AN IMPROVED LOW COMPLEX SPATIALLY SCALABLE ACC-DCT BASED VIDEO COMPRESSION METHOD

[1]G.Suresh, [2]P.Epsiba, [3]Dr.M.Rajaram, [4]Dr.S.N.Sivanandam

[1]Research Scholar, CAHCET, Vellore, India.
[2]Lecturer, Department of ECE, CAHCET, Vellore, India.
[3]Professor & Head, Department of EEE, GCE, Thirunelveli, India.
[4]Professor & Head, Department of CSE, PSG Tech, Coimbatore, India.

Abstract-In this paper, we propose a low complex Scalable ACC-DCT based video compression approach which tends to hard exploit the pertinent temporal redundancy in the video frames to improve compression efficiency with less processing complexity. Generally, video signal has high temporal redundancies due to the high correlation between successive frames. Actually, this redundancy has not been exposed enough by current video compression techniques. Our model consists on 3D to 2D transformation of the video frames that allows exploring the temporal redundancy of the video using 2D transforms and avoiding the computationally demanding motion compensation step. This transformation turns the spatial temporal correlation of the video into high spatial correlation. Indeed, this technique transforms each group of pictures (GOP) to one picture (Accordion Representation) eventually with high spatial correlation. This model is also incorporated with up/down sampling method (SVC) which is based on a combination of the forward and backward type discrete cosine transform (DCT) coefficients. As this kernel has various symmetries for efficient computation, a fast algorithm of DCT-based Scalability concept is also proposed. For further improvement of the scalable performance, an adaptive filtering method is introduced, which applies different weighting parameters to DCT coefficients. Thus, the decorrelation of the resulting pictures by the DCT makes efficient energy compaction, and therefore produces a high video compression ratio. Many experimental tests had been conducted to prove the method efficiency especially in high bit rate and with slow motion video. The proposed method seems to be well suitable for video surveillance applications and for embedded video compression systems.

Index terms: SVC (Scalable Video Coding), Group of Pictures (GOP), ACC-DCT, Spatial and Temporal Correlation.

1. INTRODUCTION

The main objective of video coding in most video applications is to reduce the amount of video data for storing or transmission purposes without affecting the visual quality. The desired video performances depend on applications requirements, in terms of quality, disks capacity and bandwidth. For portable digital video applications, highly-integrated real-time video compression and decompression solutions are more and more required. Actually, motion estimation based encoders are the most widely used in video compression. Such encoder exploits inter frame correlation to provide more efficient compression. However, Motion estimation process is computationally intensive; its real time implementation is difficult and costly [1][2]. This is why motion-based video coding standard MPEG[12] was primarily developed for stored video applications, where the encoding process is typically carried out off-line on powerful computers. So it is less appropriate to be implemented as a real-time compression process for a portable recording or communication device (video surveillance camera and fully digital video cameras). In these applications, efficient low cost/complexity implementation is the most critical issue. Thus, researches turned towards the design of new coders more adapted to new video applications requirements. This led some researchers to look for the exploitation of 3D transforms in order to exploit temporal redundancy. Coder based on 3D transform produces video compression ratio which is close to the motion estimation based coding one with less complex processing [3][4][5][6]. The 3d transform based video compression methods treat the redundancies in the 3D video signal in the same way, which can reduce the efficiency of these methods as pixel's values variation in spatial or temporal dimensions is not uniform and so, redundancy has not the same pertinence. Often the temporal redundancies are more relevant than spatial one [3]. It is possible to achieve more efficient compression by exploiting more and more the redundancies in the temporal domain; this is the basic purpose of the proposed method. The proposed method consists on projecting temporal redundancy of each group of pictures into spatial domain to be combined with spatial redundancy in one representation with high spatial correlation. The obtained representation will be compressed as still image with JPEG coder. The rest of the paper is organized as follows: Section 2 gives an overview of basic definition of three dimensional DCT. Section 3 gives the basics of the proposed method and the modifications made to improve the compression ratio and also reduce the complexity. Experimental results were discussed in section 4. The section 5 concludes this paper with a short summary.

2. DEFINITIONS

2.1. Three dimensional DCT

The discrete cosine transform (DCT)[4][7] has energy packing efficiency close to that of the optimal Karhunen-Loeve transform. In addition, it is signal independent and can be computed efficiently by fast algorithms. For these reasons, the DCT is widely used in image and video compression. Since the common three-dimensional DCT kernel is separable, the 3D DCT is usually obtained by applying the one-dimensional DCT along each of the three dimensions. Thus, the N ×N ×N 3D DCT can be defined as

$$X(u,v,w) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} x(i,j,k)C(i,u)C(j,v)C(k,w) \quad (1)$$

$$x(i,j,k) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} X(u,v,w)C(i,u)C(j,v)C(k,w) \quad (2)$$

$$\text{where } C(p,q) = \begin{cases} \frac{1}{\sqrt{N}} & ,q=0 \\ \sqrt{\frac{2}{N}}cos\left(\frac{(2p+1)q\pi}{2N}\right) & ,q\neq 0 \end{cases}$$

3. PROPOSED METHOD

The fundamental idea is to represent a video sequence with highly correlated form. Thus we need to expose both spatial and temporal redundancy in video signal. The video cube is the input of our encoder, which is a number of frames. This video cube will decomposed in to temporal frames which will be gathered into one 2D frame. The next step consists of coding the obtained frame. Normally, the variation of the 3D video signal is much less in the temporal domain than the spatial domain, the pixels in 3D video signal are more correlated in temporal domain[3].

For a single pixel model is denoted as p(x,y,t) where p is pixel value; x,y are pixel spatial coordinates; t is video instance at time. The following assumption will be the basis of the proposed model where we will try to put pixels- which have a very high temporal correlation in spatial adjacency.

$$P(x,y,t)-p(x,y,t+1)<p(x,y,t)-p(x+1,y,t) \quad (3)$$

To exploit the succeeding assumption the temporal decomposition of the 3D video signal will be carried out and the temporal, spatial decomposition of one 8x8x8 video cube [8][9][10] is presented in the Figure.1. Thus the Accordion representation (Spatial Adjacency) is obtained from the basis assumption.

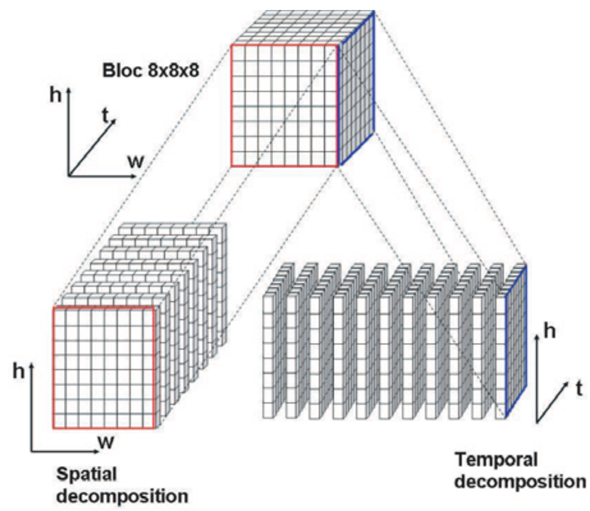

Fig.1: temporal and spatial decomposition of one 8x8x8 video cube.

Accordion representation is formed by collecting the video cube pixels which have the same column rank and these frames have a stronger correlation compare to spatial frames. To improve correlation in the representation we reverse the direction of event frames. This tends to put in spatial adjacency that the pixels having the same coordinate in the different frames of the video cube. The following example i.e., Figure.2 clearly projecting the Accordion representation also minimizes the distance between the pixel correlated in the source.

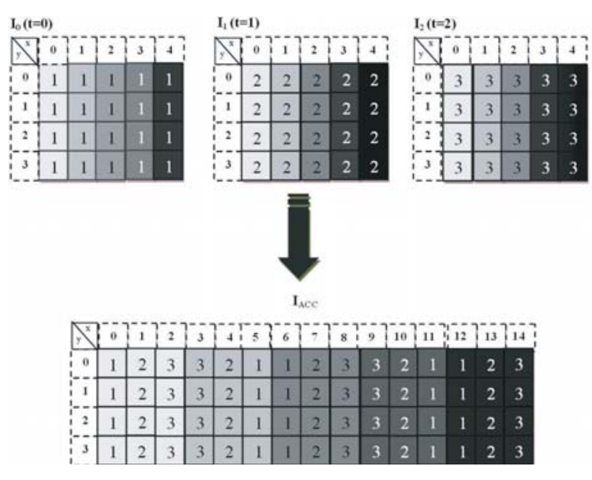

Fig.2: Accordion Representation Example

Continuation of the Accordion Representation a new concept is originated from scalable video coding (SVC)[16][17][18]19[20] technique; up/down sampling method using the DCT has a large degree of symmetries for efficient computation. Thus, a fast algorithm of the up/down sampling method is also included in our proposed method. For a performance improvement, an adaptive filtering method DCT up/down sampling is applied[13][14][15], which applies different weighting parameters to each DCT coefficient. Then we have to introduce quantization model and Entropy coding (RLE/Huffman) techniques for further performance improvement of the proposed system. A overall constructional details of the proposed model is explained in Figure.3.

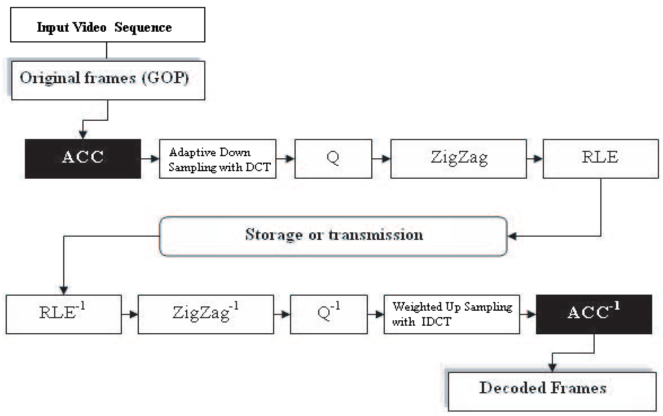

Fig.3: Complete Constructional Details of the proposed Model

3.1. Algorithm

1. Decomposition of the video in GOPs (Group of frames/pictures)
2. (a) Spatial Adjacency separation (Accordion Representation) of the GOP

For x=0: (L * N)-1 do
For y=0:(H-1) do
If ((x/N)mod 2)!=0 then
n=(N-1)-(x mod N)
else
n=x mod N
end if
IACC (x,y)=In (x/N, y)
With n= ((x/N) mod2)(N-1)+1-2((x/N) mod 2)(x mod N)

(b) For n=0:N-1 do
For x=0:L-1 do
For y=0:H-1 do
If(x mod 2)! =0 then
X ACC= (N-1)-n(x*N)
else
X ACC=n(x*N)
end if
In(x,y)=IACC(X ACC, y)
end for
end for
end for
with XACC=((x/N) mod2)(N-1)+n(1-2(x/N) mod2))+x

3. Decomposition of the resulting frame into 8x8 blocks.
4. Introduce down sampling filter/ up sampling filter with DCT.
5. Quantization of the obtained coefficients.
6. ZigZag coding of the obtained coefficient.
7. Entropy (Huffman) coding of the coefficients.

4. EXPERIMENTAL RESULTS

This section verifies the performance of the proposed low complex scalable ACC-DCT based video compression model. We summarize the experimental results with some analysis and comments. By understanding the performance of the proposed method with different GOP value the best compression rate is obtained with GOP=8. Here Figure.4 GUI model is created to integrate the encoder and decoder sections.Figure.5 shows the progress of frame separation from video sequence. Encoder model is shown in Figure.6 and Figure.7 is the example Accordion representation for one GOP video cube. Figure.8: GUI for Decoding and Validation Process. Figure.9: GUI for Reconstructed output validation, then history of entire simulation is specified as **ans.** Finally Figure.10 shows the plot between Frame number Vs PSNR(dB), Figure.11 finds the orientation flow estimation of the sample sequence and compared the strength of our proposed model with other leading standards shown in Figure.11.

Fig.4: GUI Model

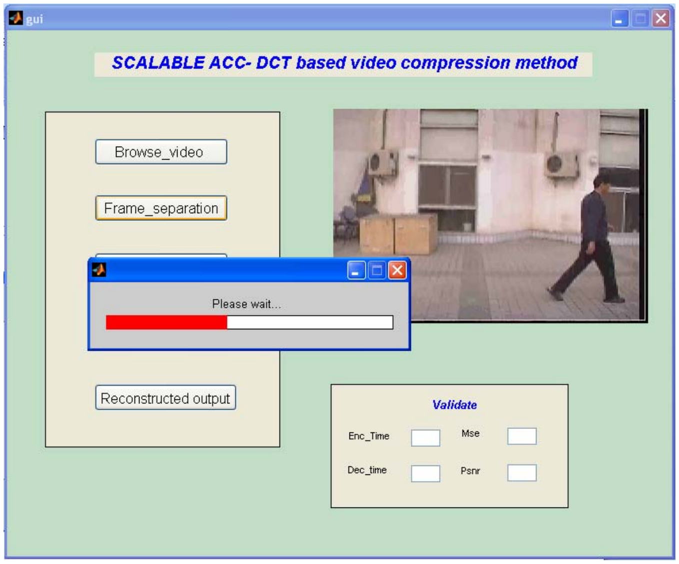

Fig.5: Frame Separation Model

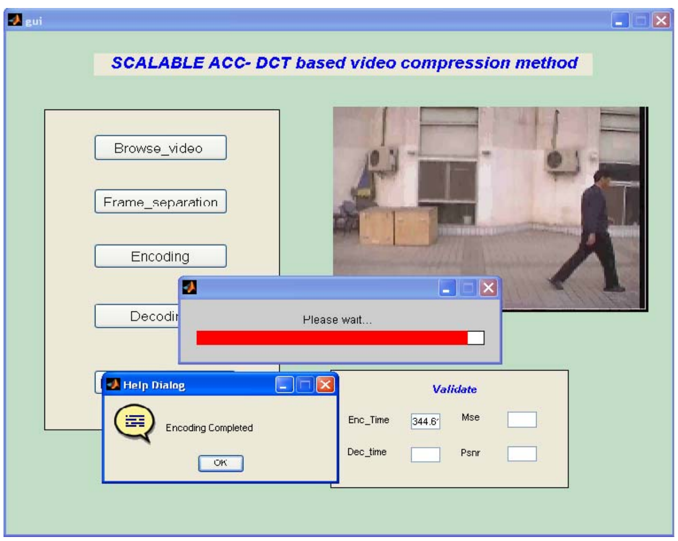

Fig.6: Encoding Model

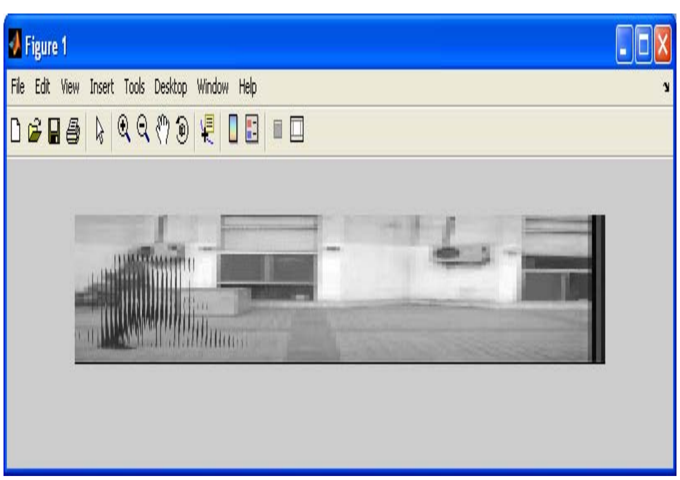

Fig.7: Accordion Representation example (Hall Monitor)

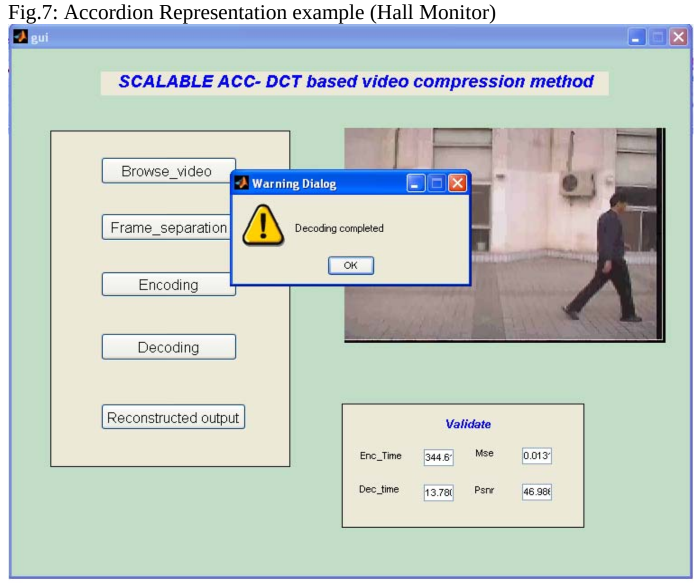

Fig.8: GUI for Decoding and Validation Process

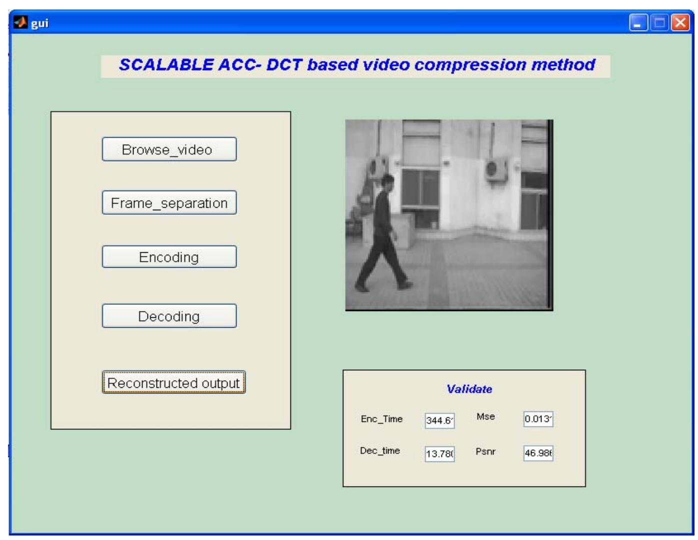

Fig.9: GUI for Reconstructed output validation

ans =

hObject: 4.0011
eventdata: []
handles: [1x1 struct]
q: 1
str1: 'frame'
str2: '.bmp'
Bitstream: {[51782x1 double]}
Bitst: 51782
j1: 2
f: 1
filename_1: '1.bmp'

Image1: [120x960 double]
row: 244
col: 356
out: [120x960 double]
Enc: [120x960 double]
r: 120
c: 960
Input_filesize: 921600
i: 120
j: 960
QEnc: [120x960 double]
ZQEnc: [1x115200 double]
Level: 8
Speed: 0
xC: {[1x115200 double]}
y: [51782x1 double]
Res: [2x4 double]
cs: 4
cc: 51782
dd: 51782
Compresed_file_size: 51782
Comp_RATIO: 71.1908
enctime: 345.6888

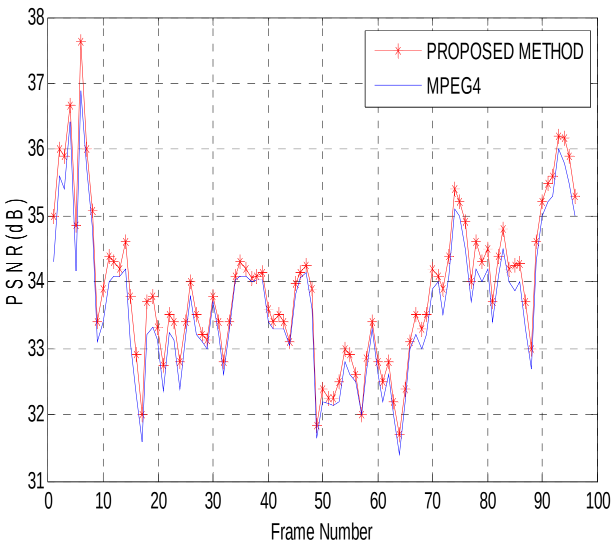

Fig.10: Frame number Vs PSNR(dB)(Hall Monitor)

The proposed method presents several advantages:

1. The proposed method transforms the 3D features to 2D ones, which enormously reduce the Processing complexity.
2. The proposed encoder and decoder are **symmetric** with almost identical structure and complexity, which facilitates their joint implementation.
3. It exploits the temporal redundancies more than the space redundancies.
4. Offers flexibility that makes it possible to be adapted to different requirements of video applications: The latency time, the compression ratio and the size of required memory depend on the value of the GOP parameter.
5. The proposed method allows the random frame access.

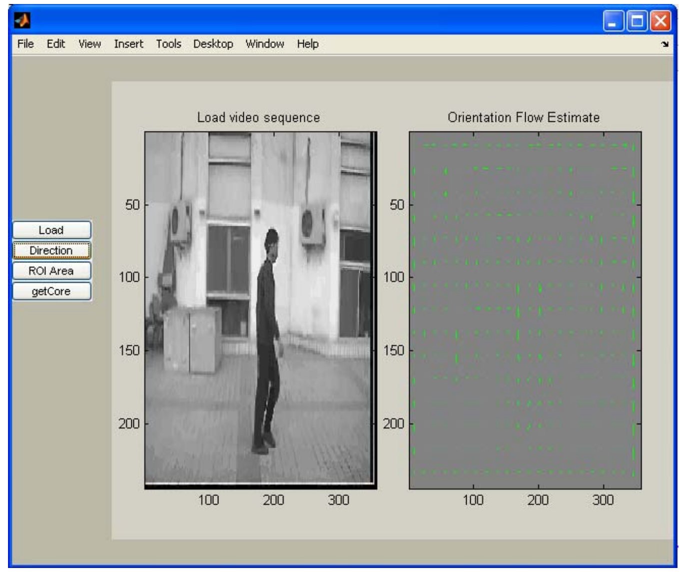

Fig.11: Orientation Flow Estimation

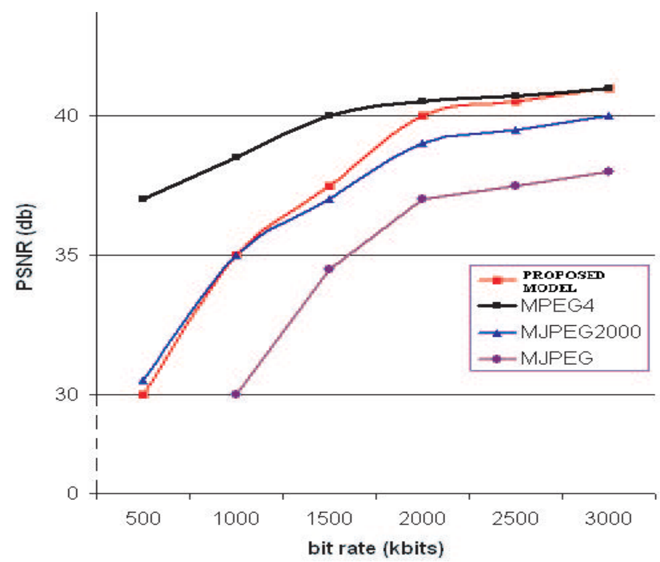

Fig.12: comparison response between bit rate Vs PSNR (dB) for different standard

5. CONCLUSION

In this paper, we successfully extended and implemented a low complex scalable ACC-DCT based video compression algorithm on MATLAB and provided experimental results to show that our method is better than the existing methods. We not only improved the coding efficiency in the proposed encoding algorithm but also it reduces complexity. As discussed in the experimental section, proposed method provides benefits of rate-PSNR performance at the good quality of base layer and low quality of enhancement layer. When SVC coding scenario meets these circumstances, proposed method should be useful. With the apparent gains in compression efficiency we foresee that the proposed method could open new horizons in video compression domain; it strongly exploits temporal redundancy with the minimum of processing complexity which facilitates its implementation in video embedded systems. It presents some useful functions and features which can be exploited in some domains as video surveillance. In high bit rate, it gives the best compromise between quality and complexity. It provides better performance than MJPEG and MJPEG2000 almost in different bit rate values. Over 2000kb/s bit rate values; our compression method performance becomes comparable to the MPEG 4 standard especially for low motion sequences. Additionally, a further development of this model could be to combine 'Accordion representation' with other transformations such as wavelet Transformation.

6. REFERENCES

[1] E. Q. L. X. Zhou and Y. Chen, .Implementation of h.264 decoder on general purpose processors with media instructions,. in *SPIE Conf. on Image and Video Communications and Processing*, (Santa Clara, CA), pp. 224.235, Jan 2003.

[2] M. B. T. Q. N. A. Molino, F. Vacca, .Low complexity video codec for mobile video conferencing,. in *Eur. Signal Processing Conf. (EUSIPCO)*,(Vienna, Austria), pp. 665.668, Sept 2004.

[3] S. B. Gokturk and A. M. Aaron, .Applying 3d methods to video for compression,. in *Digital Video Processing (EE392J) Projects Winter Quarter*, 2002.

[4] T. Fryza, *Compression of Video Signals by 3D-DCT Transform*. Diploma thesis, Institute of Radio Electronics, FEKT Brno University of Technology, Czech Republic, 2002.

[5] G. M.P. Servais, .Video compression using the three dimensional discrete cosine transform,. in *Proc.COMSIG*, pp. 27.32, 1997.

[6] R. A.Burg, .A 3d-dct real-time video compression system for low complexity single chip vlsi implementation,. in *the Mobile Multimedia Conf. (MoMuC)*, 2000.

[7] A. N. N. T. R. K.R., .Discrete cosine transforms,. in *IEEE transactions on computing*, pp. 90.93, 1974.

[8] T.Fryza and S.Hanus, .Video signals transparency in consequence of 3d-dct transform,. in *Radioelektronika 2003 Conference Proceedings*, (Brno,Czech Republic), pp. 127.130, 2003.

[9] N. Boinovi and J. Konrad, .Motion analysis in 3d dct domain and its application to video coding, vol.20, pp. 510.528, 2005.

[10] E. Y. Lam and J. W. Goodman, .A mathematical analysis of the dct coef_cient distributions for images,. vol. 9, pp. 1661.1666, 2000.

[11] *Information Technology—Coding of Moving Pictures and Associated Audio for Digital Storage Media at up about 1.5 Mbit/s: Video*, ISO/IEC 13818-2 (Mpeg2-Video), 1993.

[12] *Mpeg-4 Video Verification Model 8.0* ISO/IEC JTC1/SC29/WG11, MPEG97/N1796, 1997.

[13] *Joint Scalable Video Model JSVM-5*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-S202, Geneva, Switzerland, 2006.

[14] S. Sun, *Direct Interpolation for Upsampling in Extended Spatial Scalability*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-P012, Poznan, Poland, 2005.

[15] S. Sun, J. Reichel, E. Francois, H. Schwarz, M. Wien, and G. J. Sullivan, U*nified Solution for Spatial Scalability* Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-R018, Bangkok, Thailand, 2006.

[16] Y. Vatis, B. Edler, D. T. Nguyen, and J. Ostermann, "Motion and aliasing compensated prediction using a two-dimensional non-separable adaptive wiener interpolation filter," in *Proc. Int. Conf. Image Processing*, Sep. 2005, pp. 894–897.

[17] A. Segall and S. Lei, *Adaptive Upsampling for Spatially Scalable Coding*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-O010, Busan, Korea, 2005.

[18] A. Segall, *Study of Upsampling/Down-Sampling for Spatial Scalability* Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVTQ083, Nice, France, 2005.

[19] A. Segall, *Upsampling/Down-Sampling for Spatial Scalability*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-R070, Bangkok, Thailand, 2006.

[20] G. J. Sullivan, *Resampling Filters for SVC Upsampling*, Joint Video Team (JVT) of ISO/IEC MPEG&ITU-T VCEG, JVT-R066, Bangkok, Thailand, 2006.

Authors Details:

1.G.Suresh received the B.E Degree in ECE from Priyadarshini Engineering College affiliated to Madras University in 2000, and the M.E degree from Anna University, Chennai, in 2004. Currently he registered for Ph.D and doing research under anna university, Chennai. His research focuses on video processing, and compression techniques, Reconfigurable and adaptive logic design, and high performance clock distribution techniques.

2.P.Epsiba received the B.E degree in ECE from Ranipettai Engineering College affiliated to Madras University in 2002, and the M.E degree from Anna University, Chennai, in 2008. Currently she is working as a lecturer in C. Abdul Hakeem College of Engineering and Technology. Her research focuses on video compression and filtering techniques, Reconfigurable and adaptive logic design, and high performance clock distribution techniques.

3.Dr.M.Rajaram working as Professor and Head of the department of EEE in GCE, Thirunelveli. He is having more than twenty five years of teaching and research experience. His research focuses on networks security, image and video processing, FPGA architectures and power electronics.

4.Dr.S.N.Sivanandam working as Professor and Head of the department of CSE in PSG Tech, Coimbatore. He is having more than thirty two years of teaching and research experience. His research focuses on Bioinformatics, computer Communication, video compression and adaptive logic design.