International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-4 Issue-6, August 2015

Differential Video Encoder Design Using Cascaded DWT and DCT

Divya Dileep, Naveen S

Abstract— Digital video technology has a wide variety of applications due to its several advantages over its analog counterpart. The use of digital video has been limited by its higher bit rate requirement. In this paper a novel technique for compression of video is proposed. This technique uses difference frames and Discrete Wavelet Transform and Discrete Cosine Transform. Wavelet transform provides approximations at different levels which require very less memory for storage compared to the original data. Cosine Transform represents approximation of signal with fewer coefficients. The algorithm has been implemented using Haar wavelet, Daubechies wavelet and biorthogonal wavelet and the performance in each case is evaluated using parameters such as Mean Square Error and Peak Signal to Noise Ratio.

Index Terms— Biorthogonal, Daubechies, difference frame, Discrete Cosine Transform, Haar, video compression, Wavelet transform

I. INTRODUCTION

These days digital video has become very popular due to its ease of storage, capacity for multicasting, ability to preserve data quality even when copied. Digital video communication has been found in many applications such as broadcast, video streaming, video conferencing, multimedia messaging etc. The original video size is very huge which results in difficulties in their storage as well as transmission. Hence video compression is of vital importance which reduces the size of data without affecting visual quality much. Fortunately, the video sequence contains a good amount of redundancy within the frame and also between successive frames. The compression techniques exploit the redundancy present in the video. There are mainly two types of redundancy present in the video sequence, ie; intraframe redundancy and interframe redundancy. Intra frame coding reduces the spatial redundancy present within a frame in the video. Temporal redundancy contained between adjacent frames in the video is of concern in the case of interframe coding. Interframe coding mostly uses motion estimation/compensation which involves a prediction step in the temporal domain. The most commonly used transformation for compression is Discrete Cosine Transform (DCT) [1]. In the MPEG standard techniques, block based motion compensation is applied for compression. Motion vectors are calculated using block matching algorithms and the current frame is predicted from reference frame using these motion vectors.

Manuscript published on 30 August 2015.
***** Correspondence Author (s)
Divya Dileep, Department of Electronics and Communication, LBS Institute of Technology for Women, Thiruvananthapuram, India.
Naveen S, Department of Electronics and Communication, LBS Institute of Technology for Women, Thiruvananthapuram, India.

Then the differences between predicted and original frames are estimated, which along with motion vectors are encoded and transmitted [2]. H.261, H.262, H.263 and H.264 are video coding standards brought into line by ITU-T [3]. They perform well at high bit rates but degrade at low bit rates due to the block based DCT scheme. These techniques results in noticeable blocking artifacts which can be resolved using wavelet transform. Wavelet decomposes even complex information into its elementary forms at different positions and scales. It also provides reconstruction with high precision. Wavelet literally means a small wave oscillating zero average function. A wavelet should satisfy the following conditions

$$\int_{-\infty}^{\infty} |\varphi(t)|^2 dt < \infty \qquad (1)$$

$$\int_{-\infty}^{\infty} |\varphi(t)| dt = 0 \qquad (2)$$

Several video compression techniques incorporated wavelet transform with motion estimation. In the system proposed by Wang.et al [4], the video frame is warped into common coordinate system by mosaic technique and 3D wavelet transform is then applied on warped frame. Motion estimation along with lifting based wavelet transform was presented in [5]. There is another video encoding technique that avoids the use of motion estimation/compensation called the three dimensional (3D) coding. 3D video encoders are mostly based on three dimensional wavelet transform (3D DWT). 3D coding is used for compression of volumetric medical data [6]. A multirate scheme based on 3D coding using camera pan compensation is presented in [7]. An adaptive subband based video encoder is proposed in [8]. The number of frames used in temporal direction is more for high correlation and number of frames is less when correlation is low. C. Podilchuk et.al developed a video coding technique based on a three-dimensional spatio-temporal subband decomposition and vector quantization [9]. Y Chen et.al [10] presented a coding scheme which performs three dimensional sub band transformations on the image sequences and the result is encoded using zero-tree coding scheme. This technique provided a result almost comparable to MPEG-2. A three dimensional extension for set partitioning in hierarchical trees (SPIHT) algorithm for low bit rates is proposed [11]. A fast backward coding of wavelet trees with good coding speed and lower memory requirement than SPIHT is presented [12]. It is useful for image coding but is not extended to 3D coding. The rest of the paper is organized as follows: section II explains the proposed algorithm based on difference frames and DWT/DCT.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

In section III the performance of the proposed method is evaluated in terms of peak signal to noise ratio, mean square error and compression rate. The results of algorithm implementation using three different wavelets are compared. Finally, in section IV some conclusions are drawn.

II. PROPOSED METHOD

A video is a sequence of frames such that each frame is identical to an image. So the removal of redundant data within a frame is similar to image compression. An additional dimension is to be considered in case of video i.e. temporal dimension. In this section, a video encoder based on the application of DWT onto the difference frames is proposed.The input video can be divided into different Group of Pictures(GoPs). Each GoP may contain eight, sixteen or thirty two frames depending on the video. On each frame in the GoP, first a 2D DWT is applied. 2D DWT is implemented using separable 1D DWTs in x and y directions. At first, each frame is scanned in the horizontal direction. DWT is obtained by passing each frame through a pair of low pass and high pass filters. The filter output is then scanned in the vertical direction and again passed through low pass and high pass filters. These filters split the bandwidth of the signal into two halves, which are mirror images. The filter output is fed to a decimator so that the size of the input is preserved. The image is passed through a series of low pass and high pass filters to obtain multilevel wavelet decomposition. The block diagram for the analysis section is shown in fig1. [13].

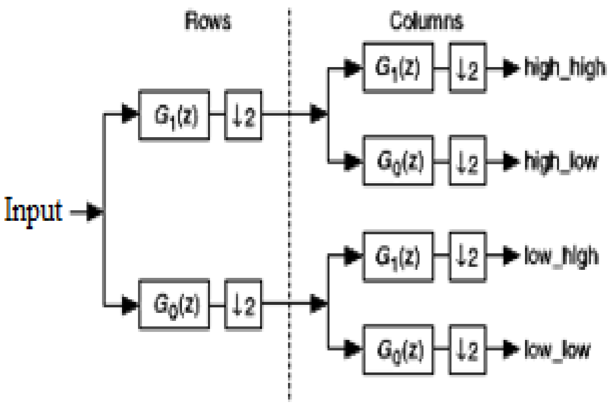

Figure 1: 2D DWT analysis section

The 1D forward transform is expressed as

$$c_n = \sum_{k=0}^{p-1} g_{0k}\, s_{2n+k},\quad 0\le n<N/2 \tag{3}$$

$$d_n = \sum_{k=0}^{p-1} g_{1k}\, s_{2n+k},\quad 0\le n<N/2 \tag{4}$$

where c_n represent the low pass down sampled result, d_n the high pass down sampled result, and g_0 and g_1, respectively, are the coefficients of the discrete-time filters G_0 and G_1 with filter p taps.

The synthesis operation is performed as follows:

$$s_n' = \sum_{k=0}^{p-1} g_{0k}'\, c_{(n-k)/2} + \sum_{k=0}^{p-1} g_{1k}'\, d_{(n-k)/2} \tag{5}$$

for $0\le n<N$.

where s_n' represents synthesized values, g_0' and g_1' are coefficients of discrete time filters G_0' and G_1'. In order to obtain perfect reconstruction, s_n' and s_n should be equal for $0\le n<N$.

The wavelet transform results in approximation, horizontal, vertical and detail coefficients. An important property of wavelet transform is the multiresolution property i.e. it provides good time resolution at high frequencies and good frequency resolution at low frequencies.

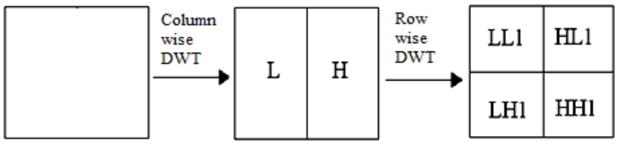

Figure 2: 2D DWT applied in spatial domain

The next step is to extract only the approximation coefficients from the set of wavelet coefficients.The mean value for other three set of coefficients are calculated and these coefficients are thresholded using this mean value or its scaled form. The first approximation frame of every GoP is set as the reference frame and the remaining frames in the GoP are represented by their difference with the first frame. The motion present within the frames is absorbed by the difference frames.

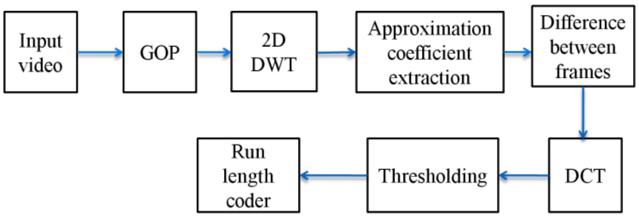

Figure 3: Block diagram of the proposed method

This is followed by the transformation of the coefficients obtained using DCT. 2-D DCT is applied on each set of coefficients. DCT doesnot introduce any loss in the source signal, it only transforms into a domain where they can be more effectively encoded.The DCT coefficient values are then thresholded and fed to the run length encoder system. The binary values of DCT coefficients are fed to run length encoder. Run length encoding is very useful for compression when the data contains repeated values. The encoding starts with the first element of the data. If the next value is same as the previous value then increment the count of consecutive values and continue examining the next values until a different value is found or reached the end of the input. Then output the value and the count. If it is different from previous value then output value followed by one.

III. PERFORMANCE EVALUATION

In this section, the proposed algorithm is tested on various standard video sequences such as ‘news’, ‘foreman’, ‘xylophone’. The implementation is performed using MATLAB R2012b on an i3-3110M CPU, 2.4GHz with Windows 8 operating system. The algorithm is performed on each video sequence using three different wavelets – Haar wavelet, Daubechies (db4) wavelet and Biorthogonal wavelet. The parameters used for evaluating the performance of the proposed algorithm are peak signal to noise ratio (PSNR), mean square error (MSE) and compression rate.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

$$\text{PSNR(dB)} = 20*\log_{10}(255/\sqrt{MSE}) \tag{5}$$

where

$$\text{MSE} = \frac{1}{m*n}\sum_{y=1}^{m}\sum_{x=1}^{n}(I(x,y) - I^{1}(x,y))^{2} \tag{6}$$

$$\text{Compression rate} = (N_a - N_c)/N_a \tag{7}$$

where N_a is the number of elements in the original frame and N_c is the number of elements in the compressed form. The proposed algorithm is implemented using Haar, Daubechies and Biorthogonal wavelets. The results obtained for the proposed method for three different wavelets are listed in Table I and Table II. In Table I, the MSE and PSNR values of 11th frame of video 'news.avi' is listed. In Table II, the MSE and PSNR values of 23rd frame of video 'news.avi' is listed. The MSE and PSNR values obtained in the three cases are plotted in fig 4 and fig 5 respectively. The algorithm is proposed on two different values of GoP i.e. 8 and 16. The MSE values obtained in both cases are shown in Table III. Figures 6 and 7 show the comparison of MSE and PSNR values for GoP equal to 16. The decompressed video frames for the three wavelets are shown in figures 8,9 and 10.

TABLE I : PERFORMANCE EVALUATION USING DIFFERENT WAVELETS IN TERMS OF MSE AND PSNR FOR 11THFRAME

Wavelet	MSE	PSNR
Haar	10.3440	31.9973
Daubechies	6.3120	34.1425
Biorthogonal	5.3664	34.8474

TABLE II : PERFORMANCEEVALUATION USING DIFFERENT WAVELETS IN TERMS OF MSE AND PSNR FOR 23RDFRAME

Wavelet	MSE	PSNR
Haar	24.2006	28.3059
Daubechies	13.0459	30.9895
Biorthogonal	12.5825	31.1465

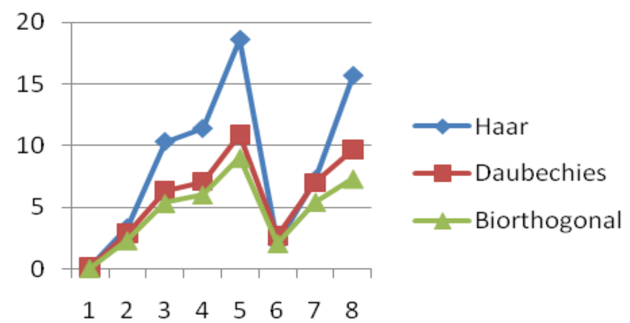

Figure 4. Comparison of MSE values for Haar, Daubechies and Biorthogonal wavelets for GoP= 8

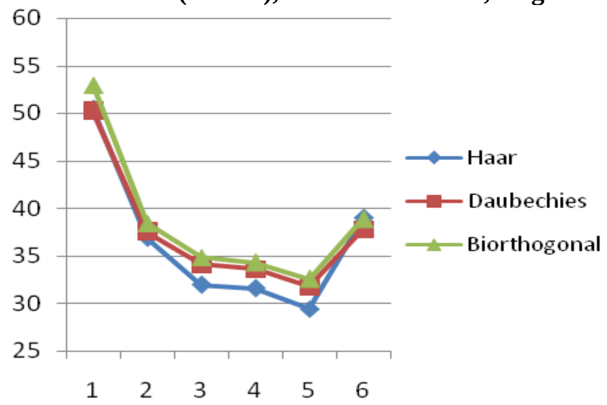

Figure 5. Comparison of PSNR values for Haar, Daubechies and Biorthogonal wavelets for GoP = 8

TABLE III : COMPARISON OF MSE VALUES FORDIFFERENT VALUES OF GOP

GoP	MSE
8	2.7312
16	11.945

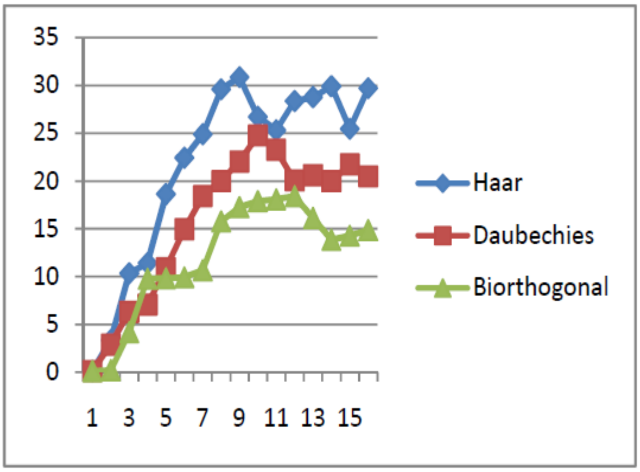

Figure 6. Comparison of MSE values for Haar, Daubechies and Biorthogonal wavelets for GoP=16

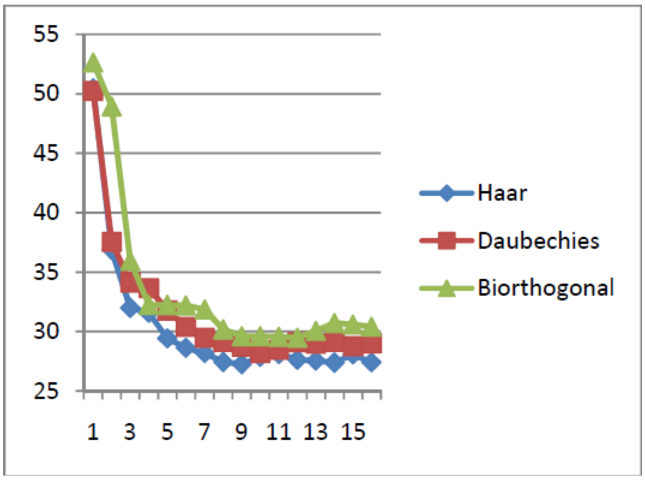

Figure 7. Comparison of PSNR values for Haar, Daubechies and Biorthogonal wavelets for GoP = 16

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Figure 8. Third frame of input video 'news.avi'

Figure 9.Third frame of 'news' reconstructed using Haar wavelet

Figure 10.Third frame of 'news' reconstructed using Daubechies wavelet

Figure 11.Third frame of 'news' reconstructed using Biorthogonal wavelet

IV. CONCLUSION

A novel video encoding technique using difference frames based on DWT and DCT is presented here.The PSNR values obtained in three cases show that a better reconstruction is possible with the biorthogonal wavelet compared to the other two for nearly same compression rate. It is seen that MSE increases as the GoP value is increased. In future this algorithm can be implemented on a dedicated embedded system which can give a superior performance in terms of processing time.

REFERENCES

[1] An overview of mpeg family and its applications, S.Vetrivel, M.Gowri, M.Sumaiya Sultana, DrG.Athisha, Indian Journal of Computer Science and Engineering,Vol. 1 No. 4 240-250, December 2010.

[2] Overview of the H.264/AVC Video Coding Standard Thomas Wiegand, Gary J. Sullivan, GisleBjøntegaard, and Ajay Luthra, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 13, No. 7, July 2003.

[3] K. Rao and J. Hwang, Techniques and Standards for Image, Video, and Audio Coding, Prentice Hall, Upper Saddle River, NJ, 1996.

[4] A. Wang, Z. Xiong, P.A. Chou, and S. Mehrotra, "Three-dimensional wavelet coding of video with global motion compensation", Proc. DCC '99, Snowbird, Utah, March 1999.

[5] B. Pesquet-Popescu and V. Bottreau, "Three-dimensional lifting schemes for motion compensated video compression," Proc. ICASSP'01, Salt Lake City, UT, May 2001.

[6] P Schelkens, A Munteanu, J Barbariend, M Galca, X Giro-Nieto, J Cornelis,Wavelet coding of volumetric medical datasets. IEEE. Trans. Med. Imaging.22(3), 441–458 (2003).

[7] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," IEEE Trans. Image Processing, vol. 3, pp. 572–588, Sept. 1994.

[8] Y. K. Kim, R. C. Kim, and S. U. Lee, "On the adaptive 3D subbandvideo coding," in Proc. SPIE, vol. 2727, pp. 1302–1312, Mar. 1993.

[9] C. Podilchuk, N. Jayant, and N. Farvardin, "Three-dimensional subband coding of video,"IEEE Trans. on Image Processing, vol. 4, pp. 125-139, February 1995.

[10] Y Chen, WA Pearlman, in Visual Communications and Image Processing,vol. 2727. Three-dimensional subband coding of video using thezero-tree method (SPIE, Bellingham, 1996), pp. 1302–1309.

[11] B J Kim, Z Xiong, WA Pearlman, Low bit-rate scalable video coding with 3Dset partitioning in hierarchical trees (3D SPIHT). IEEE Trans. Circuits Syst.Video Tech. **10**, 1374–1387 (2000).

[12] L Ye, T Karp, B Nutter, S Mitra, J Guo, in Signals, Systems and Computers,2006. ACSSC '06. Fortieth Asilomar Conference on. Three-dimensionalsubband coding of video with 3-D BCWT (IEEE, NY, 2006), pp. 401–405.

[13] Fast and Memory Efficient 3D-DWT Based Video Encoding Techniques, V. R. Satpute, Ch. Naveen, K. D. Kulat and A. G. Keskar, Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol I, IMECS 2014, March 12 - 14, 2014, Hong Kong.

[14] A Combined DWT-DCT approach to perform Video compression base of Frame Redundancy, Jasmeetkaur, Ms.Rohini Sharma, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 9, September 2012.

[15] Multicore-based 3D-DWT video encoderVicenteGaliano, Otoniel Lopez-Granado, Manuel P Malumbres and Hector Migallon, EURASIP Journal on Advances in Signal Processing 2013.

Divya Dileep is currently pursuing M.Tech. Degree in Signal Processing with the Department of Electronics and Communication Engineering, LBS Institute of Technology for women, Poojappura, Trivandrum, Kerala. She received B.Tech degree from the University of Kerala, Thiruvananthapuram, in 2013 in Electronics and Communication Engineering.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Naveen S, currently working as Assistant Professor,in department of Electronics and Communication Engineering at LBS Institute of Technology for Women, Thiruvananthapuram, Kerala, India. Has Obtained B.Tech and M.Tech degrees from University of Kerala and is currently pursuing research in the area of pattern recognition.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)