International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

Energy Efficient Devnagri Unicode Reader Design on Fpga

Nisha Sharma& Bhanisha Verma

Department of Electronics and Communication Engineering Haryana Engineering College Jagadhri, Haryana
nisha.sharma30a@gmail.com

Abstract—

Unicode and ASCII codes are gaining importance day by day with the advancement in the technology. Lot of research work is being done on the ASCII codes as well as Unicodes of different languages across the globe. As per the constitution of the India, there are 22 scheduled languages. It has been observed that amongst all the 22 languages being used Devnagari script is being the primary and most widely used script. Devnagari is used for writing the Hindi language in India. Many papers have been reviewed regarding the latest work in the areas of Unicode. It has been observed that in order to compensate the energy and power crisis being faced by the world, it has become the need of the seconds to make the energy efficient devices such as energy efficient Unicode readers for 22 scheduled languages of India. Since Devnagri being the most popular language amongst the India , so the target language of the paper is Devnagri itself and efforts have been made in the paper to search out the most suitable energy efficient technique to make the Devnagri Unicode reader device, an energy efficient device so as to contribute in the Green Communication .

Keywords—

Unicode; SSTL; Field Programmable Gate Array(FPGA); Eergy Efficient; Devnagri.

I. INTRODUCTION

Unicode font is the font that provides a unique number for each and every character irrespective of the platform, program, language and windows. It is the universally used coding system. Unicode fonts are visible in any computer system as well as in the web, it is the advantages of using Unicode fonts over ASCII fonts[1]. ASCII and Unicode donot play the same role for English character coding system but also for many languages including some major languages[1] The Organization for Unicode had been founded in 1991 to design and promote a universal, uniform, unique, unambiguous worldwide character-encoding standard.[2] The first version of standard, Unicode Standard 1.0 has been released by Unicode Organisation in the same year. As per the current survey the Unicode Standard contains 49,194 distinct coded characters. These 49,194 characters covers all the principal written languages, and various symbol systems of the world.[2]. The Unicode Searching Algorithm which is using multilevel binary tree is proposed to search the Unicode in an efficient way. The algorithm is being applied to Bangla Unicode to convert the Bijoy string into the Unicode string.[3] On Windows CE, input data are often encoded using Unicode before being processed. The shellcode should be Unicode-proof so that encoded in a way that bypasses such encoding.[4] Unicode-proof shellcode is also having a great advantage of evading the instruction detection system.[4] It has been observed that most of the legacy computer systems supports the input and display of 20,902 Han characters (Hanzis for short) encoded in Unicode 1.0. [5].In 2010, it has been observed that Unicode 6.0 has encoded the 75,616 Hanzis. But now a days , it is not easy to use these newly encoded Hanzis, on the latest computers. Most of these newly encoded Hanzis are being used in daily lives but some are only used in the ancient literatures[5]. Unicode based design of the internationalization software will bring a new trend in the industry[6]. There implies a possible doubling of data storage space and data transmission time by seeing the increasing importance of Unicode for text files, for example with Java and in some modern operating systems.[8]. Covertly communication and sending the information secretly have been of great interest for ages. It has been observed that text documents have been widely used and consequently different methods have been developed so far for hiding information in texts (text steganography).[9]. One of the severe web security problem in the near future is visual spoofing in Unicode as more and more Unicode-based web documents will be used. To check that whether a suspicious Unicode character in a word is visual spoofing or not, the context of the suspicious character is being utilized by employing a Bayesian framework. Simple context and general context are the two contexts which have been taken specifically into the consideration. Word is the simple context of a suspicious character while general context consists of all homoglyphs of the character within the Universal Character Set (UCS).[10]. The preferred means for representing symbols used in creating multimedia content, is Unicode and especially for the information that is presented in multiple

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

languages.[11]. The design and development of a novel Odia keyboard layout has been presented in this paper that can be used to type Odia texts in the standard text editor using any QWERTY keyboard. The layout of the keyboard is based on the Unicode 5.0 standard[12]. There is a continuous need for strong cryptographic algorithms with the growth of the communication/transaction over the internet. An improved encryption model, based on UNICODE and Colors map can be used which as an extension of an existing encryption algorithm[13]. Phonemes and allophones of Malayalam and their transliteration and presentation on the web have been focused in the paper. Tables consisting of Malayalam phonetic symbols with their respective Unicode numbers are the outcomes of the study.[14]. In Bengali the sorting order is different from suggested Character order in unicode by the governing authority. Correct order of Bengali words cannot be yield by simply letter by letter comparison. The situation has become more complicated by the presence of modifier characters in Bengali. The applicability of this proposed algorithm is to any chosen sorting order.[15]. One of the most effective method for digital document protection is text watermarking schemes. The invisibility and robustness of embedded watermark can be detracted by the direct implementation of text watermarking to the web document. By using Unicode based web document protection technique, the above problem can be solved.[16]. A reasonable method for displaying online processing and auto type selection has been designed in the paper. Realizations of the inputting and displaying techniques of the minority scripts with Unicode have also been done.[17]. The process of a collection of documents for extracting the specific information is called the process of information extraction (IE). The information is well structured on the Web in format of HTML or XML. Learning techniques have been used for pattern matching in the content and for the work of IE from structured documents (in HTML or XML).[19]. Smart cards are playing an important role in day today's life. Smart cards are so easy to carry, there is nothing wrong in saying that world is fully in the hands of Smart cards and It is secure even if it is stolen. Since with the increase in hackers and eavesdroppers in today's life, smart cards are becoming insecure. It has become easy for the cryptanalysts to encode the encrypted pin. So it is the need of the hour to secure the Smart cards. In the New encryption and a new hash algorithm have been applied in this paper to keep the information secure in Smart cards and to save it from the forgery attacks. O ne of the main advantage of the paper is that it will take 10^{1084} of years to decrypt the supercomputer which have not been achieved in the previous works.[20].

II. LANGUAGED AND TECHNOLOGIES USED

Energy Efficient Unicode reader design frequency scaling technique has been applied to this PUR design and it is being tested with device operating frequency of 25 MHz, 125 MHz, 625 MHz, 1 GHz and 25 GHz and by varying the different IO standards of HSTL Logic Family.[7].

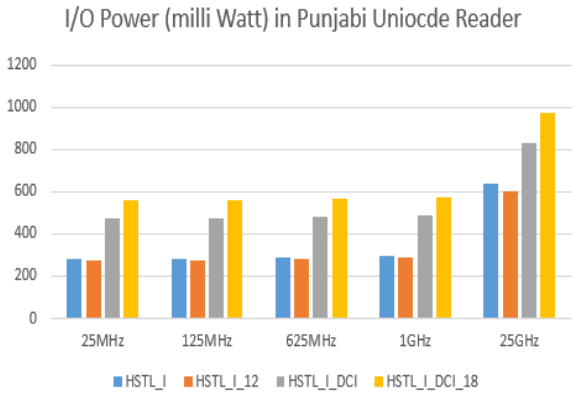

Figure1: Power Dissipation on Different Frequency[7]

There is a wide research gap in designing an energy efficient hardware which is being used in the text analysis. When device operating frequencies are being scaled down from 250GHz to 200GHz, 150GHz and 100GHz respectively in Bengali Unicode Reader and room temperature is 25 degree Celsius there is 60.01%, 39.98%, 20% reduction in Clock Power. It has been observed that there is 26.56%, 64.73%, 79.46% significant reduction in leakage power on scaling down of ambient temperature from100 C to 75 C, 50 C and 25 C.[18].

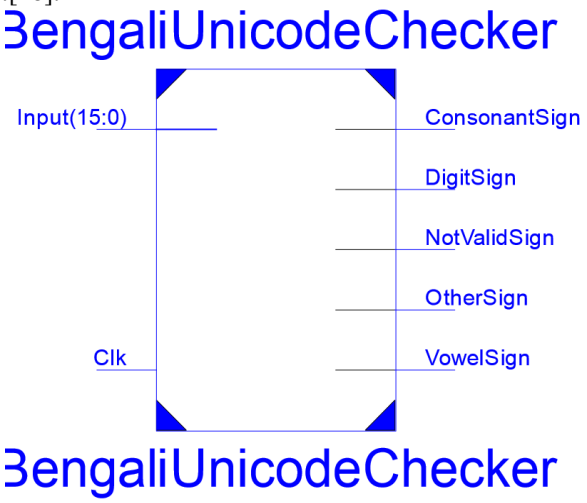

Figure 2: Schematic of Bengali Unicode Reader[18]

The following research work has been focused on designing an energy efficient Greek Unicode reader based on Design Goal Technique for natural

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

processing.[21] In the paper power analysis have been done at different frequencies keeping the temperature constant at 25 degree Celsius and maintaining the constant air flow.[21].Total power consumption at a frequency of 2.2GHZ is shown in the figure below.

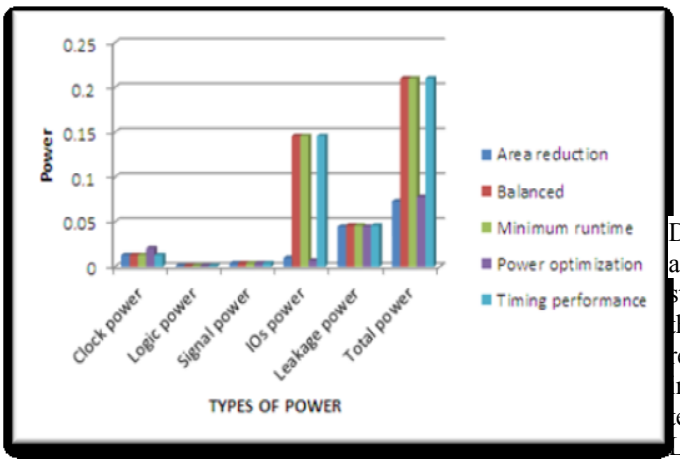

Figure 3: Power Dissipation at 2.2 GHz[21]

The new approach of Voltage Scaling along with the Clock Gating called as (CGVS) have been introduced in the following paper, The main aim of the paper is the designing of a low power Devnagari Unicode Checker (DUC) by applying the CGVS technique. Frequencies have also been changed from low range of MHZ to high range of GHZ.[22] The design is being implemented at 28nm FPGA.The Unicode range for Independent Vowels, Consonants, Digits are shown in the table below.[22]

Table 1: Devnagari Sign and its Range of Unicode[22]

Devnagari Sign	Range of Unicode	Devnagari Sign	Range of Unicode
Independent vowels	0904-0914	Digits	0966-096F
Consonants	0915-0939	Sindhi implosives	097E-097F

As per the Unicode consortium the Malayalam script covers the Unicode range from 0D00–0D7F.The main focus of the paper is to designing of an energy efficient Malayalam Unicode reader. 40nm technology has been used as Virtex-6 FPGA.[23]. The frequency scaling technique has been applied and frequency has been scaled from 1MHz to 1 THz along with the usage of SSTL logic family. It has been observed that maximum power savage of 32-93% have been done by using the SSTL18_1 IO standard of SSTL logic family. [23].

Table 2: Power Dissipation with Class 1I Signaling at SSTL18_II[23]

S.No.	Frequency	VIRTEX-6	SPARTAN-6
1.	1 THz	37.860	27.190
2.	100 GHz	4.764	2.820
3.	10 GHz	1.390	0.361
4.	1GHz	1.052	0.111
5.	100MHz	1.018	0.086
6.	10MHz	1.015	0.084
7.	1MHz	1.014	0.083

Devanagari script is one the most widely used scripts among all the scripts being used in INDIA. Different IO standards have been applied to input and output sides of the circuit in order to avoid the transmission line reflections.[24] In the paper, design has been implemented Virtex-6 along with the frequency scaling technique has been applied to different IO standards of LVCMOS family.[24]Different Energy Efficient Techniques being applied on Devnagri script are shownin the figure below.

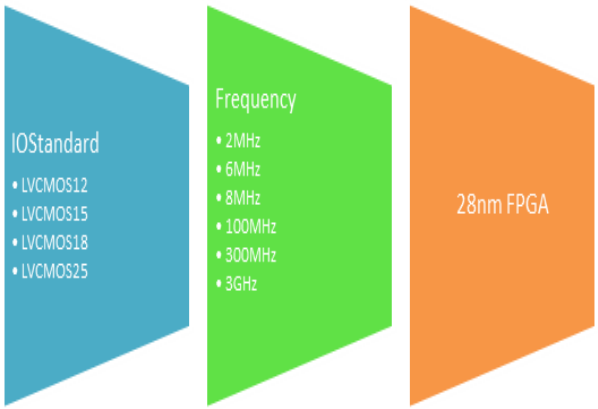

Figure 4: Different Energy Efficient Techniques of Devanagari Unicode Reader[24]

The main area of focus in the paper is the designing of power efficient LATIN Unicode reader design. The design has been operated at different frequencies of mobile architecture. Temperatures of 4 different regions have been taken as a reference to make the design thermally stable.[25] The power of 40nm (Artix-6) technology has been compared to 28nm (Kintex-7) technology and the most efficient technique have been considered further for analysis. Also clock power has not been changed in both the cases (Gated and Non-gated).[25]

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

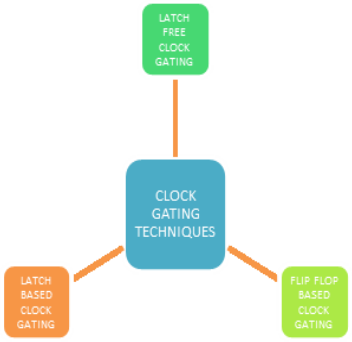

Figure 5: Clock Gating Techniques [25]

In the paper, an I/O Standard based Energy Efficient Devanagari Unicode Reader has been designed.[26] The power analyses have been done using 300MHz, 1GHz and 3GHz frequencies at different IO standards of LVCMOS (Low Voltage Complementary Metal Oxide Semiconductor) logic family.[26] 28 nm Kintex-7 FPGA have been used. The junction temperature has also been changed at different IO standards of LVCMOS. [26].

Figure 6: Specification of Energy Efficient DUR[26]

In thie paper, thermal aware Gurmukhi Unicode Reader design has been focused. [27].The ambient temperature has been changed from 20 oC to 80 oC with consideration of different airflow (250 LFM-500LFM) and Heat sink (medium profile, high profile and standard).[27] With high profile Heat sink and 500LFM airflow, the design has been made more energy efficient and thermally aware.[27]. In the paper, capacitance scaling technique has been used to to make Gurmukhi (Punjabi) Unicode Reader for (Pa). [28]The output load have been changed from 5pF to 25pF to 50pF to 100pF .[28].A regular decrease of 74%, 58% and 38.67% have been observed in IO power by scaling down capacitance from 100pF to 5pf, 25pF and 50pF when the device is operating on a frequency of 1GHz.[28].

III. Conclusion

Various papers have been studied on Unicode has been studied. It has been observed that different energy efficient techniques have been applied to the Unicodes of different languages. Energy Efficient techniques have been applied to Devnagri Unicode Reader till date such as Clock Gating, Change in IO standards of (LVCMOS) logic family along with the Frequency Scaling. But SSTL logic family has not been tested yet on Devnagri Unicode Reader which is previously being done on Malayalam Unicode Reader. So this is the new approach which we can adopt to make the Devnagri Unicode Reader more Energy Efficient along with the change in the frequency from MHZ(low frequency) to THZ (high frequency)range.

IV. Future Scope

There are lot many other energy efficient techniques such as capacitance scaling, design goal, heat sink, proper air flow can be applied to the Devnagri Unicode Reader to make the device more Energy and Power efficient. Also we can apply all the energy efficient techniques to all the other 22 scheduled languages of India. Also we can work on the Ultrascale FPGA technologies such as 20nm and 16nm.

References

[1] S. S Rahaman, A. Mahmud, M. R Islam, & M. A. H. Akhand (2013, May). Design and development of a Bengali unicode font converter. In IEEE International Conference on Informatics, Electronics & Vision (ICIEV), 2013

[2] A. P. Memon,. "Study of Unicode specifications and their implementation in Arabic script languages by designing a multilingual Unicode editor." In IEEE International Multi Topic Conference, 2001.

[3] Md. Akhtaruzzaman, "Unicode Searching Algorithm Using Multilevel Binary Tree Applied on Bangla Unicode." Innovations and Advanced Techniques in Computer and Information Sciences and Engineering. Springer Netherlands, pp. 321-326, 2007.

[4] Y. Song, et al. "Unicode-proof code injection attack on Windows CE—A novel approach of evading intrusion detection system for mobile network." IEEE 3rd International Conference on

Communication Software and Networks (ICCSN), 2011.

[5] J-W.Lin, and F-S Lin. "An auxiliary unicode Han character lookup service based on glyph shape similarity." 11th IEEE International Symposium on Communications and Information Technologies (ISCIT), 2011.

[6] G.Bo, and X.Qiang. "A Solution for Developing International Software Based on Unicode." 5th IEEE International Conference on. Intelligent Systems Design and Engineering Applications (ISDEA), 2014 .

[7] B.Pandey, G.Singh "Simulation of HSTL IO standard based energy efficient Punjabi Unicode reader on FPGA" IEEE International Conference on Computational Intelligence and Communication Networks,(CICN) 2014.

[8] M.P.Fenwick, and B.Simon. "Compression of Unicode Files." Data Compression Conference. 1998.

[9] S-S. Mohammad, and S.S. Sajad, "Persian/arabic unicode text steganography." IEEE International Conference on Information Assurance and Security, 2008. ISIAS'08.

[10] Q. Bite, N. Fang, and W. Liu. "Detect Visual Spoofing in Unicode-Based Text." ICPR. 2010.

[11] F.J.Mabry, R. J. John., and J. F. Aaron "Unicode steganographic exploits: maintaining enterprise border security." IEEE Security & Privacy, Vol.5, No.5, pp32-39, 2007.

[12] C. G. Jena, D.G. Tirthankar , and B. Anupam . "Implementation of Unicode Complaint Odia Keyboard and Its Evaluation Using Cognitive Model." IEEE International Conference on Information Technology (ICIT), 2014.

[13] H. Suryavanshi, and P. Bansal. "An improved cryptographic algorithm using UNICODE and universal colors." IEEE Ninth International Conference on Wireless and Optical Communications Networks (WOCN), 2012.

[14] V. Geethakumary, and B. A. Sharada. "Malayalam speech sounds and their mapping to Unicode symbols: a case study." Language Engineering Conference, IEEE Proceedings, 2002.

[15] Md. A. Rahman, and Md A. Sattar. "A new approach to sort Unicode Bengali text." In IEEE International Conference on. Electrical and Computer Engineering, 2008. ICECE 2008.

[16] R. Jaiswal, and N. Patil. "Implementation of a new technique for web document protection using unicode." IEEE International Conference on Information Communication and Embedded Systems (ICICES), 2013.

[17] K. Ubul, et al. "The Design and Realization of Unicode Based Minorities Character Online Processing System." In Eighth IEEE/ACIS International Conference on Computer and Information Science, 2009. ICIS 2009.

[18] A. Uddin , et al. "Thermal aware energy efficient Bengali unicode reader in Text analysis." IEEE International Conference on. Optimization, Reliabilty, and Information Technology (ICROIT), 2014.

[19] D. Bhattacharyya, et al. "A Novel Approach for Designing Indian Regional Language Based Raw-Text Extractor and Unicode Font-Mapping Tool. In IEEE International e-Conference on Advanced Science and Technology, 2009. AST'09.

[20] L. M Palanivelu, et al. "'n'—Level UNICODE position character length ciphers for securing smart cards." IEEE International Conference on Computing Communication and Networking Technologies (ICCCNT), 2010.

[21] S.Madhok, B.Pandey, T.Madhok "Design Goal Based Energy Efficient Greek Unicode Reader for Natural Language Processing on 28nm FPGA", International Journal of Energy for a Clean Environment, SCOPUS Indexed (Communicated).

[22] T.Das, et al, "Low Power Devnagari Unicode Checker Design Using CGVS Approach", International Conference on Recent Advances in Mechanical Engineering and Interdisciplinary Developments (ICRAMID), 07-08 March 2014, Kanyakumari, India. (Upgraded to AMR Journal).

[23] A.Kaur, et al, "SSTL IO Standard Based Tera Hertz & Energy Efficient MALAYALAM Unicode Reader

Design and Implementation on FPGA", In Special Edition of "Managing Information Technology", Vol.6, No.48, February, 2015.

[24] N.Kaur et al, "LVCMOS IO Standard Based Low Power Devanagari Unicode Reader on FPGA", In Special Edition of "Managing Information Technology", Vol.6, No.39, February, 2015.

[25] S.Madhok, B.Pandey, T.Madhok "Clock Gating Based Energy Efficient and Thermal Aware Design Latin Unicode Reader for Natural Language Processing on 28nm and 40nm FPGA", International Journal of Energy for a Clean Environment, SCOPUS Indexed (Communicated).

[26] S.Gargrish et al, "IO Standard Based Energy Efficient Devanagari Unicode Reader on 28 nm FPGA", In Special Edition of "Managing Information Technology", Vol.6, No.39, February, 2015.

[27] A.Kaur, et al, "Thermal aware energy efficient Gurmukhi Unicode Reader for Natural Language ", 9th INDIACom IEEE Conference , 11th - 13th March, 2015.

[28] A.Kaur, et al, "Capacitance Scaling Based Gurmukhi Unicode Reader Design for Natural Language Processing", ", 9th INDIACom IEEE Conference , 11th - 13th March, 2015.