 2009, P.P. 73-82.
6. B. Sojitra and V. Dhakad, "Neural Network in Character Recognition of Gujarati Script", Journal of Information, Knowledge and Research in Computer Engineering, Vol.2, Issue 2, Nov 12 to Oct 13, P.P. 269- 272.
7. Braille- The online encyclopaedia [Online] Available: http://www.Wikipedia.com/ Braille - Wikipedia, the free encyclopedia.htm.
8. A. Chaudhary, P. Garg and A. Agarwal, "Using Rotation Method for Removal of Misalignment of Scanned Braille Pattern", Proc. Of the Second International Conference on Advances in Computing, Control and Communication, 2012, P.P. 71-75 [Digital Library].
9. M. Singh and P. Bhatia, "Automated Conversion of English and Hindi Text to Braille Representation", International Journal of Computer Applications, Vol. 4, Issue 6, July 2009, P.P. 25-29.
10. P. Blenkhorn, "A System for Converting Braille into Print", IEEE Transactions on Rehabilitation Engineering, Vol. 3, Issue 2, June 1995, P.P. 215-221.
11. S. Halder, A Hasnat, A. Khatun, D. Bhattacharjee and M. Nasipuri, "Development of Bangla Character Recognition (BCR) System for Generation of Bengali Text from Braille Notation", International Journal of Innovative Technology and Exploring Engineering, Vol. 3, Issue. 1, June 2013, P.P. 5-10.
12. Louis Braille, 1829, Method of Writing Words, Music, and Plain Songs by Means of Dots, for Use by the Blind and Arranged for Them.

13. Acharya-Multilingual Computing for Literacy and Education [Online] Available: www.acharya.gen.in/ Introduction to Braille.htm.
14. Dr. M. Y. Hassan and A. G. Mohammed, "Conversion of English Characters into Braille Using Neural Network", IJCCCE, Vol. 11, Issue 2, Jan 2011, P.P. 30- 37.
15. S. D. Al-Shamma and S. Fathi, "Arabic Braille Recognition and Transcription into Text and Voice", Proc. Of 5th International Biomedical Engineering conference, Dec 2010, P.P.227-231.
16. J. Mennens, L. V. Tichelen, G. Francois and J.L. Engelen, "Optical Recognition of Braille Writing Using Standard Equipment", IEEE Transactions on Rehabilitation Engineering, Vol. 2, Issue. 4, Dec 1994, P.P. 207-212.
17. Bharati Braille- The online encyclopaedia [Online] Available: http://www.Wikipedia.com/Bharati Braille - Wikipedia, the free encyclopedia.htm.
18. P. Blenkhorn and G. Evans, "Automated Braille Production from Word-Processed Documents", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 9, Issue. 1 March 2001, P.P. 81-85.
19. Dr. 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, August 2013, P. P. 295- 299.

X. Zhang, C. Ortega-Sanchez and I. Murray, "A System for Fast Text-To-Braille Translation Based on FPGAS" [Online] Available: http://ieeexplore.ieee.org

H. R. Shivakumar and A. G. Ramakrishnan, "A Tool that Converted 200 Tamil Books for use by Blind Students", [Online] Available: http://tdil-dc.in

22. T. Dasgupta, M. Sinha and A. Basu, "Forward Transliteration of Dzongkha Text to Braille", Proc. Of 2nd Workshop on Advances in Text Input Methods, Dec 2012, P.P. 97-106.
23. O. Khan Durrani and K. C. Shet, "A New Architecture for Braille Transcription from Optically Recognized Indian Languages", International CALIBER, Feb 2005.