Video Data Compression using Discrete Wavelet Transform

Pinak Desai[1], Dr. P. B. Mane[2]

[1]*AISSMS IOIT, PUNE*
[2]*Principal, Member, IEEE, AISSMS IOIT, PUNE*

Abstract - *The data we deal today is of significantly higher size. This further gives rise in difficulties regarding storage and transmission. So a need for compression was felt which could reduce the data size in a way aiding us in efficient storage and compression rates. In this paper video data compression using the DWT (Discrete Wavelet Transform) based wavelet morphology algorithm is proposed. DWT separates the original image into pixels and with wavelets a mathematical function is used in signal processing and compression. The proposed system deals in video data compression with image, frame by frame via the pixels resulting in better memory efficiency by reducing the file size with efficient hardware solution of low cost and power to meet the expected real time constraints.*

Key Words: **DWT, wavelet, morphology, frame, pixel**.

1. INTRODUCTION

In the modern 21st century we are surrounded by data. Data can be in various types like sound, image, video etc. This data can be stored and transmitted accordingly in a way increasing the need and demand of the data communication and multimedia storage. The revolutionary changes in the digital world has been magnificent and the videos are stores on drives, disks. Star Wars was the first movie which was shot digitally in the early 2000's while the todays digital cameras can easily captures images in the 4K resolutions which is indeed breath taking. Thus further increasing the storage requirements.

Videos are nothing but moving images. These series of image are called frames. These videos are normally in the full high definition of resolution. A video has a raw image resolution of 1920 x 1080 pixels per frame. That is over 2 million pixels per frame of the video. Now the video is running at 30 frames per second which brings the total at 62 million pixels per second. That is a huge amount of pixels being dealt. We roughly need around 177 Megabits of data per second to access the file. Now if the video is of 5 minutes duration we will be consuming a mammoth 51 Gigabytes of data. Thus a DVD would contain just a single second of this raw data. Imagine the time taken to download the data packets through the internet just to access the file. The latest generation mobile technologies will not be able to handle the data rates thus we wont be able to access content from our smartphones. As no one has an internet connection which can deliver such high speeds, bandwidth. This will be the case if we are dealing with the raw data.

Image frames consist of all the basic and raw multimedia data mostly taking up the bandwidth required for the data communication. So it is important to reduce this data with the help of different measures as the video on demand service has great requisition. So compression of this video data was required for reducing the total size of the data making it possible for us to access the content via the internet according to the available throughput rates [8] , [9] , [10].

The first step to compress this image data was taken by considering the nearby neighbouring pixels of the image frame. This pixels could be correlated to each other and the information could be redundant. In a way giving more importance to the pixels with less correlation. Thus irrelevancy and redundancy are the two basic pillars in the field of data compression [13] , [14], [15]. Redundancy deals with removing the slightly less important data which is no longer needed from the data source while irrelevancy removes the less relevant data like the pixels which are not visible to us [13]. The lossless standard helps in removing the redundancy from the original raw data and the data is reverse engineered at the receiver end. So in current scenarios the lossless standards can only reach the average compression rates while the more practical approach of lossy standard provided us with improved compression rates but during reverse engineering at the decoder the exact original information of data was not found. The research in the data compression started way back in the 80's with the CCITT standards [7]. Here binary image compressions were used. Further giving rise to different standards like JPEG, GIF. JPEG are widely used techniques for image data compression.

Image data compression standards provides us with the following benefits

1) Ease in data exchange between devices
2) Reusing hardware and software for different applications.
3) Providing benchmarks and conclusive evidence for new and alternative measures.

The most critically acclaimed standards are derived from the Discrete Cosine Transform (DCT) [7], [8]. DCT is the most widely used data compression technique. But it also has some major issues like the loss in the image quality where the image resolution is affected when we try to achieve high data compression rates by using the DCT techniques. Further research gave us conclusive evidences about the wavelets being a major leap in the field of data compression with the help of its very own mathematical approach [9]. Wavelets successfully blocked the still imaging of artefacts which were a major issue with the DCT compression technique. Thus the Discrete Wavelet Transform (DWT) formed the core in the evergreen data compression standards like the Joint Picture Expert Group (JPEG) and the JPEG 2000.

For addressing the issue of memory, the real time video processing DWT architecture was formed, implemented and further analysed in the FPGA design flow. At the centre core of JPEG 2000 architecture is a compression technique based on wavelet. This in turn had several number of advantages over the conventional Discrete Cosine Transform [7], [8], [15]. In the DCT compression we expect blockiness in combination with the compressed video. The DWT technique also have some rough edges which results in better picture rendering. Along the years many new types of algorithms with some mathematical approaches were used for analysis and processing of our signals. Keeping in regard our signal compression domain the wavelets have regained interests in research as well as practical real time applications. Majority of the real time signals like video and music are very complex in nature with nonperiodic functions and features. So extracting and analyzing measures for monitoring or recognition of these types of signals have numerous practical consequences. To compress a video file using the DWT technique by hitting the desired 1:43 compression ratio.

A codec is a software/circuitry which can compress/decompress the original raw video data. Encoder compress and decoder decompress the data. A lossy compression where some information is absent in comparison with original video. So when this data is decompressed quality degrades due to the absence of information in reconstruction of video. Codecs are used in playbacks, recordings and many other applications. A codec picks a key frame which acts as a reference frame to others. The H.261 and H.263 were two big standards which were widely used and internationally accepted now it's the H.264 standard which has taken the improvements to another level in a way making it a favourite in high definition compression, recording and distribution world. Hierarchical storage and reuse of memory without influencing performance [21], [22].

HCC (Hierarchical Configuration Context) and LDO (Lifetime based Data memory Organization) are the two approaches considered. HCC's hierarchy terminates portion of context helping in storage reduction but affecting overhead transport with fast configurations [20]. LDO's on chip data is of two types first the lifetime of data is considered so the data having a shorter lifetime is accumulated in FIFO which increases the reuse of memory which then passes the value from one stage to next and is stored in the RAM. The data of previous frames is stored for intra and inter prediction. Prediction is done by the H.264 encoder with DWT transform to obtain compressed H.264 bitstream. Logic element of 6869 having speeds of 153 MHz with power dissipation of f 515 mW is mainly used by HCC and LDO.

2. LITERATURE SURVEY

[1] Most recent technologically advanced hybrid approach of Convolution Neural Network CNN was used with JPEG compression. It shows potential for future considering its parallel style and adjustment to intrinsic nature.

[2] Compression of an image in light ray field with focusing was done with help of plenoptic camera in HEVC compression which resulted in improved focusing of shiny, bright objects.

[3] Fuzzy logic in real time were used to increase compression ratio and quality over bit stream generation.

[4] A new Integer Discrete Cosine Transform algorithm helped in reducing the cost, power consumption by minimizing some hardware components of circuit area in FPGA's and ASIC's.

[5] A block partitioning structure using Quadtree plus binary tree block structure coding was used to obtain two different sets of tertiary tree split mode and asymmetric binary tree split mode partitioning. Corner boundary fittings were improved.

[6] Wavelet based foveated compression algorithm for real time processing was used to reduce overhalls in communication.

[7] A three dimensional discrete cosine transform in real time was used for video compression.

[8] Quantization of image was achieved with discrete cosine transform method in JPEG standard with the help of pixels.

[9] A tree based approach with tilling of raw data was used for compression. This mathematical approach resulted in improved scanning with faster calculations.

[10] Parallel implementation of FPGA using CCSDS 123 compression algorithm was used.

[11] A two dimensional embedded zero wavelet and three dimensional discrete wavelet transform were compared with help of parameters like PSNR, compression ratio.

[12] A modified three dimensional discrete cosine transform method was proposed. Variable length parallel pipes were used to replace standard fixed 8x8x8 size video

cubes which showed improvements compared to typical three dimensional DCT.

[13] Lifting based illumination adaptive based transform method with classic wavelet video coding system of t+2D was used. Illumination in rate distortion were observed.

3. METHODOLOGY

Wavelets form the core of our Discrete Wavelet Transform based system. Wavelets are mathematical functions which can divide a given function in our case a continuous time signal in the form of video data which is further scaled accordingly. Every obtained scaled component is considered further with the resolution which matches its scaled version. Thus these wavelets determine the performance of our data compression technique.

The morphology transform algorithm uses a Discrete Wavelet Transform approach.. The most important features are that it will be mostly benefitting us in our approaches it consists of a faster and better computation architecture and has one of the best memory allocation method to increase efficiency. Once the DWT algorithm is applied at the encoder end for the proper encoding of the raw uncompressed data we can easily reverse this at the decoder end by applying the Inverse DWT transform algorithm without generating further problems.

In our proposed system a very high speed, reliable and efficient DWT to IDWT processor engine is expected in the implementation resulting in low powers as well. This is implemented on the Spartan 6 Field Programmable Gate Array (FPGA). Spartan 6 acts as a coprocessor for the compression and decompression process. Video compression with the help of FPGA is one of the major and reliable procedure of memory saving into the required datapath locations.

Figure 1 shows the block diagram of the proposed work. As shown in the figure the real time video is being recorded by the camera module. This camera is interfaced with our Spartan 6 FPGA. The OV 7670 CMOS video sensor module is very small in size, highly sensitive and requires low voltage. The threshold number is provided for video capturing.

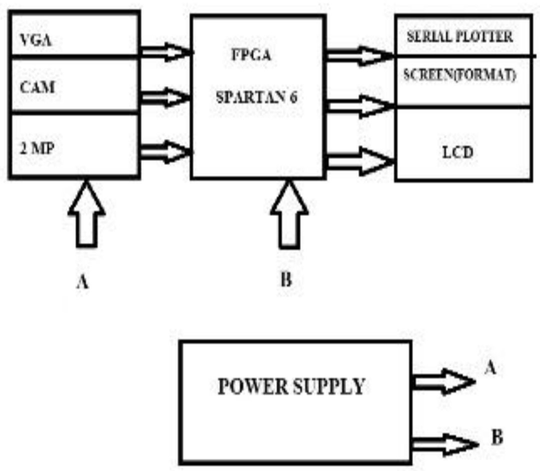

Fig -1: Proposed Block Diagram

The frames get captured till the threshold number and camera stops the recording. This process has been done in the Vivado MATLAB with the help of SCCB Bus Control acting as serial plotter. It outputs the whole frame, sampling and resolution of data. The VGA system can reach a maximum of standard 30 fps. Image frames are extracted from the recorded video and pre-processing is applied on it. It simplifies the complexities that have been occurred with the coloured image. The aim of pre-processing is an improvement of the data images that suppresses redundant distortions or enhances some image features that are required for further processing. Header file of every extracted image frame is done, in which all row's and all column's pixel values are stored, which is required in the main compression part.

3.1 Image Compression using DWT

JPEG stands for the Joint Photographic Experts Group, a standards committee that had its origins within the International Standard Organization (ISO).JPEG provides a compression method that is capable of compressing continuous-tone image data with a pixel depth of 6 to 24 bits with reasonable speed and efficiency. JPEG may be adjusted to produce very small, compressed images that are of relatively poor quality in appearance but still suitable for many applications. Conversely, JPEG is capable of producing very high-quality compressed images that are still far smaller than the original uncompressed data.

JPEG is primarily a lossy method of compression. JPEG was designed specifically to discard information that the human eye cannot easily see. Slight changes in color are not perceived well by the human eye, while slight changes in intensity (light and dark) are. Therefore JPEG's lossy encoding tends to be more frugal with the gray-scale part of an image and to be more frivolous with the color.

Here we are targeting a 30sec video. This will be a coloured video. The frames per second of the video will be 30fps. The original resolution of the video will be of 640 x 480. That is 297,600 pixels per frame of the video. So the 30fps video will have around 8,928,000 pixels. In this way we are dealing with almost 9 million pixels per second.

3.2 Steps included

1) Video is divided into frames. (Grouping)
2) A single frame is considered from n number of frames. (Single frame)
3) Resolution of the frame is represented in X x Y format
4) Thus a single image is generated using the resolution and stored as reference in RAM.
5) Pixels are extracted from this image and the difference between the frames is predicted with motion estimation which estimates the residual value obtained from DWT.
6) A deblocking filter is used to better the quality and performance by smoothening sharp edges baetween macro blocks. R, G, B values of each pixel is calculated wrt DWT prog. (Range 0-255)
7) These important values are further processed using DWT technique. (Formula equation)
8) The image is compressed upto a compression ratio of 1:43
9) Lastly IDWT is applied to obtain final compressed output. (Step 8 to 1 reversed).

4. RESULTS

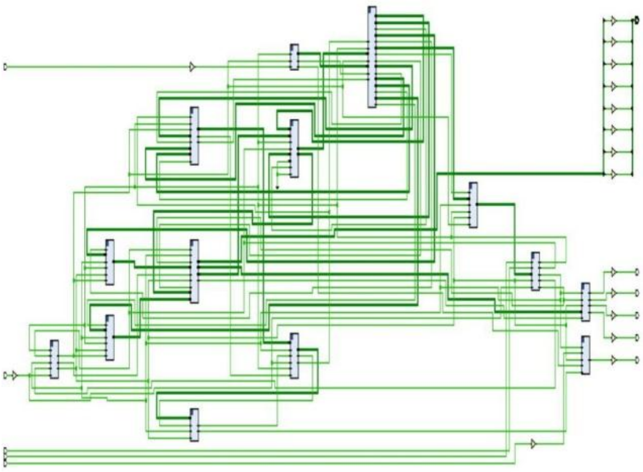

Fig -2: Elaborated Architecture

The DWT will provide us a multi- resolution representation of images. This process highlights on the simple to work with but very fast implementation of the DWT and the IDWT algorithms with the help of system C coding on Vivado XILINX environment. The figure shows elaborated architecture of system in software.

In a way the R, G, B values are quantized by the 'Sampling Quantization' method and iteration bits are developed in our case a total of 10 iteration bits as shown. Each iteration bits generating a case accordingly. In our case we will have a total of 10 cases to deal. So ideally each case will undergo the DWT algorithm and the size will be compressed of the now active signal. Complexity is dealt with precision and boosting the computational speed can be done. A two dimensional DWT will be performed on images with varying pixel sizes. The algorithm flow System parameters like the peak signal to noise ratio can also be measured. From the generated simulation results we can conclude that algorithm will be working as per our requirements. We wrap that up for obtaining the higher compression values and further higher decomposition is needed.

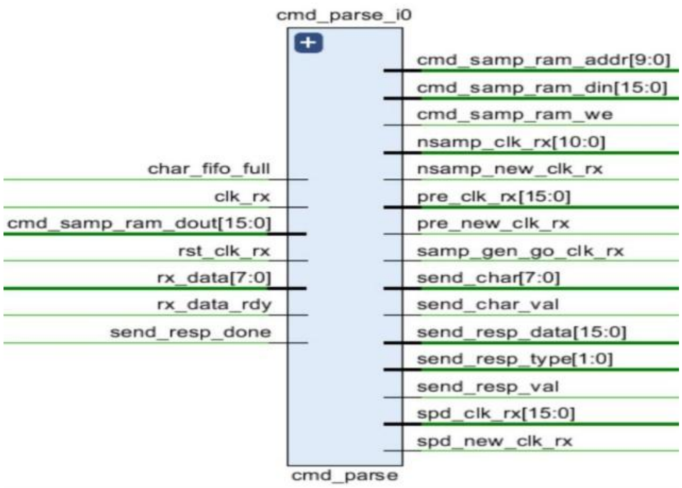

Fig -3: Single Command Design

Fig -4: Methodological Result

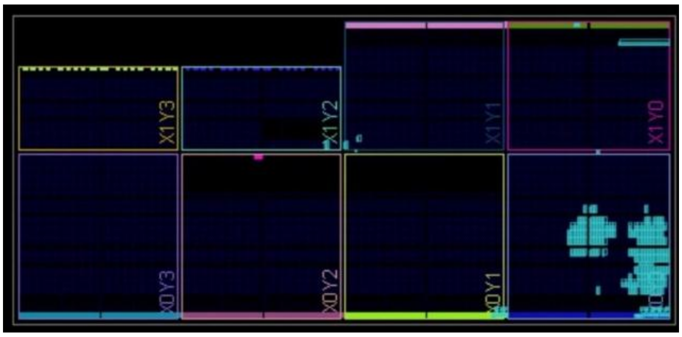

Fig -5: Graphical Result

5. CONCLUSIONS

Implementation of system using VLSI i.e. FPGA approach provides us the benefits of low memory requirement and low power consumption. The original video file size (ex- 8-10 MB) could be compressed to as low as possible (1-2 MB). The programmer can control all the aspects of the video file like image, pixels, bitrate. Thus a new and better approach prototype of data compression is expected.

REFERENCES

[1] Siwei Ma et al., Image and Video Compression with Neural Networks, IEEE, September 2018.

[2] Cristian Perra, Daniele Guisto, An analysis of HEVC compression for light field image refocusing application, IEEE, 2018.

[3] Upendra Kumar Srivastava, Navin Prakash, A Systematic Review on Real Time Video Compression and Enhancing Quality Using Fuzzy Logic, IJCSE, 2018.

[4] Jiajia Chen et al., Hardware Efficient Integer Discrete Cosine Transform for Efficient Image/Video Compression, IEEE, 2019.

[5] Fabrice Le Léannec et al., Highly Flexible Coding Structures for next generation video compression standards, IEEE Data Compression Conference, 2019 .

[6] J C Galan Hernandez et al., Wavelet based Foveated Compression Algorithm for Real Time Video Processing, IEEE, 2010.

[7] T Fryza et al., Introduction to implement real time video compression method, IEEE, 2008.

[8] Anitha S. et al., Image Compression Using Discrete Cosine Transform & Discrete Wavelet Transform, International Journal of Scientific & Engineering ResearchVolume2, Issue8, 2014.

[9] Kai-Lung Hua et al., Inter Frame Video Compression With Large Dictionaries of Tilings: Algorithms for Tiling Selection and Entropy Coding, IEEE, 2012.

[10] Milica Orlandic, Johan Fjeldvedt and Tor Arne Johansen , A Parallel FPGA Implementation of the CCSDS-123 Compression Algorithm , 2019.

[11] V.R. Satpute et al. , Comparative analysis of video compression mechanisms using 3D-DWT based video encoding along with EZW , 2017.

[12] Radu Radescu, An Efficient Solution for Video Compression Using an Original Modified Algorithm Applied to 3D Discrete Cosine Transform, November 2018.

[13] Maryam Haghighat et al., Lifting-based Illumination Adaptive Transform (LIAT) usimg mesh-based illumination modelling, IEEE, 2017

[14] Artem Gorodilov et al. ,Neural Networks for image and video compression, IEEE International Conference on Artificial Intelligence Applications and Innovations , 2018.

[15] Jaya Krishna Sunkara et al., A New Video Compression Method using DCT/DWT and SPIHT based on Accordion Representation.

[16] R. Mehala , Comparision of DCT & DWT in Image Compression Techniques , IJARTET 2016.

[17] Ken Cabeen and Peter Gent,"Image Compression and the Descrete Transform"Math 45, College of the Redwoods.

[18] A. S. Lewis and G. Knowles, "Image Compression Using the 2-D Wavelet Transform" IEEE Trans. On Image Processing, Vol. I. NO. 2, PP. 244 -250, April 2016.

[19] Ken Gaben and Peter Gent, Image Compression and the Discrete Transform, IEEE Trans. On Image Processing, Vol. I No. 2, PP. 244-250.

[20] Yansheng Wang et al., On Chip Memory Hierarchy in One Coarsed Grain Reconfigurable Architecture to Compress Memory Space and to Reduce Reconfiguration Time and Data Reference Time, May 2014.

[21] T. Geng et al., Parallelization of computing intensive tasks of H.264 high profile decoding on Reconfigurable multimedia system, Jan 2010.

[22] M. K. Ganesan et al., H. 264 decoder at HD resolution on a coarse grain Dynamically Reconfigurable Architecture, Aug 2007.

[23] Greg Ames, "Image Compression", Dec 07, 2002.