Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.6, No.4, 2015
www.iiste.org

Efficient Algorithms/Techniques on Discrete Wavelet Transformation for Video Compression: A Review

Surabhi K. Vyas (Corresponding author)
Student (VLSI Design), Mody University, Laxmangarh, Sikar
E-mail: survyas1991@gmail.com

Dr. K. G. Sharma
Faculty of Engineering, Mody University, Laxmangarh, Sikar
E-mail: shrama.kg@gmail.com

Mr. Dharmendra Kumar
Senior Design Engineer, SiliconMentor, Greater Noida, Utter Pradesh
E-mail: dharmendra@siliconmentor.com

Abstract
Visualization is the most effective and informative form for delivering any information. There are various techniques for video compression such as Motion Estimation and Compensation, Discrete Cosine Transformation, Discrete Wavelet Transformation etc. Wavelet transforms have been triumphant in high rates of compression as well as maintains good video/image quality. In this paper, the implementation of different algorithms of three dimensional wavelet transformations for video compression is presented.
Keywords: Video compression, Temporal decomposition, Discrete Wavelet Transform (DWT), 3D Wavelet Transform.

1. Introduction

As the technology has advanced so are the increasing demands of people. The area of visualization has extended to its best including areas of video capturing, storing, High Definition (HD) displaying and many more. The increasing feature in multimedia and use of storage devices such as Compact Disk (CD), Digital Versatile Disk (DVD) led to emphasize on its storage so that its processing can be done easily without any loss of information. Video is required in any of the stream which may be conveying information, in gaming, movie, extracting relevant information and one such application of real-time is the space application of video. For sending and receiving any real-time video, our major field which needs to be stressed upon is its processing with minimum consumption of memory capacity as possible. This lead to the very arising topic of re-searches which is 'Video Compression'. The challenging task in video compression is the intensity of data which is required for storage and transmission, this is quiet impractical. As there is limited bandwidth capacity, the data within this needs to be compressed. Various techniques have been employed for video compression. Some of them are Motion Estimation and Compensation, Block Matching Algorithm, Discrete Cosine Transformation (DCT), Vector Quantization (VQ), Fractal Compression (FC), Discrete Wavelet Transformation (DWT) etc.

In this paper, DWT technique has been used which has proven to be advantageous in comparison to DCT. The same process like any other technique is followed, where the transformation into the components of frequency of an image is done. DCT picks a small desirable data of an image and then performs further processing. Its result is represented in the form of hierarchy where each frequency band relates each layer [1]. As mentioned that DCT can be applied on a small part of an image and DWT can be applied on the entire image at the same time. So, DWT techniques can be extended to video schemes also.

2. Literature Review

In 2001, the proposed efficient algorithm of 3D wavelet transformation shows the better results in terms of low clock latency and lower memory demands using efficient way of temporal decomposition for video compression [2].

In 2006, the proposed 3D wavelet transformation algorithm was compared with 3D traditional algorithm and results exhibits lower memory demands and better performance of video [3].

In 2007, the proposed algorithm was used for real time applications by implementing it on FPGA to reduce the chip area and delay using FIR filters [4].

In 2013, lifting based scheme was used instead of Conventional convolution based DWT to reduce the computational complexity and memory requirements [5].

In 2014, video encoding techniques based two mechanism were compared with each other. As the level of Spatial DWT and Temporal DWT increases, efficient utilization of memory is required [6].

3. Related Work

1. E. Moyano, F.J. Quiles, 'A. Garrido, 2L. Orozco-1 Bardosa, 3J. Duato presented a paper in 2001 and found that the GOF based 3D-WT algorithm was less efficient because it required large memory space to process large group of frames at the same time. Maximum number of temporal decomposition levels were required to get the maximum number of high temporal frequency bands. As the numbers of decomposition levels increased large delay would be produced.

This proposed algorithm works by taking the first four frames from the original sequence and two low and two high frequency bands are provided as a resultant. After that, for the next temporal decomposition it takes two new frames from the original sequence and two low frequency bands from the result of last temporal decomposition step. So in the last step, we get all high temporal frequency bands except two i.e. low frequency bands [2]. Results were found by using an Intel Pentium III (650 MHz) processor and 128 MB of RAM for 'Trevor' video sequence, which has 64 frames and each of 512 × 512 pixels and 256 grey levels [2].

Advantages: This algorithm is more efficient than the GOF based 3D-WT algorithm. The main advantage of this algorithm is that it provides high temporal frequency bands by only one level of temporal decomposition. Also, no longer delay is produced as 4 frames are processed form the original sequence.

2. Yongjian Nian, Lehua Wu, Shibiao He and Yongjun Gu proposed a new encoding algorithm in 2006, the traditional 3D-WT algorithm was not suitable for many applications which require large GOF to get most temporal redundancies in input sequence without much motion and also large buffer requirements for large GOF.

This proposed algorithm is slightly different from the previous algorithm in terms of temporal decomposition. According to this algorithm video sequence of 16 frames is firstly divided into 4 groups of frames of one low temporal frequency band and three high temporal frequency bands. After the decomposition levels, one low frequency band and 15 high frequency bands H_1-H_{15} are provided, this is very easy to compress. This new algorithm requires small number of quantization steps because last decomposed low frequency band L_5 has lowest resolution and it is known that maximum information is available in L_5, which is very important for the Human eye. Experimental results are compared in terms of computational complexity and memory requirements [3]. When the temporal decomposition of 16 frames is done, then the complexity of multiplication in computing is given as

$$(p \times q - 1) \times (Ltaps + Htaps) \quad \text{..Ref [3]}$$

Where p × q is the frames in original video sequence, Ltaps indicates the number of low-pass filter coefficient and Htaps indicates the number of high-pass filter coefficient. Complexity of computing addition is given as

$$(p \times q - 1) \times (Ltaps \times Htaps - 2)/2 \quad \text{..Ref [3]}$$

After the calculations have done, it is shown that computational complexity is not different from the traditional algorithm [3].

Table -I Performance table for Miss American and Salesman Ref [3]

Miss American		**Salesman**	
CR	**PSNR(dB)**	**CR**	**PSNR(dB)**
80	39.1186	30	37.5756
135	38.0949	55	35.4486
250	36.0649	105	32.2637
530	33.3475	162	30.0485
800	31.8522	360	27.5647

Figure 1 and table-1 shows comparison results of compression Ratio (CR) and Average Peak signal to noise ratio PSNR (dB) for Miss American and Salesman frames of 352 × 288 sizes on luminance component. Total number of frames used is 80 but every time size of input sequence will be 16 frames. When CR=250 of 15^{th} frame of Miss American and CR=105 of 22^{nd} frame of Salesman, then quality of reconstructed frames (b) and (e) are shown below:

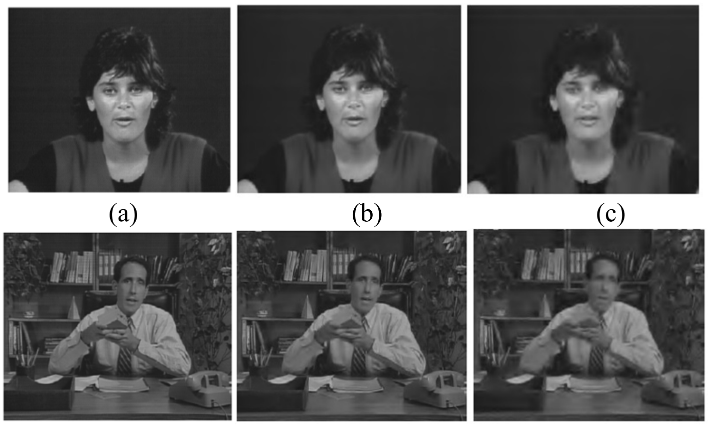

Figure-1(a) Original frame of Miss American. (b) Reconstructed Frame at CR= 135. (c) Reconstructed Frame at CR= 530. (d) Original Frame of Salesman. (e) Reconstructed frame at CR= 55. (f) Reconstructed frame at CR= 162. Image Courtesy Ref [3]

Advantages: When frame is reconstructed, this proposed algorithm provides the good degree of compression in Miss American frames as compared to Salesman frames. The reason behind it that there is much motion and detailed information in background of Salesman frames which is very difficult to compress.

3. In 2007, Samar Moustafa Ismail, Ali Ezzat Salama and Mohamed Fathy Abu-El Yazeed used hardware designing approach and FPGA implementation for novel 3D-V algorithm to make it usable for real-time video applications. This proposed algorithm used two filters Direct FIR filters and Transposed FIR filters are compared for designing the basic block of hardware by using pipelining approach [4]. Both filters have some trade-off and limitations according to their applications used. In this proposed work, basic block of design is 1D-DWT filter bank. Daubenchies-2 filter is used with 4 coefficients so Transposed FIR filter uses 4 symmetric taps. A delay element, a signed constant coefficient multiplier (KCM) and a register is carried out by the each tap of Transposed FIR filter structure [4]. By using these elements, redesigning of filter is not necessary. For the same process parameter is also a varying factor which may be used desirably [4].

Figure 2 shows the software simulation results of two FIR filters in terms of clock latency and chip area utilization. As the number of taps increases, latency of Direct FIR filter increases, but latency remains same in case of Transposed FIR filter.

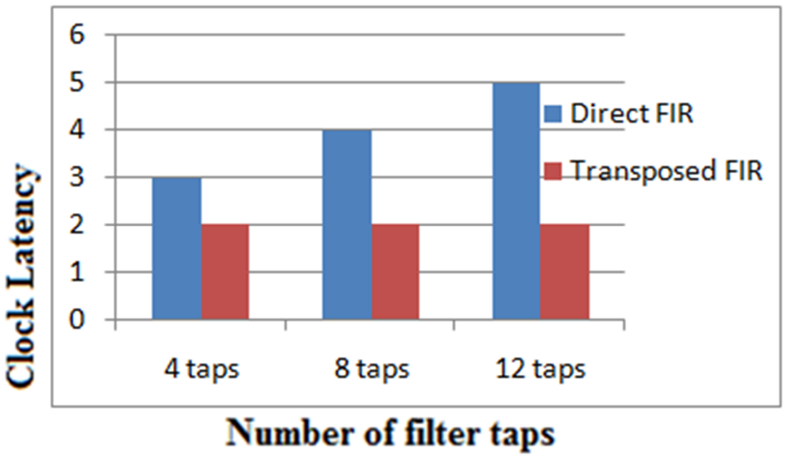

Figure-2 Comparison of two filters design by clock latency Vs number of filter taps. Ref. [4]

Figure 3 shows the comparison of logic cells and number of flip-flops which are used by Direct and Transposed FIR filter. Since the filter block used in this design is core of the design, therefore device area utilization is decreases as the number of taps increases in Transposed FIR filter structure.

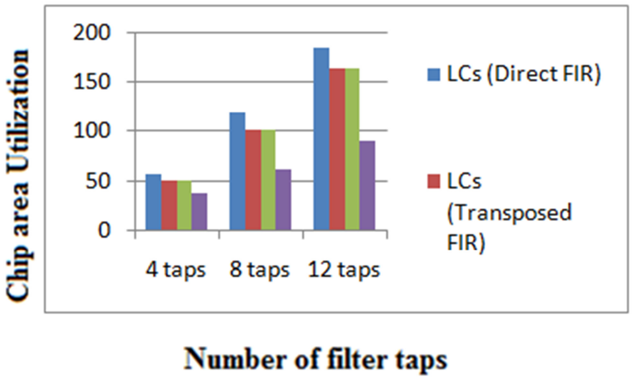

Figure-3 Comparison of two structures in terms of chip area utilization Vs number of filter taps. Ref [4]

Table-II Chip resource utilization for different frame size. Ref [4]

Frame Size		64 × 64	100 × 100	128 × 128	160 × 160	256 × 256	512 × 512
Frequency (MHz)		10.7	10.7	10.7	10.7	10.7	10.7
Resouces	Available	Used %	Used %	Used %	Used %	Used %	Used %
IOs	470	19.36	19.36	19.36	19.36	19.36	19.36
LCs	9984	7.94	8.84	8.62	8.46	8.72	7.94
DFFs	10160	2.74	3	2.83	2.88	2.96	2.74

Table-2 shows area utilization of the given resources at different frame size of the video sequence. The different frame sizes of any video sequence can be supported by the all necessary features of designed system with approximately same LCs and DFFs chip utilization [4].

Hardware implementation results: The FLEX 10KE family of the Embedded Programmable Logic Devices provided by Altera Corporation is used for designing of 3D-V algorithm on Hardware. All necessary features which are important to implement the whole design system on a single device like density, speed are provided by Flexible Logic Element Matrix (FLEX) architecture. The device EPF10K200SFC672-1 is used for the resources like IOs, LCs, DFFs.

Advantages: The main focus of this proposed work is selection of filter structure for the 1D-WT block. By the comparison of Direct and Transposed FIR filter structure, Transposed FIR filter structure gives the better results because of its less chip area utilization and clock latency. In addition, this design is suitable for any frame size of video sequence because of the Daubenchies-2 coefficients.

4. In 2013, Ch. Madhuri, T. Guru Charan, Y. Shekar proposed an architecture of 3D-WT. Frequently, there are certain instances when images need more accuracy and precision without loss of any information content like that in case of medical images. This paper includes the lifting based scheme (5, 3) filter which gives lossless mode of information. Further the lower computational complexity and reduced memory requirements in its implementation adds to its merits. Last but not the least, its lower implementation area and reduced power consumption makes its usage very popular over conventional convolution based DWT for image coding. An efficient approach is to implement a method whose coefficients are exactly given by finite precision numbers allowing complete lossless encoding [5]. In the following review, the design of Lossless 3-D DWT (Discrete Wavelet Transform) using Lifting Scheme Architecture has been modelled using the Verilog HDL and its utility is verified using the Model sim tool and can be synthesized using the Xilinx tool.

The design architecture discussed in this paper has been developed upon Xilinx programmable device (FPGA) XC4VFX140 with speed grade of 12 through the Xilinx ISE 7.1i tool. A consistent word length of 17 bits has been maintained throughout the processor to sustain sufficient data depth. Due to pipelining of multipliers, the critical path for the processor consists of single adder, making it comparatively faster [5]. Simulation is performed using ModelSim XE III 6.0a, which gives a set of results completely handshaking with the results from MATLAB 7.0.0, where a model of the hardware is developed. By using the proposed architecture design, improvements are shown in terms of hardware. Table-3 shows the generic features which are used to show the proper results using the proposed algorithm:

Table-III Overall design report. Ref [5]

Custom frame size	256 × 256
Group of frames (GOF)	Infinite
Maximum clock frequency	321 MHz
No. of occupies Slices	1776 (2%)
Total number of input LUT's	2188 (1%)

Advantages: In lifting based 3-D DWT architecture no restriction on the group of frames and subsequent improvements make it possible to have half latency during processing and a double throughput [5]. Once the functional check is ensured, discrete wavelet transform is synthesized by using Xilinx tool in Spartan 3E FPGA family. Hence it has been observed that the discrete wavelet transform (DWT) works at a maximum clock frequency of 99.197 MHz respectively.

5. V. R. Satpute, Ch. Naveen, K. D. Kulat and A. G. Keskar in 2014 presented that for any Video processing algorithm, efficient memory utilization is major criteria. In this paper, two techniques- Spatial and Temporal DWT are used for Video Encoding. There are two mechanisms which are based on levels of Spatial and Temporal DWT. This paper deals with the advantage of temporal DWT on video with compare to 2-D spatial DWT [6]. The implementation of Spatial DWT method restricts its usage to only 2-D images i.e. x and y-axis and video is considered as 3-D object. Instead of applying directly Spatial DWT can be apply indirectly to the video. This procedure can be applied by two main mechanisms which are discussed in the given paper.

Mechanism-1: The mechanism includes the multi level of spatial 2D-DWT and same level of temporal DWT. According to the Mechanism-1, 3 levels are applied of 2D Spatial DWT and Temporal DWT and after all the 3 steps resultant arrangement will look like Multi Resolution Analysis structure which is mainly applied on 3 axes i.e. row, column and time [6]. This mechanism can also be named as 3D multilevel pyramid decomposition. At last in this procedure, frames are reconstructed with the help of high pass components of highest level to get back the low pass components of previous level [6]. This process is repeated until the original video is reconstructed.

Mechanism-2: This mechanism process is same as the mechanism-1 but with slightly different modification in levels and reconstruction process. As the levels of temporal and 2D Spatial DWT are different so the reconstruction process is also different from the Mechanism-1 [6].

For the results, videos are taken from the internet and some of the captured with the help of Canon camera. These mechanisms doesn't apply on whole video, it uses the higher levels of temporal and 2D spatial DWT which reduces the utilization of system memory. Results of both mechanisms are compared in terms of DWT, IDWT (Inverse DWT), MSER (Mean square error ratio) [6].

4. Conclusion

In this review paper techniques or algorithms of 3D-WT are discussed for Video Compression. The Discrete Wavelet Transform (DWT) is advantageous over many Video Compression techniques e.g. DCT etc. We can conclude that by using different efficient algorithms, we can get the high degree compression results as compared to other many techniques e.g. DCT. DCT has the main disadvantage of DCT is that it splits the image into blocks at the higher compression ratio which in result gives blur edges of an image. This disadvantage can be overcome by using DWT, however DWT gives lower quality of image/video but high quality compression can be achieved by DWT at lower bit rates. Including this feature, the given algorithms are then implemented on hardware, which gives better results in terms of shorter response time and less chip utilization area. So, DWT performs better compression because of time-scale representation.

5. Future Work

As the area and the speed of the video encoding techniques are focused in the research work, the research can be extended through hardware utilization and optimization in latency. Research work will be aimed at the reduction of the number of DFFs in the filter design. The reduction of the path delay or the latency of the filter design for the DWT will also be the proposed work.

References

[1] Aydin Bicer, "Motion Compensated Three Dimensional Wavelet Transform Based Video Compression and Coding" Thesis of Natural and Applied Sciences of Middle East Technical University, January 2005.
[2] E. Moyano, F.J. Qui;es, 'A. Garrido, 2L. Orozco-1 Bardosa, 3J. Duato, "Efficient 3D wavelet Transform Decomposition for Video Compression," Second International Workshop on Digital and Computational Video (DCV'01) pp. 118, 2001.
[3] Yongjian Nian, Lehua Wu, Shibiao He and Yongjun Gu, "A New Video Coding Based on 3D Wavelet Transform," IEEE Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06),

2006.
[4] Samar Moustafa Ismail, Ali Ezzat Salama and Mohamed Fathy Abu-El Yazeed, “FPGA Implementation of an Efficient 3D-WT Temporal Decomposition Algorithm for Video Compression,” IEEE International Symposium on Signal Processing and Information Technology, 2007.
[5] Ch. Madhuri, T. Guru Charan, Y. Shekar, “Implementation for 3-D Discrete Wavelet Transform by Using Efficient Architecture,” International Journal of Computer Science and Mobile Computing, Vol. 2, Issue 11, pg. 203-209, November 2013.
[6] V. R. Satpute, Ch. Naveen, K. D. Kulat and A. G. Keskar, “Fast and Memory Efficient 3D-DWT Based Video Encoding Techniques”, Proceedings of the International Multi Conference of Engineers and Computer Scientists, Vol 1, 2014.