Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 70 – 77

Procedia Computer Science

International Conference on Computational Modeling and Security (CMS 2016)

Enhanced Scalable Video Coding Technique with an Adaptive Dual Tree Complex Wavelet Transform

Sunil Kumar B S[a], Dr. A S Manjunath[b], Dr. Christopher S[c], Dr. Rajalaxmi Menon[d]

[a]*Associate Professor, CSE,GMIT,Davngere- 577004, India*
[b]*Professor, SIT, Tumkur – 572 103, India*
[c]*Secretary Dept. Of Defence & R&D, Director General, DRDO, New Delhi-110011, India*
[d]*Scientist –G, CABS, DRDO, Bengaluru – 560 037, India*

Abstract

The scalable Video is the most suitable video format for the current day scenario where there are different varieties of electronic gadgets used to communicate and play back. As the device resolution varies, the codec designed for Scalable video (H.264/AVC) is the only solution for its support. But the challenge is in its transmission due to the limited network resources in wireless networks viz; bandwidth. This introduces jitter during video playback. This needs an effective Compression technique to reduce the possible transmission time. The paper presents a novel Adaptive Dual Tree Complex wavelet Transoform (ADT-CWT) using the efficient Motion Block Estimation (MBE), Diamond search with Large Diamond Search Path (LDSP) and a Small Diamond Search Path (SDSP). This technique effectively improves the performance of Scalable Video in wireless transmission. The PSNR and the Compression Ratio (CR) are considered for the comparison.

Peer-review under responsibility of the Organizing Committee of CMS 2016

Keywords:Compresion Ratio (CR); Dual Tree Complex Wavelet Transform; PSNR; Scalable Video Coding

1. Introduction

In past few decades, Compression of video has been developed as an integral part of the total communication[1]. The applications of Digital video communication in the present day are Broadcast services, storage of video, wire and wireless network services, and in the terrestrial channels. The solution is video compression[2,3,4]. All video coding standards make use of the redundancy inherent within digital video information in order to reduce its bit rate[5]. The different scenarios of Digital video applications of are; i) Broadcasting, Pay on view service, terrestrial and cable transmission channels, ii) Interactive video communication over wired and wireless in the Internet and Local network, and iii) Various storage formats[7,8]. The classification of video compression methods are;

1877-0509
Peer-review under responsibility of the Organizing Committee of CMS 2016
doi:10.1016/j.procs.2016.05.179

i) Prediction, ii) Transformation, iii) Quantization, iv) Entropy coding[6]. A video has sequence of frames or images, in general, in the order of time. This hints that compression of video can be performed by prediction over the previous frame. Compression can be done by the subtraction of present and previous frame and to code only the residual part of error[9]. This can even b e done by searching the right part of image subtracted from the previous frame. The reason is that the video contains much spatial and temporal redundancy. In a single frame, nearby pixels are often correlated with each other. This is called spatial redundancy, or the intra frame correlation. Another one is temporal redundancy, which means adjacent frames are highly correlated, or called the inter frame correlation[10,11].

Scalable Video Coding (SVC) is a technique that has been in the signal processing community for already sometime[12]. SVC has designed to cater the video of lower resolutions, lower rate of frame transmission and also reduced quality, to the heterogeneous devices having lower resolution displays and reduced computational capabilities[13]. Lagrangian technique helps to produce an optimal bit allocation to an encoder, having bit rate constraint. In contrast to conventional non layered video coding, SVC is able to serve all users by single bit stream of transmissio[14]. Modern video transmission system challenges are effectively solved by the SVC. This paper refers the term 'Scalability' as preparing a frame with reduced bits to suit the video for various devices in the network[15,16,17].

SVC is a bitstream which is encoded by sampled bits of the original video. The encoder is a complex one and it produces many sets of lower resolutions. SVC encoder has a compression technique involved in it to support the transmission to the direct end users on the broadcast system. In the research, different SVC methods were utilized for performing video coding process. Among them, wavelet based SVC is the most widely used technique to encode and decode the video frames. In this wavelet-based SVC technique, initially the frames are encoded in the base layer with low resolution and then these encoded frames are decoded by applying inverse process. The differences between the original and the decoded frames are calculated. This difference is then given as input to the wavelet process. The sub bands obtained from the wavelets are given as input to the coding techniques to acquire the high resolution bit streams. The most commonly used wavelet techniques like DCT and DWT have major drawbacks in their process. The disadvantage of DCT is that its vulnerability to block noise and restricted scalability, whereas DWT compress the image with less quality. Such drawbacks in the wavelet technique make the image blurred and provide low PSNR value due to the presence of small amount of noise. Also, the application of wavelet decomposition degrades the SVC performance. Moreover, there is no standardization exists for selecting the suitable techniques to solve those drawbacks. If all the aforesaid drawbacks in the literary works are solved, then the image quality is improved to better PSNR value. As there are no solutions for such issues, it is motivated to do research work in this direction.

The main aim of this research is to provide a better SVC (Scalable Video Coding) technique for video by solving the drawbacks that currently exist in the literary works. Hence, we intended to propose an enhanced SVC technique using wavelet decomposition. In the proposed technique, the frame's bit streams will be processed in two layers. Initially an enhanced zero coding model will be used at the base layer to make less memory storage of the low resolution bit streams. The variation between the low resolution decoded frame and the original frame will be computed. The difference value will be then given to the wavelet technique; here we will exploit a novel wavelet technique as adaptive complex wavelets transform (ACWT). The ACWT will be used to overcome the drawbacks of other wavelet methods and also it is a complex-valued extension of the standard discrete wavelet transform (DWT). The wavelet adaption is fashioned to select the wavelet coefficients based on the nature of the video. The selected wavelet coefficients by ACWT will be given to the quantization and coding techniques to acquire high resolution bit streams at the enhancement layer. Thus, the image can be coded more effectively by achieving higher PSNR ratio with high quality. The technique will be implemented in the working platform of MATLAB and the results will be analyzed to demonstrate the performance of the proposed wavelet-based SVC technique.

2. Proposed Methodology

The main aim of this research is to provide a better SVC (Scalable Video Coding) technique for video by solving the drawbacks that currently exist in the previous works. Hence, we intended to propose an enhanced SVC technique using wavelet decomposition. In this paper, a video compression scheme based on the Adaptive Dual Tree Complex Wavelet Transform (ADT- CWT) representation and motion compensation is presented. For representing images and video signals, the multi- resolution/ multi-frequency nature of the standard discrete wavelet transform is an ideal tool. In the multi-resolution motion compensation approach, motion vectors in higher resolution are predicted by large pattern diamond search (LPDS) for fast block matching the motion vectors in the lower resolution are predicted by small pattern diamond search(SPDS) refined at each step. A method based on variable block-size motion compensation scheme in which the size of a block is adjusted to its size of input video frame in this paper, has been proposed. This scheme gives a significant portrayal of the intrinsic motion structure and also impressively lessens the searching and matching time. After wavelet decomposition, each scaled sub-frame has a tendency to have distinctive statistical properties. Every motion frames are encoded and decoded using SPIHT encoder and decoder. An adaptive truncation process is implemented and a bit allocation scheme similar to that in the transform coding is inspected by adapting to the local variance distribution in each scaled sub-frame. The block diagram of proposed method is shown in figure.1.

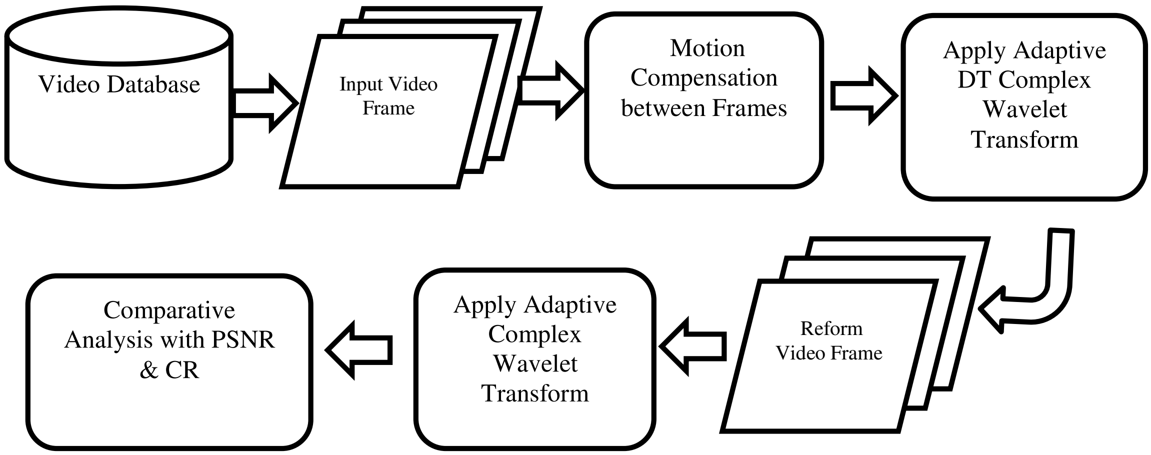

Fig. 1. Block diagram of proposed ADT-CWT on Scalable Video

3. Technology

The proposed method has many of the technologies which are combined and are made the necessary one to produce the better results in the performance of video delivery. The following are the technologies adopted in the implementation;

- Diamond search algorithm
- Adaptive Dual Tree Complex Wavelet Transform (ADT-CWT)

3.1 Diamond Search Algorithm

Diamond Search (DS) algorithm is one of the most efficient techniques of Block Matching Algorithm (BMA) this helps in faster Motion Estimation (ME). The DS is like Four Step Search (4SS). There are number of variants in the New Cross Diamond Search (NCDS); Small Cross-Shaped Pattern (SCSP), Large Cross-Shaped

Pattern (LCSP), Small Diamond-Shaped Pattern (SDSP) and Large Diamond- Shaped Pattern (LDSP) Shown in Fig 2 (a) & (b). These are rather using as an alternatives can be used in a combination to achieve the best performance out of the core Diamond Search (DS). The two different sorts of fixed search patterns SDSP and LDSP are used by DS algorithm. In Fig 3 DS is illustrated example of the search path with motion vector (-4,-2) in five search steps—four times of LDSP and one time SDSP. In the example the LDSP has nine checking points the centre point is surrounded by eight points this forms a diamond outline. The SDSP comprising of five checking points those create a smaller diamond outline.

At the beginning LDSP is used and when the Minimum Block Distortion (MBD) happens at the center point then it moves to last step. The subsequent step, aside from the last step, use LDSP, is recurrently utilized till it reaches the center point.

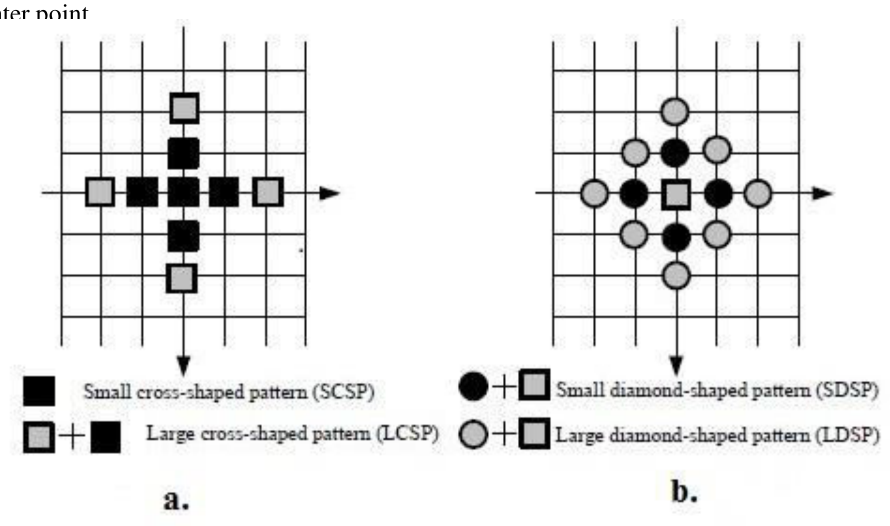

Fig. 2. Search Patterns (a) SCSP and LCSP and (b) SDSP and LDSP

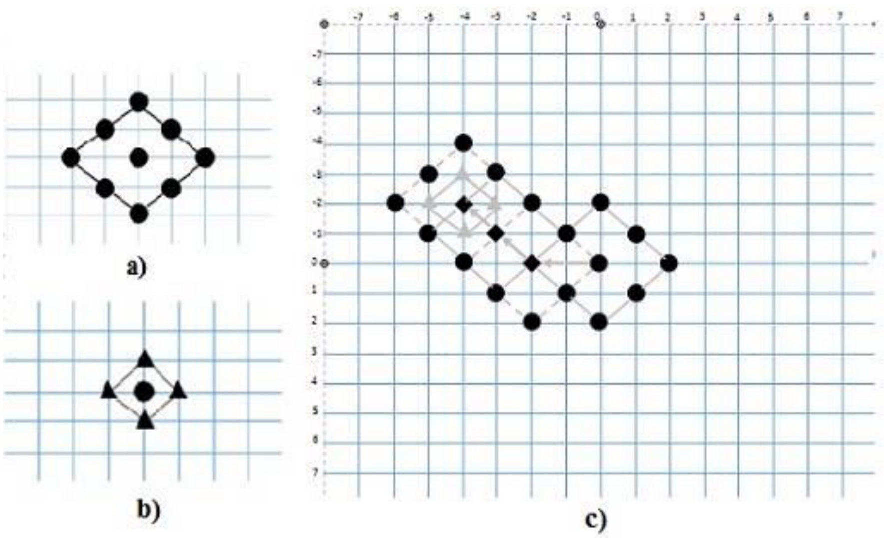

Fig. 3. Search Patterns (a) Large Diamond Search Path (LDSP), (b) Small Diamod Search Pattern (SDSP), and (c) Search path example to the motion vector (-4,-2) in five search steps—four times of LDSP and one time SDSP.

At last, SDSP is acted in place of LDSP to find out the MBD point. Among the five checking points in SDSP, the MBD point found in this stage produces the best matching block. DS is one of the most efficient which renders fast motion estimation algorithm consistently for the image sequence with wide range of motion content. It overrides the well-recognized Three-Step Search (TSS) algorithm and accomplishes close Mean Square Error (MSE) performance contrasted with New Three-Step Search (NTSS) algorithm which will reduce the computation by 22% roughly.

After completion of diamond search, the motion between the determined video frames gives motion vectors. These vectors contain rapid pixel changes or high motion portion between the consecutive frames. Then, APIHT method has been applied on highly motion blocks. After that ADT-CWT has been applied on the whole frame.

3.2 Adaptive Dual Tree Complex Wavelet Transform (ADT-CWT)

The Complex Wavelets Transforms (CWT) uses complex-valued filtering (analytic filter) that separates the real or complex signal into real and imaginary parts in transform domain. The coefficients of real and imaginary will compute both the amplitude and phase information, this type of information is useful to precisely describe the energy localization in oscillating functions (on the basis of wavelet). The edges in signal processing applications manifest them as oscillating coefficients in the wavelet domain. The value of these coefficients provides the strength of the singularity whereas the phase indicates its location. After getting motion blocks from DS algorithm, SPIHT has been applied for each.

The flowchart of Set portioning using SPIHT algorithm is shown in Fig. 4. In the process the given image is divided into 10 sub-bands. At this point the said method finds the maximum and the iteration number. In the second stage, the technique places the DWT coefficients into sorting pass which finds the coefficients which are significant among others and the sign of the these significant coefficients are embedded. In the third step, the significant coefficients are placed into the refinement pass that use two bits to exact the reconstruct value for closing to real value. The above three steps are iterative, next iteration decreases the threshold $T_n = T_{n-1} / 2$ and the reconstructive value $R_n = R_{n-1} / 2$. Next step, the encoding bits access entropy coding and then transmits.

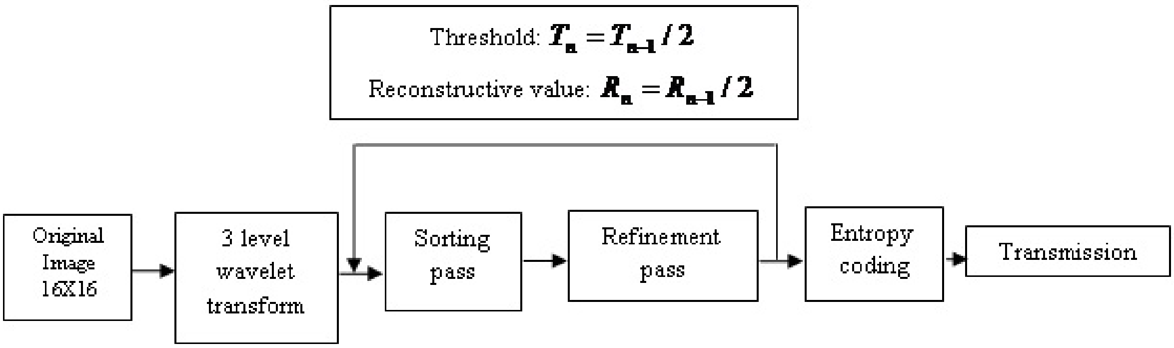

Fig. 4. Flow of Set Partitioning using SPIHT

3.2.1 Dual Tree Complex Wavelet Transform

The dual-tree complex wavelet transform utilizes two real DWTs. These two transforms together provide an overall transform. The first DWT transform point toward the real part and the second transform point toward the imaginary part of Eqn. (1).

$$\psi(s) = \psi_j(s) + j\psi_j(s) \tag{1}$$

A directional filter of DTCWT is applied on the frame for decomposition purposes. The resultant images are the output of the one directional filter from the low pass and high pass filters. The DTCWT coefficients that are achieved from the first filter bank are known as the real part and the coefficients that are achieved from the other filter are known as the imaginary part. The real part of the image contains less significant data in comparison with the imaginary part which comprises of more information. In each comparison of motion we have obtained an initial frame or i-frame and it is taken for DTCWT decomposition. Before decomposition it is necessary to encode and decode the high motion estimated region using SPIHT. DT-CWT is applied to full size of frame after encoding of blocks using the adaptive method and SPIHT.

4. Experimental Setup and Results discussion

The algorithms were written in the MATLAB code and are run with an Intel processor of 2.66GHz clock cycle having a RAM of 4GB. The simulation is run on MATLAB2003a version. The experiment was conducted on both the Non adaptive and the adaptive technique to show the results improvisation. The input is a standard 'VIPtraffic' of 352 x 254 resolution for 30 frames at the rate of 30fps. The search range is relative to the temporal displacement of +/- 8 pixels per frame. The results of both the Peak Signal to Noise Ratio (PSNR) and the Compression Ration (CR) are compared and are tabulated in TABLE 1.

TABLE 1: Performance Comparison of PSNR over CR

	PSNR		CR	
	Non Adaptive Method	**Adaptive DT-CWT Method**	**Non Adaptive Method**	**Adaptive DT-CWT Method**
Frame1	28.3573	27.7263	0.9698	0.9671
Frame2	28.2684	27.6859	0.9684	0.9651
Frame3	28.3146	27.7974	0.9716	0.9677
Frame4	28.4549	27.8940	0.9644	0.9610
Frame5	28.4269	28.0600	0.9655	0.9620

a. b. c.

Fig.5. VIP traffic video frame @ 352x240 resolution (a) Original Video frame (b) Non Adaptive method (c) Adaptive DT CWT method.

The CRs of the Adaptive DT CWT are improved over the Non Adaptive method. The CR of five frames are indicating a clear improvisation however the PSNR values are remain good with little less compare to the Non adaptive method which is not considerable. It is important to consider the visual quality of the low temporal resolution frames. Without motion compensation, the low-pass frames contain disturbing ghosting artifacts. This is largely avoided by motion compensation with both motion models. We note that even when there is no increase in reconstructed PSNR, as with test sequence, the visual quality (Fig 5) of the low-pass frames is still significantly improved.

5. Conclusions

The paper presents an Adaptive Dual Tree Complex Wavelet Transform method. This gives an improved performance of scalable video compression over the Non adaptive method. The motion compensated diamond search algorithm using SDSP and LDSP for temporal wavelet decomposition of video sequences has been extended to incorporate SPIHT DT-CWT model. In many cases this leads to improved CR performance and visual quality in comparison to a block-based model. Future work will be directed at implementing superior motion models, and optimizing rate allocation for the motion information.

References

1. Sullivan and Wiegand, "Video Compression - From Concepts to the H.264/AVC Standard", In Proceedings of the IEEE, Vol. 93 , No. 1, pp. 18-31, 2005.
2. Rutika Joshi, Rajesh Kumar Rai and Jigar Ratnottar, "Review of Different Standards for Digital Video Compression Technique", International Journal of advancement in electronics and computer engineering, Vol. 1, No. 1, pp. 32-38, April 2012.
3. Ortega and Ramchandran, "Rate-distortion methods for image and video compression: An Overview", IEEE Signal Processing Magazine, Vol. 15 , No. 6, pp. 23-50, 1998.
4. Sullivan and Wiegand, "Rate-distortion optimization for video compression", IEEE Signal Processing Magazine, Vol. 15 , No. 6, pp. 74 - 90, 1998
5. Simon S. Lam, Simon Chow and David K. Y. Yau, "An algorithm for lossless smoothing of MPEG video", In Proceedings of the conference on Communications architectures, protocols and applications, Vol. 24, No. 4, pp. 281-293, Oct. 1994.
6. Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard and Ajay Luthra, "Overview of the H.264/AVC video coding standard", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 13 , No. 7, pp. 560- 576, July 2003
7. Jerome Edward Lengye, "Compression of time-dependent geometry", In Proceedings of the symposium on Interactive 3D graphics, pp. 89 – 95, 1999
8.Atul Puria, Xuemin Chenb and Ajay Luthrac, "Video-coding using the H.264/MPEG-4 AVC compression standard", Signal Processing: Image Communication, Vol. 19, No. 9, pp. 793–849, October 2004.
9.Weiyao Lin, Panusopone, K., Baylon and Ming-Ting Sun, "A Computation Control Motion Estimation Method for Complexity- Scalable Video Coding", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 20 , No. 11, pp. 1533-1543, 2010.
10. Tarek Ouni, Walid Ayedi and Mohamed Abid, "A complete non- predictive video compression scheme based on a 3D to 2D geometric transform", International Journal of Signal and Imaging Systems Engineering, Vol. 4, No. 3, pp. 164-180, 2011.
11. Baig, Yousuf, Lai, Edmund M.K., Punchihewa and Amal, "Distributed Video Coding Based on Compressed Sensing", IEEE International Conference on Multimedia and Expo Workshops, pp. 325- 330, 2012.
12. Chris Develder, Peter Lambert, Wim Van Lancker, StefaanMoens, Rik Van de Walle, JelleNelis, Dieter Verslype, Steven Latré, Nicolas Staelens and Nick Vercammen, "Delivering scalable video with QoS to the home", Telecommunication Systems, Vol. 49, No. 1, pp. 129-148, 2012
13. Jong-Seok Lee, Francesca De Simone, Naeem Ramzan, Zhijie Zhao, Engin Kurutepe, Thomas Sikora, Jorn Ostermann, EbroulIzquierdo and Touradj Ebrahimi, "Subjective evaluation of scalable video coding for content distribution", In Proceedings of the international conference on ultimedia, pp. 65-72, 2010.
14. TienAnh Le, Hang Nguyen and Hongguang Zhang, "Eval SVC - An evaluation platform for ScalableVideo Coding transmission", In Proceedings of the IEEE 14th International Symposium on Consumer Electronics, pp. 1-6, 2010.
15. Amer, I., Lucarz, C., Roquier, G., Mattavelli, M., Raulet, M., Nezan and Deforges, "Reconfigurable video coding on multicore", IEEE Signal Processing Magazine, Vol. 26 , No. 6, pp. 113-123, 2009.
16. Lipeng Du, Nugraha, I.B. ; Morita and Lin Zhang, "An experimentation testbed for scalable video coding on IMS-based IPTV", In Proceedings of the 2nd IEEE International Conference on Network Infrastructure and Digital Content, pp. 24-26 Sept. 2010.

17. Ahn, Jae-Kyun, Yeong Jun Koh, and Chang-Su Kim. "Efficient Fine Granular Scalable Coding of 3D Mesh Sequences", IEEE Transactions on Multimedia, Vol. 15 No. 3, pp. 485-497, 2013.