RESEARCHERID
THOMSON REUTERS
[AIVESC-18]
IC™ Value: 3.00
ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

VHDL IMPLEMENTATION AND TESTING OF DATA COMPRESSION ALGORITHM IN DIGITAL COMMUNICATION

Amal Anilkumar[1], Mrs. Sarika K. T.[2], Divya Devadas[3], Ann Babu Treeza[4] and Ajeesh K.S[5]

[1]Department of Electronics & Communication Engineering, Vidya Academy of Science & Technology,Thrissur, India
[2]Assistant Professor , Department of Electronics & Communication Engineering, Vidya Academy of Science & Technology, Thrissur, India
[3]Department of Electronics & Communication Engineering, Vidya Academy of Science & Technology, Thrissur, India
[4]Department of Electronics & Communication Engineering, Vidya Academy of Science & Technology Thrissur, India
[5]Department of Electronics & Communication Engineering, Vidya Academy of Science & Technology Thrissur, India

DOI: 10.5281/zenodo.1227209

ABSTRACT

While transmitting data in digital communication, data bits should be as minimum as possible. There are several methods to compress the data. Source coding theorem says that the source encoder is said to be efficient when coding efficiency approaches unity. Minimum value of code length can be found using source coding theorem. In source coding, a prefix code is constructed based on the set of symbol and their probabilities or frequency. This enables faster writing and reading, less disk space, variable dynamic range and faster file transfer. Our project is aimed to compress data through VHDL implementation using Shannon- Fano and Huffman algorithm, their average code word length is determined and the efficient one among them are used in communication. Hence bandwidth available is utilized more efficiently. Shannon-Fano and Huffman is used in an implode compression method which are used in zip of rar format. The project is an effort to find out the optimum code for data compression. Synthesis and simulation of the source coding theorem is done using VHDL programming on software Xilinx ISE 14.2, simulated using Modelsim SE6.3f.

KEYWORDS: VHDL, SHANNON FANO, HUFFMAN

I. INTRODUCTION

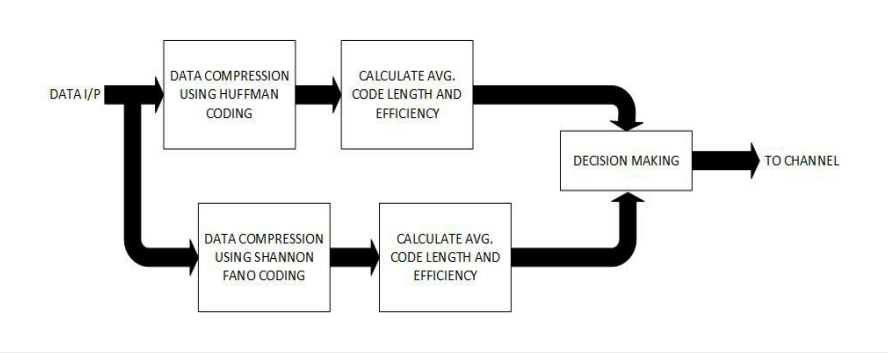
Data compression is the science of representing information in a compact form. It has been one of the critical enabling technologies for the ongoing digital multimedia revolution for decades. Without compression techniques none of the ever growing Internet digital TV mobile communication or increasing video communication would have been practical developments.

In telecommunication, a communication system is a collection of communication networks, transmission systems, relay stations etc. A communication system mainly comprises of a transmitter section that transmits the message to be transferred and a receiver section that receives the message, which are interconnected by means of a suitable channel. Such a system is said to be efficient if the data received at the receiver is exactly what was transmitted. But it is a well-known fact that, the message transmitted can get distorted due to several disturbances in its path. Nowadays, as demand is continuously increasing for development of reliable telecommunication and wireless systems, it is important to detect and correct errors in the information received through communication channels. In digital communication, the transmission of information is done in digital or encoded form. That is, we encrypt the signal for the added security to our information signal being transmitted.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Digital communication system is a system which makes use of binary 1's and 0's for the transmission and reception. In order to reduce the occurrence of error in digital communication system, the message signal will be coded using an encoder. At the receiver section, a decoder is used for the purpose of decoding the signal. The coding concept was introduced for the detection of error in the data. Actually coding theory is the branch of mathematics that is concerned with accurate and efficient transfer of data across noisy channels. The primary goal of coding theory is the encoding of information, easy transmission of encoded messages, fast decoding of received information and correction of errors introduced in the channel.

The transmitter section consists of source encoder and channel encoder and the receiver section comprises of channel decoder and source decoder connected using a channel. Message sources are the origins of message or input signal. A source encoder is used for the purpose of converting symbols to binary representation. The output of source encoder is fed to channel encoder. The channel encoder converts message into code words by adding parity digits for the purpose of error free transmission through channel. This codeword is then transmitted through the channel. A communication channel is simply referring to the medium by which a signal travels. There are different types of channels according to the type of transmission like optical fibers, coaxial cables or air in case of wireless communication. It is in this phase of communication that most of the error can occur. Errors can be of two types - single bit error and burst error. Single bit error can be caused by inverting a single bit and more than one bit inversion can be indicated by burst errors. Noise can get interfered with the message in the channel. The signal that is passed through the communication channel must be effectively captured by a receiver. The codeword transmitted through the channel is first subjected to channel decoding. Channel decoder obtains the transmitted codeword and this is then given to source decoder. Source decoder will give out the transmitted message. Source coding can be done using different methodologies like Huffman and Shannon-Fano coding.



In this project, the main focus of the work is on the source encoding. The most commonly used channel coding techniques are Shannon fano code and Huffman code which are having importance in real time applications. The main objective of the project is to implement both codes and determine which is more efficient, are programmed using Vhdl language. Synthesis and simulation of the source coding theorem is done using VHDL programming on software Xilinx ISE 14.2, simulated using Modelsim SE6.3f

II. METHODOLOGY

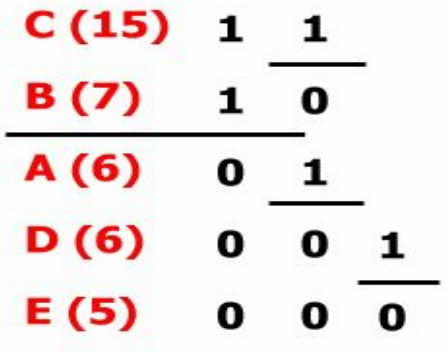
In this project mainly concentrated in the source encoding. Source coding is the process of removing redundancy from source symbols, which reduces data size.

Source coding theorem: The source encoder is said to be efficient when coding efficiency approaches unity. Minimum value of code length can be found using source coding theorem. Source coding is an important part of any communication system as it helps to use disk space and transmission bandwidth efficiently. Source coding can be either be lossy or lossless. It codes data to more efficiently represent the information. In case of lossless encoding, error free reconstruction of source symbols is possible whereas exact reconstruction of source symbols is not possible in case of lossy encoding. Source coding theorem is called noiseless encoding theorem as it establishes an error free encoding. It is also called Shannon's first theorem.The most commonly used source coding techniques are Shannon-Fano code and Huffman code which are having importance in real time applications. The main objective of our project is to implement a system to compare efficiency of both data encoded by Shannon-fano and Huffman, transmit the more efficiently encoded data. Codes are programmed using VHDL language and are simulated using Xilinx ISE 14.2 and ModelSim SE 6.4 simulator.

In today's world of digitalization, all the data which is to be processed of transmit or receive that should contain memory or bits as minimum as possible. So to reduce the bits or compress the data there are several technique,

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

one of which technique is Shannon-Fano algorithm in which we construct a prefix code based on the set of symbol and the probabilities of frequency.

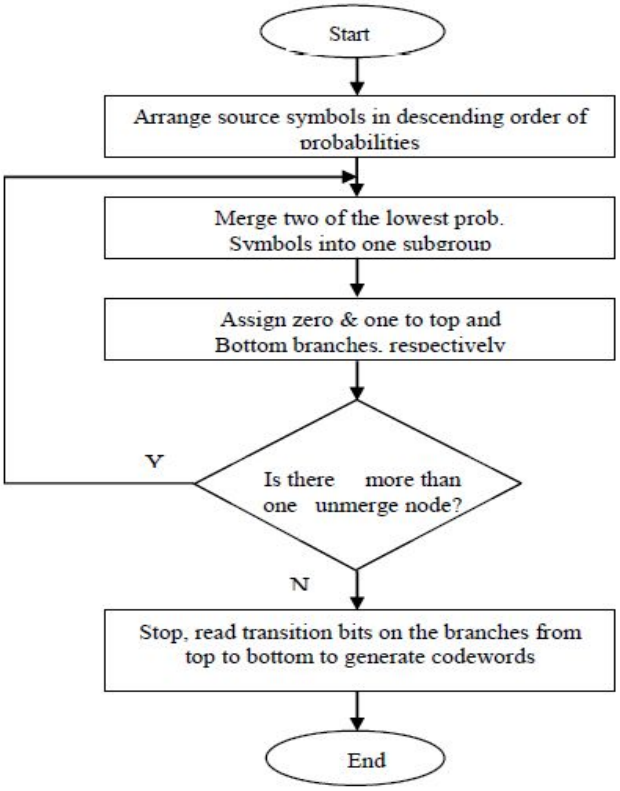


In Shannon-Fano algorithm sorting of different symbol is done according their frequency or occurrence. The sorted symbols with their occurrence are divided into two parts according to their probabilities. The first part is the highest probable symbol and second part is the remaining symbols and this remaining symbols are again divided into two parts, one part is second highest probable symbol ad second part is remaining other symbols, in this way all the symbols are separated and after dividing all the symbols separately, now we ready to assign bits to this sequence. In which the first part which is divided into two parts in which the highest probable symbols assign with a single bit '0' ad remaining part assign to a bit one. Similarly in the second iteration the second highest probable symbol are again assign to '0' and remaining parts are assigned to '1' in this way bits are assigned to this tree. While encoding to each symbol we follow the path from top to respective symbols and take all the bits in that path to generate a proper encoded bits for that symbol. So now we can observe that first highest probable symbol assign with '0', second highest probable symbol encoded to '10' third highest probable symbol encoded to '110', fourth highest symbol encoded to '1110' in this way we encoded all the symbol of the given data to reduce the bits and finally we can verify by counting all the bits of the encoded symbol that our resultant data gets compressed. Huffman coding is another source encoding technique. In this project we will compare the data encoded by both the technique and transmit the more efficiently compressed data. A minimal variable-length character coding based on the frequency of each character. First, each character becomes a trivial binary tree, with the character as the only node. The characters frequency is the tree's frequency. Two trees with the least frequencies are joined as the sub trees of a new root that is assigned the sum of their frequencies. This is repeated until all characters are in one tree. One code bit represents each level. Thus more frequent characters are near the root and the coded with few bits, and rare characters are far from the root and are coded with many bits. Also known as static Huffman coding. Huffman's algorithm provided the first solution to the problem of constructing minimum-redundancy codes.

The input data is given to both the source encoder simultaneously. In the source encoding section both the data gets compressed according to their frequencies or occurrences.In Shannon-Fano algorithm sorting of the different symbol is done according to their frequency or occurrence.Symbols with their occurrence are divided in two parts according to their probabilities. Then in this sorted symbols the first part is the highest probable symbol and second part is the remaining symbols, and now this remaining symbols are again divided in two parts one part is 2nd highest probable symbol and second part is remaining others symbols.In each sorting highest probable symbol is assigned '1' and the remaining symbols assigned '0'.and the codes are read after the final sorting.

In Huffman coding variable-length character coding based on the frequency of each character. First each character becomes a trivial binary tree, with the character as the only node. The characters frequency is the tree's frequency. Two trees with the least frequencies are joined as the sub trees of a new root that is assigned the sum of their frequencies. This is repeated until all characters are in one tree. One code bit represents each level. Thus more frequency characters are near the root and coded with few bits, and rare characters are far from the root and are coded with many bits. Also known as static Huffman coding. Huffman's algorithm provided the first solution to the problem of constructing minimum redundancy codes

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7



After data compression using both the algorithms, the compressed data then given to the next block where its code length and efficiency both calculated using the equations described earlier. After their efficiencies are calculated, in the next block, efficiency of both compressed data are compared. Then in the next block decision is made by selecting the more efficiently compressed data. The entire program is written using VHDL language. It stands for VHSIC Hardware Description Language. This language commonly used to write text models that describe a logic circuit. Such a model is processed by a synthesis program, only if it is part of the logic design.

III. SOFTWARE DETAILS

Xilinx ISE(Integrated Synthesis Environment) is a software tool produced by Xilinx for synthesis and analysis of HDL designs, enabling the developer to synthesize (”compile”)their designs, perform timing analysis, examine RTL diagrams, simulate a design’s reaction to different stimuli, and configure the target device with the programmer. Xilinx ISE is a design environment for FPGA products from Xilinx, and is tightly-coupled to the architecture of such chips, and cannot be used with FPGA products from other vendors. The Xilinx ISE is primarily used for circuit synthesis and design, while ISIM or the ModelSim logic simulator is used for system level testing. Other components shipped with the Xilinx ISE include the Embedded Development Kit (EDK), a Software Development Kit (SDK) and ChipScope Pro.

Very High Speed IC Hardware Description Language is commonly used to write text models that describe logic circuits. It is used in electronic design automation to describe digital and mixed-signal systems such as field-programmable gate arrays and integrated circuits. This hardware description is used to configure a Programmable logic device (PLD), such as a field programmable gate Array (FPGA), with a custom logic design. The general format of a VHDL program is built around the concept of BLOCKS which are the basic building units of a VHDL design. Within these design blocks a logic circuit of function can be easily described. A VHDL design begins with an ENTITY block that describes the interface for the design. The interface defines the input and output l1ogic signals of the circuit being designed. The ARCHITECTURE block describes the internal operation of the design. Within these blocks are numerous other functional blocks used to build the design elements of the logic circuit being created. After the design is created, it can be simulated and synthesized to check its logical operation. SIMULATION is a bare bones type of test to see if the basic logic works according to design and concept.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

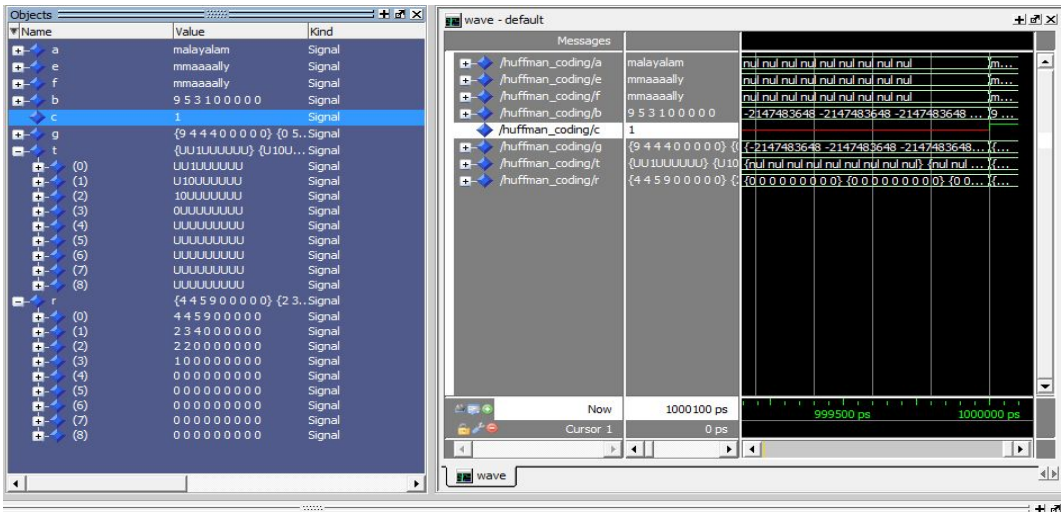
Xilinx ISE 14.2:-The ISE Text Editor allows to create, view, and edit text files, such as ASCII, UCF, VHDL, Verilog, and TCL files. Use the software to do the following
Create files using features such as auto-completion of text, file insertion, and pre built code from the ISE Design Suite Language Templates. Edit files using commands to indent, convert, and comment text. Navigate files using bookmarks and Go To commands. Find and replace text. ModelSim SE}
\qquad ModelSim is a Multi-language HDL simulation environment by mentor graphics for simulation of hardware description languages such as VHDL, Verilog and system C and includes a built in C-debugger. ModelSim can be used independently, or in conjunction with Intel Quartus Prime or Xilinx ISE/Vivado. Simulation is performed using the graphical user interface or automatically using scripts.

IV. RESULTS AND DISCUSSION

SCREENSHOT 1. SHANNON FANO OUTPUT

The Vhdl program for Shannon Fano and Huffman algorithm was executed using the Xilinx software and the result is obtained in ModelSim. Here the word Malayalam is used for data compression. Above Figure shows the simulation result for Shannon fano algorithm. In the program we mentioned 'a' as the input message. As per given in the program, our input message is a nine letter word, here we obtained the results for word 'Malayalam'. Variable 'b' mention another array in which occurrences or probabilities of the characters are stored.

Variable 'c' is the control signal if '1' the below mentioned programs will run or if '0' jump to the next section. ‘e’' is the probabilities mentioned in this array, here since the total length as per the given program is nine, the probabilities obtained by dividing the occurrence of each character by nine. Obtained the result or the encoded binary data in this 'g' array which is two dimensional array. Below figure shows the simulation result of Huffman algorithm.



V. CONCLUSION

From the observation it can be found that the data compression using Shannon-fano and Huffman algorithm has been very successfully implemented using VHDL coding using ModelSim SE 6.4 simulator and the data gets compressed using these technique. The project helps to represent the data with minimum no of bits to save the bandwidth. Compress the data using Shannon-Fano and Huffman algorithm. Compare the efficiency of each compressed data, then decide and transmit more efficiently compressed data. This method enable less disk space, variable dynamic range, faster file transfer, faster writing and reading.

VI. ACKNOWLEDGEMENTS

During the course of project work several persons collaborated directly and indirectly with us. Without their support it would be impossible for us to finish our work. That is why we wish to dedicate this section to recognize their support. We are thankful to our project coordinator and to our project guide for their valuable advice and guidance towards this work. We received motivation, encouragement and hold up from them during the course of work. We are grateful to express our thanks to all the faculty members of our department for their support.

VII.REFERENCES

[1] *M. Vaidya and E. S. Walia, "Data compression using Shannon-Fano algorithm implemented by VHDL," 2014 International Conference on Advances in Engineering \& Technology Research (ICAETR - 2014), Unnao, 2014*

[2] *G. E. Atkin, B. Patel and J. Rahhal, "Combined source coding and modulation for fading channels," IEEE 43rd Vehicular technology Conference, Secaucus, NJ, 1993, pp. 491-494*

[3] *N. Dhawale, "Implementation of Huffman algorithm and study for optimization," 2014 International Conference on Advances in Communication and Computing Technologies (ICACACT 2014), Mumbai, 2014, pp. 1-6.*

[4] *K. A. Babu and V. S. Kumar, "Implementation of data compression using Huffman coding," 2010 International Conference on Methods and Models in Computer Science (ICM2CS-2010), New Delhi, 2010, pp. 70-75.*

[5] *Alka P. Sawlikar, Z. J. Khan, S. G. Akojwar, "Power Optimization of Wireless Sensor Networks Using Encryption and Compression Techniques", Electronic Systems Signal Processing and Computing Technologies (ICESC) 2014 International Conference on, pp. 222-227, 2014.*

[6] *Srikanth Immareddy, Sravan Kumar Talusani, "Multi level Huffman test data compression with multiple scan chain method", Computer and Communications Technologies (ICCCT) 2014 International Conference on, pp. 1-6, 2014*