Efficient Accordion DWT based Video Coding

N.R.Gawande
Dept. of Computer Engineering
Vishwakarma Institute of
Technology, Pune, India

P.P.Ghadekar
Dept. of Computer Engineering
Vishwakarma Institute of
Technology, Pune, India

M.L.Dhore
Dept. of Computer Engineering
Vishwakarma Institute of
Technology, Pune, India

ABSTRACT
Among all multimedia applications, transmission of video frames requires large bandwidth and more bytes for storage. To reduce transmission bandwidth and storage memory, video compression is necessary. The proposed technique compresses the video by reducing the spatial, spectral and temporal redundancies of the input video. Considerable amount of redundancy is often present between a set of frames since only a small portion of each frame is involved with any motion that is taking place. It is called temporal redundancy. Accordion is a method that converts the temporal redundancy into the spatial redundancy. Transformation of the video from 3D to 2D that allows exploring the temporal redundancy of the video using 2D transforms and avoiding the computationally demanding motion recompense step. Accordion transformation converts the spatial and temporal correlation of the video signal into a high spatial correlation. This technique transforms each group of pictures(GOP) into one picture with high spatial correlation.The main advantage of applying Discrete Wavelet Transform(DWT) is to achieve high compression while maintain reconstruction quality. Since, the large data block size is considered compared to Accordion-Discrete Cosine Transform (ACC-DCT); there will be less probability of occurrence of the blocking artifact. The compression ratio was found more for those videos having less motion and vice- versa. Many experimental tests have been conducted to prove the technique efficiency especially in high bit rate and with slow motion video. Since motion is low, temporal redundancy is high and it is expected that ACC-DWT becomes efficient.

General Terms
Video compression, temporal redundancy

Keywords
Accordion, GOP, ACC-DCT, DWT, and CR

1. INTRODUCTION
Video files are going to touch every area of information technology. Video is incorporated into applications, captured off the TV or the DVD for use as attachments to multimedia e-mail items. Video compression is a technique used to reduce redundancy and so to reduce number of bits required to transmit or store video data. Video is compressed by reducing the total number of bits by removing redundancy. Block matching techniques are the most popular and efficient of the various motion estimation techniques. Block matching algorithm in video compression select the current frame and divides into blocks. These, helps to find motion vector for each blocks within a search range find a best match that minimize an error measure. For the entire motion based video compression process, motion estimation is the most computationally expensive and time consuming process [1]. Principle of the 3D-DCT transform is based on the correlation between several frames in the temporal dimension. The 3D-DCT (Three-Dimensional Discrete Cosine Transform) [2] has a serious disadvantage if the video sequence's cut is present in the input data. In this case, the frame transparency is performed. The aim was to eliminate transparency of the different video sequences inside the input of the 3D-DCT encoder [3]. The simplest way to extend intra-frame image coding techniques to inter-frame video coding is to consider 3-D waveform coding. 3D-DCT includes the time as third dimension into the transformation and energy compaction process. 3D-DCT does not require the computationally intensive process of motion estimation. 3D DCT based video compression techniques treat video as a succession of 3D blocks or video cubes, in order to exploit the DCT properties in both spatial and temporal dimensions.

The Accordion-DCT coding technique [4] based on the same vision. Accordion-DCT technique puts in priority the exploitation of temporal redundancy, which is more important than the spatial one. The Accordion technique 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 2-DCT will be applied on 8x8 data block of an Accordion frame. Disadvantage of Accordion-DCT is Bblocking artifact may appear due to the higher numbers of blocks used. The noticeable "blocking artifacts" across the block boundaries cannot be neglected for higher compression ratio. The quality of the reconstructed images is degraded by the "false contouring" effect for specific images having gradually shaded areas. The rest of the paper is organized as follows: Section 2 gives an overview of Accordion 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 the paper with a short summary.

2. ACCORDION-DCT
Accordion representation with Discrete Cosine Transform [4] technique 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 2-DCT will be applied on 8x8 data block of an Accordion frame. Discrete Cosine Transform (DCT) is applied to reduce spatial redundancy. Normally human eyes are less sensitive to AC coefficient and highly sensitive to DC coefficient. So AC coefficient can be eliminated. The compression in DCT can only be achieved by quantization and scaling. Higher scaling factor will be needed to achieve higher compression. It degrades the reconstruction quality of each block. Blocking artifact may appear due to the higher numbers of blocks used. The noticeable "blocking artifacts" across the block boundaries cannot be neglected for higher

compression ratio. In addition, the quality of the reconstructed images is degraded by the "false contouring" effect for specific images having gradually shaded areas [5]. The false contouring occurs when smoothly graded area of an image is distorted by an aberration due to heavy quantization of the transform coefficients. The effect looks like a contour map. In order to achieve high compression while maintain reconstruction quality the proposed method is implemented. Section 3 gives the basics of the proposed method and the modifications made to improve the compression ratio and also reduces the complexity.

3. PROPOSED METHOD

A Video Sequence usually possesses high correlations among and within its frames. Exploiting the temporal and spatial correlations enables processing and better understanding of the video sequence. The input is video cube which is converted into many numbers of frames. Several works proved that the video signal variation in 3D is much less in the temporal field compared to the space variation. Then, the pixels in 3D video are more correlated spatially than temporally [4]. This was expressed as follows:

- I : Pixel intensity value
- x, y : Space coordinates of the pixel
- t : time

Assumption:

$I(x,y,t)- I(x,y,t+1)< I(x,y,t)-I(x+1,y,t)$

This assumption is the base of the representation in Accordion. It aims to put in space adjacency the pixels having a high temporal correlation. The video signal undergoes a space-time decomposition illustrated by the fig.1, the Accordion representation is built by carrying out the temporal decomposition of the video. The resulting frames are called temporal frames which will be gathered into one 2D image that collects the video pixels having the same column rank referring to the 3D representation of the video. As depicted in fig.2, the temporal frames are turned over horizontally (Mirror effect), the last stage consists of successively projecting this frame on a 2D plan called the IACC: Accordion. This projection is obtained by traversing the temporal frames while reversing the direction from one frame to another.

	1	2	3	4
1	100	125	104	111
2	120	110	250	123
3	250	112	122	110
4	122	100	94	122
5	100	98	123	115
6	123	110	122	250

Image 1

	1	2	3	4
1	123	250	112	140
2	250	122	120	112
3	123	120	123	122
4	130	133	122	250
5	98	115	96	110
6	122	122	100	11

Image 2

	1	2	3	4
1	123	105	111	100
2	123	150	234	250
3	56	110	250	132
4	100	113	123	104
5	250	114	100	112
6	122	250	111	98

Image 3

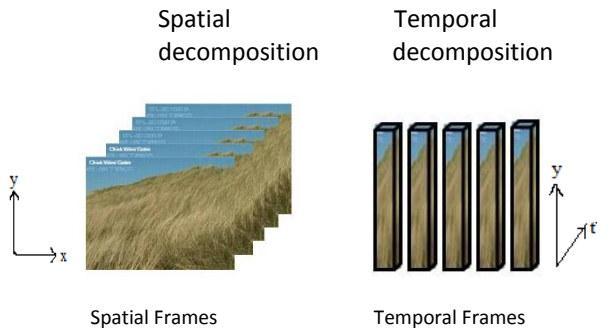


Fig 1: Video Decomposition

The purpose of changing the direction of course from a temporal frame to another is the use of spatial correlations in frames ends. However, it does not neglect the use of spatial correlations. It also minimizes the distances in the 2D representation between the pixels correlated in the source. Thus, Accordion transformation tends to transform temporal correlation in the 3D original video source into a high spatial correlation in the resulting 2D image (IACC). It aims to put in priority the exploitation of temporal correlation.

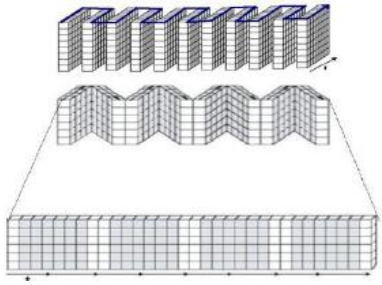


Fig 2: Accordion representation

Let us consider the example of 3 video frames to which we would like to construct the IACC: The Accordion image size (X_{acc},Y_{acc}) is given by:

$$\begin{bmatrix} X = X \\ Y = Y * N \end{bmatrix}$$

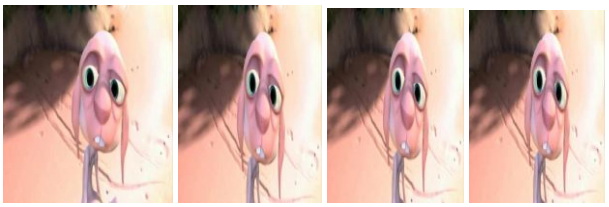
With X and Y are the frame sizes; N is the number of video frames.

	1	2	3	4	5	6	7	8	9	10	11	12
1	100	123	123	105	250	125	104	112	111	100	140	111
2	120	250	123	150	122	110	250	120	234	250	112	123
3	250	123	56	110	120	112	122	123	250	132	122	110
4	122	130	100	113	133	100	94	122	123	104	250	122
5	100	98	250	114	115	98	123	96	100	112	110	115
6	123	122	122	250	122	110	122	100	111	98	111	250

Image IACC

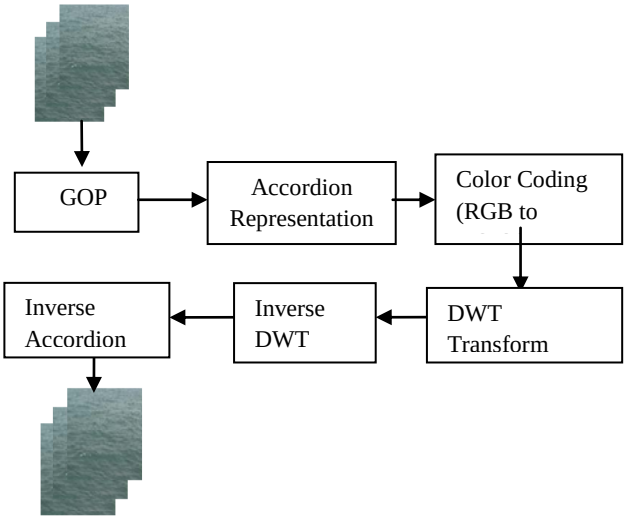
Fig 3: Accordion transformation example

For instance, in Fig.3, columns 1 and 2 are adjacent in IACC and are also temporally adjacent to the video source. Columns 3 and 4 are adjacent in IACC and are spatially adjacent in the video source. Thus, the adjacent columns in the representation IACC are also adjacent to the video source either temporally or spatially. Now, Every Accordion color video frames represented by 24 bits, (8bits for each Red, Green and Blue color) were consists of high amount of spectral redundancy. The spectral redundancy was reduced by converting the color format from RGB to YC_bC_r color format. YCbCr color space is used as the input for the next stage of DWT[6]. In Accordion, group of four consecutive frames were taken and merge to create a single frame/image. Consider the input video of size M x N x3x 500 . Fig. 4(a), (b), (c) and (d) shows four consecutive frames and Fig.4(e) shows its accordion representation. The corresponding column pixels were places adjacent to each consecutive frame's column. From input video the consecutive frames from 104 to 108 frame numbers are shown in Fig.(a), Fig.(b), Fig.(c) and Fig.(d) respectively. The accordion function was computed on these frames which represents the temporal redundancy as spatial representation.



(a) Frame 104 (b) Frame105 c)Frame 106 d)Frame 107

Fig 4: Accordion of Figure (a) to (d)
Original Video Frames



Reconstructed Video Frames

Fig 5: Flow Diagram of the proposed method

Let us note that:

1) L and H are the length and the height of the video frames.
2) N is the number of frames of a GOP (Group of Picture/Frames).

Algorithm 1: Algorithm of Accordion:

1. Decomposition of the video in GOPs (Group of frames/pictures)

2. Spatial Adjacency separation (Accordion Representation) of the GOP

N= Group of pictures (here N=9/10/11)

L= no of columns in images

m=a row vector it saves the index of columns of GOP in 1 to N or N to 1 sequence.

for i1 from 1 to 3

Here 1-10, 21-30, 41-50... Column of m allotted the value from 1-N i.e. 1 to 10

Here 11-20, 31-40, 51-60... Column of m allotted the value from N-1 i.e. 10 to 1.

```
y=1;
for j from 1 to L
        for x from  1 to N
          Iacc(:,y,i1)=I(:,j,i1,m(y));
         y=y+1;
          end
end
end
```

Algorithm 2: Algorithm of Inverse Accordion:

N= Group of pictures. (here N=10)

L= Number of columns in images.

m=a row vector it saves the index of columns of GOP

```
for i1 from 1 to 3
```

Here 1-10, 21-30, 41-50... Column of m allotted the value from 1-N i.e. 1 to 10

Here 11-20, 31-40, 51-60... Column of m allotted the value from N-1 i.e 10 to 1.

```
y=1;
for j from 1 to L1
         for x from 1 to N
        I1(:,j,i1,m(y))=Iacc1(:,y,i1);
        y=y+1;
        end
end
end
```

The 2-DWT applied on the data block of an each Accordion frame. The main advantage of applying two stage of DWT is to achieve high compression while maintain reconstruction quality. The motivation behind applying DWT is to represent the most important information in few low frequency components.

4. EXPERIMENTAL RESULTS

Many experiments had been conducted in order to study the performances of proposed technique by considering different kinds of videos. Different videos were collected from various sources. The compression performance of the proposed (ACC-DWT) technique was compared to the performance of the ACC-DCT coding technique. By studying the performances of the proposed technique with different N values, it's pointed out that N presents the number of frames of the video cube that forms the "IACC" frame. The best compression rate is obtained with N=10. By increasing the GOP's frames number there is compression improvement this is due to the exploitation of the temporal redundancies. Better compression performance can be achieved by increasing the value of N, so that the coder exploits more temporal redundancies. But by considering all type of videos proposed algorithm gives good PSNR and CR for GOP 10.

Fig.6 Shows PSNR Comparison of ACC-DWT-YC_bC_r and ACC-DCT-YC_bC_r of different GOP. Different GOP size considered. For GOP size 10 ACC_DCT, and ACC_DWT algorithm gives PSNR 31 dB, 32.6 dB, respectively. For the GOP size 10 both algorithms gives good PSNR than GOP 9 and GOP 11.

Video: 'grass.avi'

Video Parameters: 25.00 frames per second, RGB24 320x240.120 total video frames available.

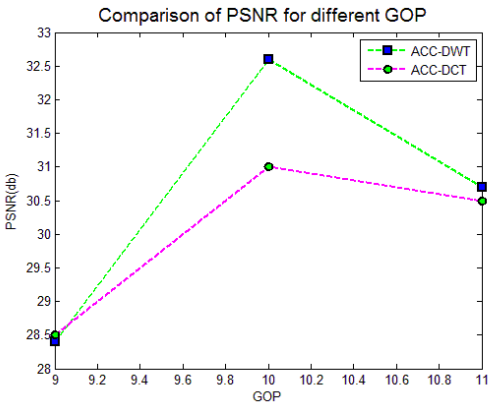


Fig 6: PSNR Comparison of ACC-DWT- YC_bC_r and ACC-DCT-YC_bC_r for different GOP

Fig.7 shows Comparison of ACC-DWT- YC_bC_r and ACC-DCT-YC_bC_r of different GOP. For GOP size 10 Accordion_DCT, and the proposed Accordion_DWT algorithm are 79.8 dB, 97.2 dB, respectively. For the GOP size 10 both algorithms gives good compression ratio.

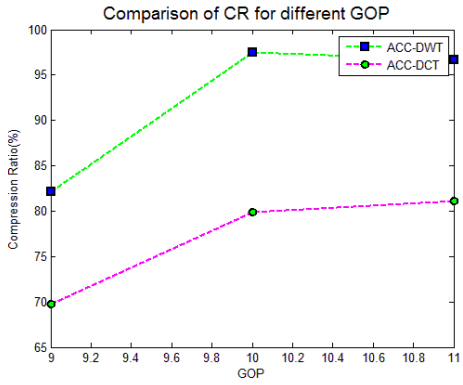


Fig 7: CR Comparison of ACC-DWT- YC_bC_r and ACC-DCT-YC_bC_r for different GOP

For demonstration video kid.avi is considered having size of 720x576.The video is compressed using the proposed(ACC-DWT-YC_bC_r, standalone DCT, DWT, ACC_DCT_RGB, ACC_DWT_RGB, ACC_DCT_YC_bC_r. Fig.8 represents the comparison of PSNR using the proposed Accordion DWT-YC_bC_r algorithm, stand alone DWT, DCT, Accordion-DCT-YC_bC_r, Accordion-DWT-RGB and Accordion-DCT-RGB algorithm. It is observed that the proposed algorithm has higher PSNR as compared to other algorithms.

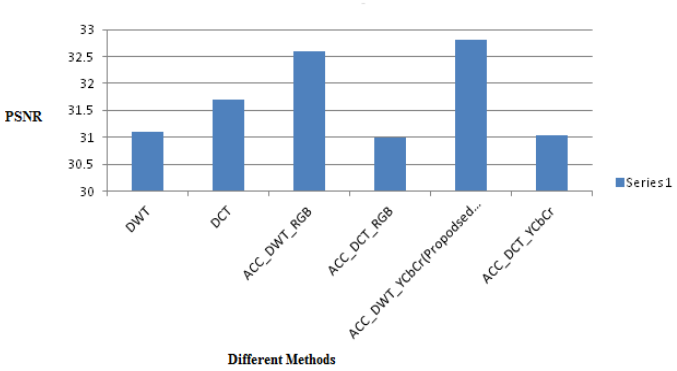


Fig.8 PSNR Comparison of all algorithms

The algorithm is also applied on several video sequences, such as, dynamic textures, videos having linear and nonlinear motion. The results are compared with the JPEG-based DCT, Acc_RGB_DWT, ACC_RGB_DWT, and ACC_YC_bC_r_DCT.

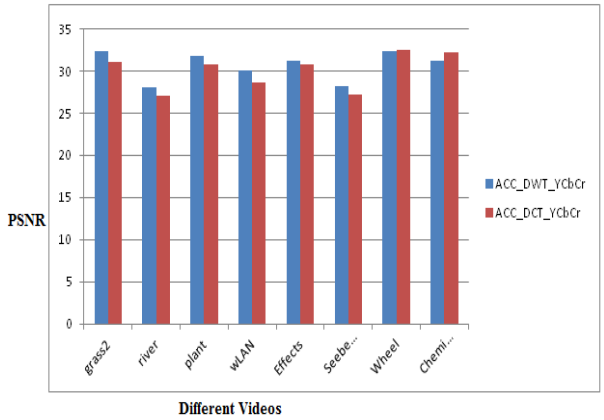


Fig.9 PSNR Comparison of different videos for proposed Accordion DWT and ACC-DCT

Among the studied videos, plant video provides better compression. The "plant" sequence contains involve less motion than the other studied sequences. Fig.11 shows the frames of 'plant' sequence along with other reconstructed frames using the proposed Accordion DWT algorithm and ACC-DCT. The PSNR using Accordion_DCT and the proposed Accordion_DWT algorithm are 32.9 dB, 29.8 dB respectively. The false contouring effect due to extreme compression is clearly visible in the frame reconstructed using the JPEG-based DCT, On the other hand, it is clearly visible that the proposed algorithm has better reconstruction quality compared to ACC-DCT method.
Video Parameters: 25.00 frames per second, RGB24 720x576x3x100.

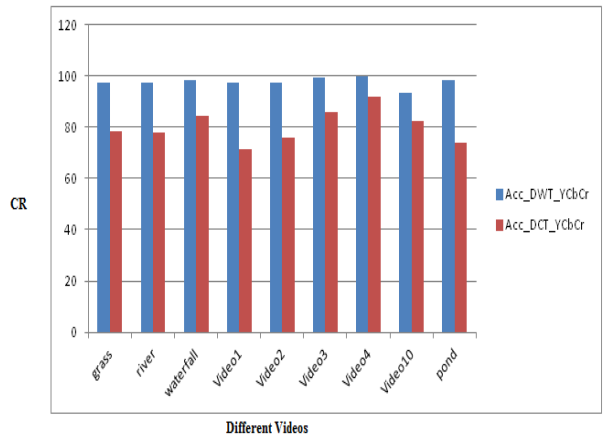


Fig.10 CR Comparison of different videos for the proposed algorithm(Accordion DWT) and ACC-DCT

Frame 10 Frame 30 Frame 50
a)Original video frames ('plant.avi')

Frame 10 Frame 30 Frame 50
b)Proposed ACC-DWT-YC_bC_r(PSNR=32.3)

Frame 10 Frame 30 Frame 50
c)ACC-DCT-YC_bC_r(PSNR=31.8)

Fig.11: PSNR Comparison (a) Original Frames of video ('plant.avi'), reconstructed frames using (b) proposed Accordion_DWT, (c) Accordion_DCT.

Fig.12 shows the some frames of 'WLAN' sequence along with other reconstructed frames using DCT, DWT and proposed schemes. The CR using Accordion_DCT and the proposed Accordion_DWT algorithm are 99.2 dB, 84.9 dB respectively. The DCT algorithm has the lowest CR value . Video Parameters: 25.00 frames per second, RGB24 720x576x3x80.

Frame 30 Frame 50 Frame 100
a) Original Video frames ('WLAN.avi')

Frame 30 Frame 50 Frame 100

b) Proposed ACC-DWT-YC_bC_r(CR=99.2)

Frame 30 Frame 50 Frame 100

c)ACC-DCT-YC_bC_r(CR=84.9)

Fig 12:CR Comparison (a) Original frames of video ('WLAN.avi'), reconstructed image using (b)Proposed Accordion_DWT,(c) Accordion_DCT.

Table 5.4 represents the comparison of PSNR and CR for Different videos using the proposed Accordion DWT algorithm and Accordion DCT algorithm. It is observed that the proposed algorithm gives higher PSNR and CR value as compared to Accordion DCT algorithm. The deviation of PSNR value between the proposed algorithm and DCT is consistent for all kind of video sequences.

Table 1.Comparison of the proposed ACC-DWT-YC_bC_r and ACC-DCT-YC_bC_r

Different Videos	Methods	PSNR(db)	CR(%)
'Wheel.avi'	ACC_DWT_YC_bC_r	32.370	97.688
	ACC_DCT_YC_bC_r	31.869	79.205
'Chemical process.avi'	ACC_DWT_ YC_bC_r	31.243	91.03
	ACC_DCT_ YC_bC_r	31.17	77.210
'Kid.avi'	ACC_DWT_ YC_bC_r	29.843	95.60
	ACC_DCT_ YC_bC_r	30.855	78.055
'Grass2.avi'	ACC_DWT_ YC_bC_r	32.414	97.916
	ACC_DCT_ YC_bC_r	31.04	81.8

4.1 Subjective quality measurement:

For the end user, the visual perception of the reconstructed image is essential. In some cases the objective quality assessment does not give proper information about the quality of the reconstructed frames. In such scenarios, it is important to analyze the reconstructed frames using subjective analysis. Hence the subjective measure is considered, viewers focus on the difference between reconstructed and original frames and correlates the differences.

Table 2. Subjective quality measurement table

Visual Quality	Very Bad	Bad	Poor	Fair	Good	Excellent
Proposed approach using ACC-DWT-YC_bC_r						
Plant	-	-	-	4	8	8
Grass	-	-	-	6	10	4
WLAN	-	-	-	6	10	4
Proposed approach using ACC-DCT-YC_bC_r						
Plant	-	-	-	4	8	8
Grass	-	-	-	8	10	2
WLAN	-	-	-	8	12	2

In the proposed technique, the DWT is exploited in temporal domain. Actually, the application of the DWT on IACC allows the transformation from the spatial domain to the frequency domain. It eliminates the high spatial frequencies of "IACC" frame which actually present the high temporal frequencies of the 3D signal source. Thus, a strong quantification does not affect the quality of image but rather affect the fluidity of the video. In one hand, ACC-DWT affects the quality of some frames of a GOP, but on the other hand, it provides relevant quality frames in the same GOP. In video compression, such feature could be useful for video surveillance field; generally, we just need some good quality frames in a GOP to identify the objects (i.e. person recognition) rather than medium quality for all the frames.

5. CONCLUSION

In this paper, an Accordion scheme with the DWT algorithms under high compression ratio constraint for video compression has been presented. The algorithm was tested on several types of videos such as, natural, medical, and dynamic textures. It was also tested on videos having different motion like linear and non-linear motion. The results of the exhaustive simulations show consistent improved performance for the accordion DWT compared to the accordion DCT. The new scheme has also reduced the false contouring effects and blocking artifacts significantly which occurs in the videos reconstructed using DCT algorithm at higher compression ratio. The analysis also showed that for a fixed level of PSNR, the number of bits required to transmit would be less than those required for other schemes. 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. The proposed method gives best results for GOP size 10 in terms of PSNR and CR. By increasing the GOP's frames number there is compression improvement this is due to the exploitation of the temporal redundancies.

It was observed that the proposed algorithm performs better than the existing algorithms. By considering different types of videos the proposed algorithm gives average PSNR 32.4db and compression ratio 98%. The best results were given with plant sequence. Since motion is low, temporal redundancy is high and it is expected that ACC-DWT becomes efficient.

6. REFERENCES

[1] N.Boinovi and J. Konrad,"Motion Estimation and Motion Compensated Video Compression Using DCT and DWT", vol.20, pp. 510.528, 2005.

[2] G. M.P. Servais, "Video Compression using the three dimensional discrete cosine transform" in Proc. COMSIG, pp. 27.32, 1997.

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

[4] G.Suresh, P.Epsiba, Dr.M.Rajaram, Dr.S.N.Sivanandam "Scalable ACC-DCT Based Video Compression Using Up/Down Sampling Method", IJCNS) International Journal of Computer and Network Security, Vol. 2, No. 6, June 2010.

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

[6] Design of a New Video Compression Algorithm Using Accordion Function International Journal of Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-1, Issue-6, May 2013

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

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