Transliteration of Indian Ancient Script to Braille Script using Pattern Recognition Technique: A Review

Kirti Nilesh Mahajan, PhD

Professor (B.Sc., MCM, MCA, Ph.D., APHRM)
Bharti Vidyapeeth Deemed University
Institute of Management & Entrepreneurship
Development, Pune

ABSTRACT

Strenuous research has been done on pattern recognition and a huge number of research works have been published on this topic during the last few decades. The Indian ancient scripts are a golden treasure of not only Asian continent, but also the whole world. Many researchers and an organizations still working on the appropriate character recognition of Indian ancient scripts. This paper presents some research work that has been significant in the area of pattern recognition and also highlights the review of existing work done on the Indian ancient scripts. The purpose of this research work is to study the appropriate identification of the Indian ancient script using pattern recognition techniques. We have studied the different pattern recognition techniques and its categorized steps. In current research work we have used MODI script. Modi is Indian ancient script which is normally preferred in western and southern part of India. This paper shows the importance of MODI script and by using technique of pattern recognition we transliterate into the readable and an audible format for visually impaired people. The review is categorized into four sections. The Section 1 present the Background of Research paper, Section 2 covers the introduction part of MODI Script, Braille Script and Pattern Recognition Technique's, which are the primary part of our research, Section 3 describe the review of literature, and at the end in Section 4 we conclude the paper.

Keywords

Indian Ancient Script, Pattern Recognition, MODI Script, Braille Script.

1. INTRODUCTION

The first most important factors that drive us to implement this system is because of less interaction of blind people with outer world. Using such system they can contribute their own thoughts, ideas for not only educational purpose but also for other administrative work. The second significant aspects that motivate us, is because of less availability of Ancient Script Transliteration software for visually impaired people. According to World Health Organization (WHO) 285 million peoples are Visually Blind in overall world in that 39 million peoples are completely blind and 246 million peoples have low vision [1]. Also, according to the Government of India, Ministry of Home Affairs Office of Register General and Census Commissioner, India report of the 2011 Census that count of Visually Disabled population in India is 18.16 % and in Pune district the count is 16.60% following chart show the classification of the disabled population by Gender wise of India and Pune District[2].

Niket Pundlikrao Tajne

Research Scholar, Bharti Vidyapeeth Deemed University Institute of Management & Entrepreneurship Development, Pune

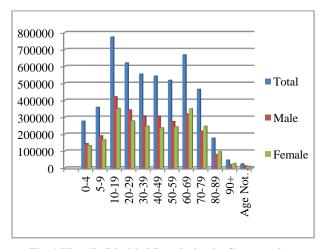


Fig. 1 Visually Disabled Population by Sex age wise-Census of India

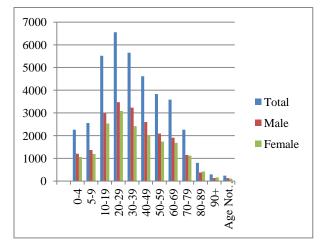


Fig. 2 Visually Disabled Population by Sex age wise-Census of India in Pune District

2. INTRODUCTION TO TECHNIQUES

Ancient languages are a Significant for any region. Many Handwritten recognition systems were developed for different foreign languages like Japanese, Chinese, French, German, etc. But still works on Indian ancient languages are unbeatable arena. The conversion process work for any primary language of any region to braille script is still going on not only India but also in other countries. In India an initiative was taken by the Government of India for the development of Indian language by starting a program TDIL (Technology Development for Indian Languages) [34]. In this section we present the introductory part about MODI Script, Braille Script and Pattern Recognition Techniques.

2.1 MODI script

MODI is one of the script used to write the Marathi language, which is primarily used in the state of Maharashtra located in the western part of India. The original meaning of MODI word is inspired from the Persian word "Shikasta" which means breaking in English and in Marathi it means "Modne" [3,31]. There are some evidences about the origin of this script, says that the proper MODI alphabet originated in the 17th century. Which was used for Administrative work, Education, and other routine activities. But actually MODI script was developed by Hemandpant or Hemadri in 12th century. So from the 12th century till 17th century there are many changes has been made in writing style of MODI." After 17th century, it was used rapidly in western and southern part of India till mid of 19th century. Following table shows the different name with centuries of MODI script from past 600 to 700 years [4, 5].

Table 1. Name with centuries of MODI script

Sr. No.	Century	Name of MODI script
1	12 th Century	Proto MODI or Adyakalin
2	13 th Century	Yadavakalin
3	14 th to 16 th century	Bahamanikalin
4	17 th Century	Shivakalin
5	18 th Century	Peshvekalin

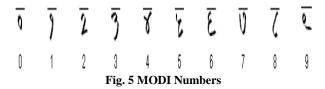
Also the MODI was used in the primary school books during the 19th century. The MODI alphabet are variant of the Marathi Language and it has 14 vowels, 34 consonant letters and 10 Numbers [3].



Fig. 3 MODI Vowels

χ	KA	$\overline{\mathcal{A}}$	NYA	ŧŢ	DHA	8	LA		
र्ग	KHA	J	TTA	7	NA	Ū	VA		
ग्	GA	ช	TTHA	IJ	PA	Ţ	SHA		
घ	GHA	3	DDA	प्र	PHA	Ą	SSA		
<u>5</u> .	NGA	\$	DDHA	£	BA	Ţ	SA		
उ	CA	ण्	NNA	भ	ВНА	Ŋ	HA		
ঘ	CHA	7	TA	χ	MA	$\overline{\mathfrak{p}}$	LLA		
ष	JA	\mathfrak{t}	THA	पु	YA				
झ	JHA	IJ	DA	J	RA				
	Fig. 4 MODI Consonants								

Fig. 4 MODI Consonants



2.2 Braille Script

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 not a language, it is Script which primarily used by the blind people. Initially the Braille was existence by Charles Barbier for soldiers in 1825. It uses a 6x2 matrix in which a maximum 12 dots in a single cell. In 1829 Louis Braille introduced the first Braille system for common people [26]. It uses a 3x2 matrix in which a maximum of six dots can be embossed. The size of this matrix allowed $64(2^6 = 64)$ possibilities for creating different characters [25]. Braille system is totally depending on the cell which is made up by six dots numbering from 1 to 6. The six dots are arranged with the group of three, vertically. The structure of a single cell of the braille script as shown in Figure 4[21, 24-28].

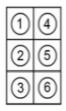


Fig. 6 Braille Cell

Depend on the arrangement of dots in a cell, the cell is categorized into alphabets, numbers and punctuation marks and the combination of these categories we can make any words. The visually impaired people read these cells through the touch of his/her fingertips. Reading and writing rules are same for different categories of letters, number's, and punctuation marks. For example, following Figure show Braille structure of numbers which has two cells the first cells represent the given character is number and the second one is an actual number [40].

:::::	::::	:::::		::::
1	2	3	4	5
:: ::	:: ::	::	:: ::	:: ::
6	7	8	9	0

Fig. 7. Braille Number System

Table 2. List of Pattern Recognition Applications

Problem Domain	Applications	Input Pattern	Pattern Classes	
Bioinformatics	Sequence Analysis	DNA / Protein Sequence	Known Types of Genes or Pattern	
Data Mining	Searching For meaningful patterns	Points in multidimensional space	Compact and well separated clusters	
Document Classification	Internet Search	Text Document	Semantic Categories	
Document Images Analysis	Optical Character recognition	Document image	Alphanumeric characters, word	
Industrial Automation	Printed circuit board inspection	Intensity or range image	Defective / non-defective nature of product	
Multimedia Database retrieval	Internet search	Video Clip	Video genres (e.g. Action, Dialogue, etc.)	
Biometric recognition	Personal identification	Face, Iris, Fingerprint	Authorized users for access control	
Remote sensing	Forecasting crop yield	Multi spectral image	Land use categories, growth pattern of crop	
Speech Recognition	Telephone directory	Speech waveform	Spoken words	
Medical	Computer aided diagnosis	Microscopic image	-	
Military	Automatic target recognition	Optical or infrared image	Target type	
Natural language processing	Information extraction	Sentences	Parts of Speech	

2.3 Pattern recognition technique

According to Wei-Lun chao [14] "Pattern Recognition is a Process that is taking in raw data and making an action based on the category of the pattern". Also the Americans mathematician and philosopher Nobert Wiener[13] "A Pattern is essentially an arrangement. It is characterized by the order of the elements of which is made, rather than by the intrinsic nature of these elements". The Research of pattern recognition Methods are Statistical Pattern Recognition, Data Clustering, Neural Network, Template matching model, syntactic model[15,16], We can utilize the techniques of Pattern Recognition in many fields by many ways. Following table show the pattern recognition problem domain list with its applications [8-12].

Many types of Approaches are present in the pattern recognition and all are very useful following table show the pattern recognition approaches.

3. REVIEW OF LITERATURE

The Review of Literature section we categorized into two subsections. In first section we present the existing literatures work of MODI script, Braille Script and some other languages. In Second section we present literatures of reviews, issues of MODI script, Braille script transliteration system using Pattern Recognition techniques.

3.1 Literature based on Existing work

Nikisha B. Jariwala, Bankim Patel (2015), has proposed the Review on conversion system of Gujarati to Braille script also they addressed the problem of Gujarati to Braille script conversion like in braille one cell is made up of 6 dots, so total 64 characters can be formed through it. But in Gujarati there are total 75 characters. Also braille word building also depends on the pronunciation of the words it is spelled as it is pronounced [26]. Er. Vandana, Er. Nidhi Bhalla and Ms. Rupinderdeep Kaur (2012), has proposed the Gurmukhi to Braille conversation. They accept the input in the form of string or character, tokenize that character and match the

tokenize character with existing token from database and then braille Processor converts the character input into the braille code or cell[23].

Dr. B. L. Shivakumar, M. Rajasena thipathi (2013), researcher gave the methodological implementation of the English text to braille code conversion system [21]. Tirthankar Dasgupta, Manjira Sinha and Anupam Basu (2012), The Researchers proposed the system which converts the Dzongkha (Bhutanese script) to Braille Script. They accept the input from QWERTY keyboard, used the Unicode Dzongkha script and match that Unicode of Dzongkha script with Braille script Unicode and finally generate the output. Demerits of the proposed system is first they accept input from QWERTY keyboard which is quite uncomfortable for visually challenged people and second demerit is they perform the conversion only in one way[20].

P H Rathod, M L Dhore, R M Dhore(2013), they proposed system which perform the transliteration of Hindi to English and Marathi to English machine transliteration using SVM (Support Vector Machine) there are total three phases in proposed system architecture first phase is preprocessing on raw input data second phase is Training data which has again two sub phases like training data and other is features on which the data is to be trained and last phase is Testing phase which accept two files one is model file that generated at the time of Training phase and second file is Testing file[19].

John Esan O. Ananayo, john Daniel B. Cayaos, Fancis Gerome T. Rosal,(2011), have addressed some merits like formed accurate meaning, easily adopt modern language, no maintance for the Dictionary and spelling doesn't matter and it has also some demerits cannot fix grammar, some kankanaey is limited to word per word, more bilingualr corpora are needed[17]. Abdul malik S, Al Salman Ali El-Zaart, Yousef Al Subaibani Khaled Al-Hakail, Abdul Gumaei (2014), they proposed the system which is divided in to two section first is hardware and second is software section in software is used the C# language and software section is again divided into three different steps that are scanned image, preprocessing the image and finally generate the ASCII code of the character [30].

Pallavi Aggarwal, Yashasvi Rawal, (2014), describe the methodology in to some steps these are Reading an Image, Converting to Gray Scale Image, Converting to Binary Image, Detecting Edge in the Image, Image Dilating using Edges, Image Filling[27].

Rahul Raghvendra Joshi, Maya Shelke, Joya Prakash, Swapnil Bhadhane(2013), has proposed a system interface that does automatic transliteration and translation of Marathi to English language. Which has total six steps these are accept the input from the keyboard, splitting the words into small parts i.e. perform the parsing phase on words, check with existing database, transliterated the words, translate into the English word, and finally generate the finished output[22]. K. Karthik, D. Vaishnavi, E l. Ramkumar, I. S. Nivethitha (2015), has proposed the system which transliterated process for German and Tamil for experimental purpose and they have shown good results based on pattern recognition [18].

Latha R. Nair, David Peter S., (2012), represent the analytical study of Indian language translation tools with appropriate result in text as well as in audio format. Also they proposed the optical character recognition methodology for conversation of Germen and Tamil languages [29]. Saad D. Al-Shamma and Sami Fathi, (2010), has proposed the

methodology of Arabic to Braille recognition into text which has seven steps like Braille document scanning, Image preprocessing, Image filtering, Braille Cells and Dots Framing, Decimal Braille Code Generation, Braille letter recognition and transcription, and Braille word recognition and transcription[24].

A S Ramteke, G S Katkar(2013), has suggest the different methodology used for the recognition of off line MODI script like classification method, Preprocessing, segmentation, Neural Network, Neuro Fuzzy Logic[6]. Sonali R. Mayne, Monika P. Dhobale, Rohini H Jadhav, Rushikesh A Ambildhok (2015), has focus on proposed system of handwritten MODI characters conversion in English for that researcher describe the step by step preprocessing algorithms like gray scale algorithm for image processing, Otsu's method for image thresholding, Boundary Detection, Cropping, Noise Removal, Stentiford thinning algorithm, Scaling and Template Matching. Another significant approach of this research article is researcher mention that some demerits of the research like they formed the final output based on limited samples and another demerits is because of less people how experts in MODI scripts so it take more time to convert MODI manuscript in English[7].

3.2 Literature based on Issues and Reviews

Vandana Pushe, Rupinderdeep Kaur (2011), and Salah C. A. Ranjith Ram(2015)" combinally concluded that in these literatures the researchers find the different issues for converting the Punjabi to Braille scrip and Malayalam Text to Braille. Basically researcher focused on open source, lack of database, scanning problem, and symbol problem the existing system is not making any difference between vowels and consonants [33, 32].

Amarpreet Kaur, Er. Jyoti Rani, (2014), and Veerpal Kaur, Amandeep kaur Sarao and Jagtar Singh(2014), researchers gave the implementation of Punjabi to Hindi and Hindi to English language Transliteration by using Statistical translation technique which has one to one mapping of characters[39,37]. Devinder Brar, Er. Rishamjot Kaur, (2014), has Describe the Significant of Transliteration, types of transliteration, difference between Transliteration and Transcription and finally the small design and implementation tool for the transliteration from English to Punjabi language which has four steps like accept input, tokenize the input, match or mapping the tokenize input with existing data and finally generate the final output which is in Punjabi language [38].

Sumita Rani, Dr.Vijay laxmi(2013), have focus on a proposed system of transliteration from Punjabi to Hindi language in that researcher focus one to one character mapping using statistical pattern recognition technique[36]. Salah C. And A. Ranjith Ram(2015), has 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 by using SVM (Support Vector Machine) also the researchers suggest by using SVM method transliterated very fastly, develop output in low cost with accurate solution[35].

4. CONCLUSION

After recapitulating related research paper we concluded that not much work is done for Indian ancient script like MODI script. Number of organization's / institutions are working on transliterating MODI script in to some primary languages but still ample amount of work is required for visually impaired

people. Major weakness of transliteration from MODI script to Braille script are varying writing style of each individual character in MODI script, Second issue in converting any language to braille is scanning problem, symbol problem, and limited database. The third problem is Original document has an unwanted line character which changes the meaning of the existing Document. Next issue is poor word selection of the character from the existing document may change the meaning of the existing ancient document and many more. It is concluded that this study is helpful for MODI script writer, reader and most importantly visually impaired people so they can contribute their thoughts and ideas. Also it is shown that the existing systems has some lacunae which is not satisfactory and hence there is a need to develop a new technique which can provide the expected results.

5. REFERENCES

- [1] The World Health Organization (WHO) website. [Online] Available.http://www.who.int/mediacentre/factsheets/fs2 82/en
- [2] The Census of India website. [Online] Available. http://www.censusindia.gov.in/DigitalLibrary/Archive_h ome.aspx
- [3] Mr. Rakesh A. Ramraje, "HISTORY OF MODI SCRIPT IN MAHARASHTRA", Global Online Electronic International Interdisciplinary Research Journal (GOEIIRJ), Volume-II, Issue-I, June 2013.
- [4] Sadanand A. Kulkarni, Prashant L. Borde, Ramesh R. Manza, Pravin L. Yannawar, "Offline Handwritten MODI Character Recognition Using HU, Zernike Moments and Zoning", Available. https://www.researchgate.net/publication/263390464_Offline_Handwritten_MODI_Character_Recognition_Usin g HU Zernike Moments and Zoning.
- [5] Prof. Mrs. Snehal R. Rathi, Rohini H. Jadhav, Rushikesh A. Ambildhok, "RECOGNITION AND CONVERSION OF HANDWRITTEN MODI CHARACTERS", International Journal of Technical Research and Applications, Volume-3, Issue-1 PP 128-131, Jan-Feb 2015.
- [6] A S Ramteke, G S Katkar, "Recognition of Off-line Modi Script: A Structure Similarity Approach", International Journal of ICT and Management, Volume-1,Issue-3, Feb-2013.
- [7] Sonali R. Mayne, Monika P. Dhobale, Rohini H Jadhav, Rushikesh A Ambildhok, "Recognition and Conversion of Handwritten Modi Characters", International Journal of Advanced Research in Computer Science and Software Engineering, Volume-5, Issue-4, 2015.
- [8] S. P. Shinde, V.P. Deshmukh, "Implementation of Pattern Recognition Techniques and overview of its applications in various areas of artificial intelligence", International Journal of advances in Engineering and Technology, Volume I, Issue 4, PP 127-137, September 2011.
- [9] Rajleen Kaur, "An Overview of Pattern Recognition Applications and Its Approaches: A Review", International Journal of Science and Research, Volume 3, Issue 11, November 2014.

- [10] Divya Prabhu K. P. , Priyanka Prabhu, "Overview of Pattern Recognition Techniques", Available: http://israse.com/digital/assets/papers/1ICRASE-12.pdf
- [11] V. Sucharta, S. Jyothi, "A Survey on Various pattern Recognition methods for the identification of a Different types of images", Publications of Problems and Application in engineering Research, Volume 4, Issue 1, 2013.
- [12] M. Subba Rao, Dr. B. Eswara Reddy, "Comparative Analysis of Pattern Recognition Methods: An Overview", Indian Journal of Computer Science and Engineering (IJCSE), Volume 2, No. 3, Jun-Jul 2011.
- [13] Eva R. Toussaint, Godfried T. Toussaint, "What is a Pattern?", Proceedings of Bridges 2014: Mathematics, Music, Art, Architecture, Culture, 2014.
- [14] Introduction to Pattern Recognition by Wei-Lun Chao Available: http://disp.ee.ntu.edu.tw/~pujols/Introduction%20to%20 Pattern%20Recognition.
- [15] Jie Liu, Jigui Sun, Shengsheng Wang, "Pattern Recognition: An overview", IJCSNS International Journal of Computer Science and Network Security, Volume 6, No 6 June 2006.
- [16] Priyanka Sharma, Manavjeet Kaur, "Classification in Pattern Recognition: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013.
- [17] Jan Esan O. Ananayo, John Daniel B. Cayaos, Francis Germoe T. Rosal, "Translation Algorithm English to Kankanaey", Proceeding of the 8th National Natural Language Processing Research Symposium De la Salle University Manila, November 2011.
- [18] K. Karthik, D. Vaishnavi . Ramkumar, I. S. Nivethitha, "Language Translator Using OCR In LabVIEW", IJIRST –International Journal for Innovative Research in Science & Technology, Volume 1, Issue 8 January 2015.
- [19] P H Rathod, M L Dhore, R M Dhore, "HINDI AND MARATHI TO ENGLISH MACHINE TRANSLITERATION USING SVM", International Journal on Natural Language Computing (IJNLC), Volume 2, No 4, August 2013.
- [20] Tirthankar Dasgupta, 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), Page 97-106, December 2012.
- [21] Dr.B.L.Shivakumar, M.Rajasena thipathi, "English to Braille Conversion Tool using Client Server Architecture Model", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013.
- [22] Rahul Raghvendra Joshi, Maya Shelke, Joya Prakash, Swapnil Bhadhane, "DEVELOPING THE TRANSLATION INTERFACE FOR MARATHI TEXT", International Journal of Advanced Engineering Research and Studies, Volume 2 Issue 3, April-June 2013
- [23] Er. Vandana, Er. Nidhi Bhalla, Ms. Rupinderdeep Kaur, "Architecture of Gurmukhi to Braille conversion

- system", IRACST International Journal of Computer Science and Information Technology & Security (IJCSITS), Volume 2 No. 2, April 2012.
- [24] Saad D. Al-Shamma and Sami Fathi, "Arabic Braille Recognition and Transcription into Text and Voice", 5th Cairo International Biomedical Engineering Conference, 2010.
- [25] Bijet Maynoher Samal, K.Parvathi, Jitendra Kumar Das, "A BIDIRECTIONAL TEXT TRANSCRIPTION OF BRAILLE FOR ODIA, HINDI, TELUGU AND ENGLISH VIA IMAGE PROCESSING ON FPGA", IJRET: International Journal of Research in Engineering and Technology, Volume 4, Issue 7, July 2015.
- [26] Nikisha B. Jariwala, Bankim Patel, "Conversion of Gujarati Text into Braille: A Review", International Journal of Innovations & Advancement in Computer Science IJIACS, volume 4, Issue 1, January 2015.
- [27] Pallavi Aggarwal, Yashasvi Rawal, "Applications of MATLAB's Toolbox to Recognize Handwritten Characters Part 2: Experimental Results", International Journal of Engineering Research and General Science, Volume 2, Issue 4, June-July 2014.
- [28] Aisha Mousa, Hazem Hiary, Raja Alomari, and Loai Alnemer, "Smart Braille System Recognizer", IJCSI International Journal of Computer Science Issues, Volume 10, Issue 6, November 2013.
- [29] Latha R. Nair and David Peter S., "Machine Translation Systems for Indian Languages", International Journal of Computer Applications, Volume 39, No. 1, February 2012
- [30] Abdul malik S, Al Salman Ali El-Zaart, Yousef Al Subaibani Khaled Al-Hakail, Abdul Gumaei, "Designing Braille Copier Based on Image Processing Techniques", International Journal of Soft Computing and Engineering (IJSCE), Volume 4, Issue 5, November 2014.
- [31] Shubhangi Bhatambrekar, Niket Tajne, "Context Free Grammar (CFG) for MODI Script", International Conference on Computer Science, Data Mining and Mechanical Engg., April 2015.

- [32] Salah C. A. Ranjith Ram, "Issues in Transliterating Malayalam Text to Braille", International Journal of Engineering Research & Technology (IJERT), Volume 4, Issue 5, May 2015.
- [33] Vandana Pushe, Rupinderdeep Kaur, "ISSUES IN CONVERTING PUNJABI TO BRAILLE", Asian Journal Of Computer Science And Information Technology1:4, 2011.
- [34] Kulkarni Sadanand A. Borde Prashant L. Manza Ramesh R. Yannawar Pravin L., "Review on Recent Advances in Automatic Handwritten MODI Script Recognition", International Journal of Computer Applications, Volume 115, No. 19, April 2015.
- [35] Salah C. and A. Ranjith Ram, "A Review Paper on Malayalam Text to Braille Transliteration", International Journal of Current Engineering and Technology, Volume 5, No 4, August 2015.
- [36] Sumita Rani, Dr.Vijay laxmi, "A Review on Machine Transliteration of related languages: Punjabi to Hindi", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 2, Issue 3, March 2013.
- [37] Veerpal Kaur, Amandeep kaur Sarao and Jagtar Singh, "A Review on Hindi to English Transliteration System for Proper Nouns Using Hybrid Approach", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 5, September 2014.
- [38] Devinder Brar Er. Rishamjot Kaur, "A Review of Transliteration system from English to Punjabi", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 7, July 2014.
- [39] Amarpreet Kaur and Er. Jyoti Rani, "A Review on a Web Based Punjabi to Hindi Statistical Machine Translation System", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 8, August 2014.
- [40] Christine Lauenstein, "On the Compatibility of the Braille Code and Universal Grammar", Institute of English Linguistics Universität Stuttgart, 2007.

[41]

IJCA™: www.ijcaonline.org 38