# Comparison of 3D-DWT Based Video Pre-Post Processing Techniques

V. R. Satpute, *Member, IEEE,* Ch. Naveen, Student, *Member, IEEE*, A. G. Keskar, *Member, IEEE* and K. D. Kulat, *Member, IEEE*

***Abstract*— In this paper we are going to discuss about video encoding and decoding techniques which are applied on videos before compression and after decompression respectively. Here we are using Spatial and Temporal Discrete Wavelet Transform (DWT) to encode and decode the videos. In this paper we are comparing 4 different video encoding and decoding mechanisms on the basis of time taken for encoding and decoding, system memory usage and the mean square error (MSE) obtained after reconstruction. As we are applying these mechanisms before and after the video compression they may be called as pre/post processing techniques. Here, we also discuss about Multi-Resolution Analysis (MRA) on video in temporal axis. In this paper Haar wavelet is taken as the reference due to its ease of implementation.**

***Index Terms*— Discrete wavelet transform, MRA, Video compression, Haar Wavelet, Temporal DWT**

## I. Introduction

We are in the era of fast and exponentially developing technology. From the invention of wheel to the utilization of satellite services there is exponential growth in technology which is drawing the attention of huge community of people, which in turn effecting the limited resources. As in the communication world, we are facing problems due to the limitation of the frequency band which cannot be created. So, the only solution is to use it as effectively as possible by some methods like multiplexing the data, sharing the channel, compressing the data etc. In this paper we focus on the techniques which improve the compression by keeping the base of compression techniques untouched. In the sense, we are going to discuss the pre/post processing techniques which are to be applied on the videos for achieving good compression using the existing compression techniques such as Embedded Zero Wavelet (EZW) [6], 3D-SPIHT [15,16], etc. In this paper we are not going to discuss the EZW but we will focus only on the post processing techniques and its importance in achieving high compression without changing the hardware requirements. We use concepts of Spatial [3,11] and Temporal DWT [4,11] and apply them on videos in different order which makes each mechanism different from other in the sense of memory utilization or video quality maintenance or time taken for encoding and decoding or generation of correlation between the pixels at higher levels of DWT [13]. The application of Spatial and Temporal DWT combinedly is called as 3D-DWT [7,9,10]. In section-2 we are going to discuss about the 4 different mechanism and their individual importance over other mechanisms. In section-3 the diagrammatical view of all the mechanisms are shown. In section-4 we are going to compare the mechanisms based on the time taken for encoding/decoding and MSER [12,14] obtained. In section-5 we will discuss about the results and finally we conclude our work in section-6.

## II. Four Mechanisms and their Importance

In this section let us discuss about each mechanism and its importance in detail. Here details about each mechanism are provided individually one after the other.

1. **Mechanism-1:** In mechanism-1 multi-level spatial DWT is applied and on the resultant output only single-level temporal DWT is applied [1]. The advantage of mechanism-1 is that it gives “superior compression over the normal EZW” applied on video frame by frame as it has more zeros at the second half of the video which occur due to the application of the temporal DWT. But, by limiting the temporal DWT level to one we are not allowing the more zeros to be generated, which will be the drawback of mechanism-1. The drawback of limiting the temporal DWT level to one has been lifted in mechanism-2.
2. **Mechanism-2:** In mechanism-2 we can apply multi-level spatial DWT and also multi-level temporal DWT [1]. Here, on the raw video we are going to apply the multi-level spatial DWT and on the resultant we are going to apply the multi-level temporal DWT. The advantage of this mechanism is that it will generate the “correlation between the pixels” which are at different levels of DWT both in spatial and temporal. The concept of correlation is the basic idea behind the implementation of EZW [6,8]. So, by taking this correlation into consideration we can generate more zero trees which are helpful in achieving the high compression rate. But the heavy computation’s performed on each pixel of video at the time of applying multi-level Spatial DWT and multi-level temporal DWT may lead to the degradation at the time of reconstruction.
3. **Mechanism-3:** As of mechanism-2 we can choose different level for spatial DWT and different level for temporal DWT, but here in mechanism-3 we are going to fix the same level for both spatial and temporal DWT by choosing the minimum level from both [2]. The genuine

V. R. Satpute is with Electronics Engineering Department, VNIT Nagpur, as an Assistant Professor. (E-mail: vrsatpute@ece.vnit.ac.in ).
Ch. Naveen is with Electronics Engineering Department, VNIT Nagpur. (E-mail: naveench@ieee.org ).
Dr. A. G. Keskar and Dr. K. D. Kulat are with Electronics Engineering Department, VNIT Nagpur, as Professor. (E-mail: agkeskar@ece.vnit.ac.in, kdkulat@ece.vnit.ac.in ).

reason behind selecting the minimum level and fixing it as the common level for both spatial DWT and temporal DWT is due to its property of encoding the video by applying spatial DWT and temporal DWT simultaneously one after another on the low pass obtained. Here, on the raw video we are going to apply the temporal DWT of level-1 and on the low pass frames of the resultant output we are going to apply the spatial DWT of level-1. Now again we are going to take the low pass components from the resultant output and apply temporal DWT of level-2 and again on the low pass frames of the resultant output apply the spatial DWT of level-2 and so on. We can applying temporal and spatial DWT in the same way until we get either of the number of rows or number of columns or number of frames becomes odd or one. This type of treatment leads to the generation of "Multi Resolution Analysis (MRA)" on videos in both spatial and temporal axis, which can be seen diagrammatically in figure (3). The advantage of mechanim-3 is that it is "memory efficient" than mechanism-2 and superior in generating zeros over mechnaism-1. But coming to the generation of zero trees mechanism-2 is superior to mechanism-3 due to its inherent property of generating the correlation between the pixels. But this mechanism is more beneficial when the quality of video and system memory requirement is taken into consideration. This is because of the fact that we are applying the spatial and temporal DWT simultaneously which reduce the computations on pixels of video and in this case we are applying the temporal DWT first and then taking only low pass components into consideration for applying the spatial DWT by reducing the system processing memory at further increased levels of DWT. So, there is a trade-off between generating more number of zero trees and quality of reconstruction. To minimize this trade-off between generation of zero trees and quality of reconstruction we can go for next mechanism i.e., mechanism-4.

4. **Mechanism-4:** In mechanism-4 we have facility of selecting the different level for spatial DWT and different level for temporal DWT but the way of encoding the video differs [2]. Here, we apply the multi-level spatial DWT on the raw video and on the low pass component of the resultant output we are going to apply the multi-level temporal DWT. So, here we had reduced the number of computations on each pixel of video which helps in good quality of reconstruction and the correlation between the pixels at the last level of temporal DWT are preserved. So, this technique of encoding can give the optimum quality of reconstruction with the optimum compression rate.

Hence, this mechanism can be the trade-off mechanism between mechanism-2 and mechanism-3 with the combined qualities of both the mechanisms. Now, by looking the four different mechanisms mentioned we can go for opting the suitable post processing technique which suits the application of user. For better compression rate go for mechanism-2 and for better quality go for mechanism-3 and for the optimum results one can opt for mechanism-4. All these four mechanism are summarized in Table-1.

TABLE I
SUMMARIZATION OF MECHANISM-1 TO MECHANISM-4

| | **Spatial Level (Ls)** | **Temporal Level (Lt)** | **Importance** |
|---|---|---|---|
| Mechanism-1 | X | 1 | |
| Mechanism-2 | Xs | Xt | Generation of Correlation |
| Mechanism-3 | X | X | Quality of reconstruction |
| Mechanism-4 | Xs | Xt | Provides trade-off |

Here, Xs, Xt indicates the different levels for spatial DWT and temporal DWT respectively, X represents the same level for both spatial and temporal DWT for mechanism-3 and level for spatial DWT for mechanism-1.

## III. DIAGRAMMATICAL REPRESENTATION OF THE MECHANISMS

In section-2 four different mechanisms and their importance in video post processing has been discussed. So, here let us see how the encoded video looks after applying the four different mechanisms.

1. Mechanism-1: The diagrammatical view represents the level-3 spatial DWT and single level temporal DWT in figure (1).
2. Mechanism-2: Here the diagrammatical representation is shown for level-3 spatial DWT and level-4 temporal DWT in figure (2).
3. Mechanism-3: Here in the diagrammatical representation spatial level and temporal level applied is 3 as shown in figure (3).
4. Mechanism-4: For mechanism-4 level-3 spatial DWT and level-3 temporal DWT is represented in diagrammatical representation in figure (4).

TL, TH, SL represents temporal low pass, temporal high pass, spatial DWT respectively, and1, 2, 3 represents the corresponding DWT level applied.

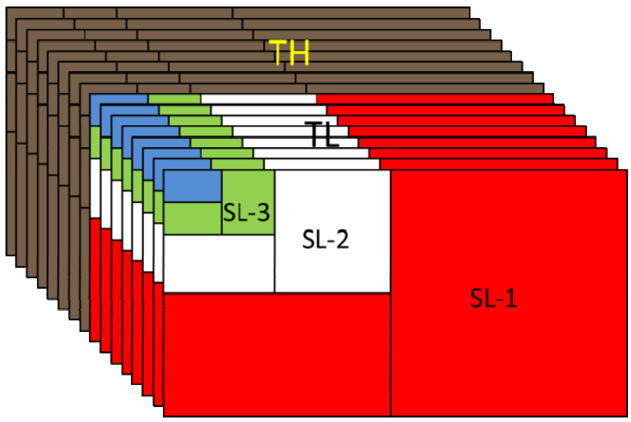



Fig. 1: Diagrammatical representation of mechanism-1 with spatial level-3 and temporal level-1

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

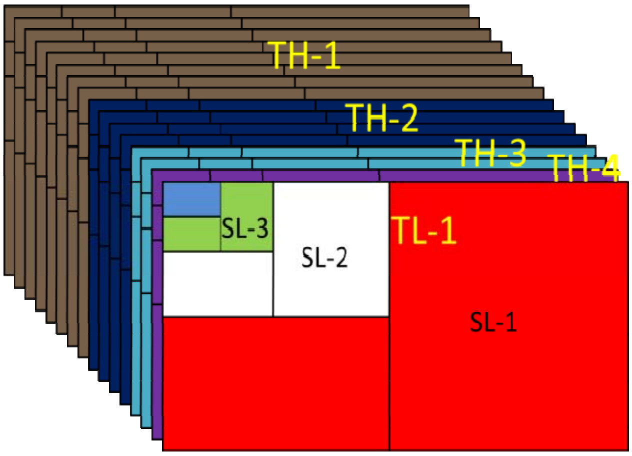



Fig. 2: Diagrammatical representation of mechanism-2 for spatial level-3 and temporal level-4

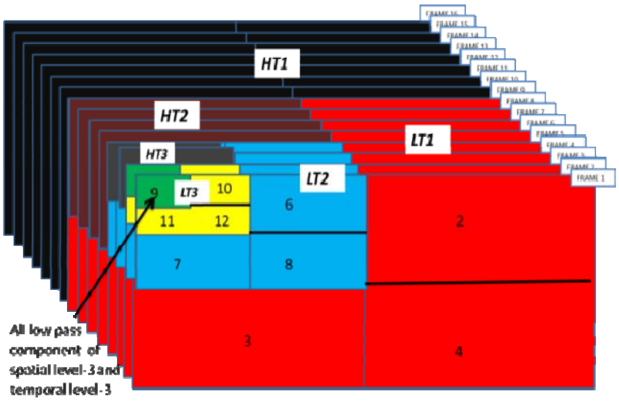



Fig. 3: Diagrammatical representation of mechanism with level-3 spatial and temporal level DWT.

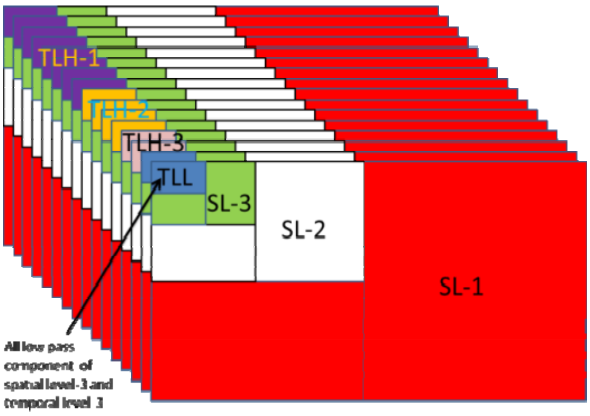



Fig. 4: Diagrammatical representation of mechanism-4 with spatial DWT level-3 and temporal DWT level-3.

## IV. Comparison of Mechanisms

For the comparison of four mechanisms which are discussed are tested on various standard videos, videos downloaded from internet and videos captured in the laboratory under non – standard conditions with a Canon Power shot A460 Digital camera having Approximately 5.3 Million Pixels 1/3.0 inch type CCD sensor with automatic exposure control and with video size of 480 by 640 pixels as given in table – 2.

Here, we are presenting the comparison between the mechanisms, based on the time taken to transform and reconstruct the video and also the MSER [12,14] after reconstruction in table 3 (for the temporal level-1 and spatial level-1) and in table 4 for multi- level DWT (for spatial level-2 and temporal level-2). The outputs presented in table-3 and table-4 are executed on a computer with the configuration: Intel i7, 3.4 GHz processor with 32 GB of RAM and 1 TB of HDD. ("NA" for table-4 means Not Applicable)

TABLE II
List of Videos Used for Testing

| Video No. | Size | No. of Frames/ sec | TIME (in sec) | Total number of frames |
|---|---|---|---|---|
| Video 1 | 120x160 | 15 | 8 | 120 |
| Video 2 | 480 x 640 | 10 | 14 | 140 |
| Video 3 | 480 x 640 | 10 | 13 | 128 |
| Video 4 | 240 x 320 | 29 | 18 | 540 |
| Video 5 | 576x1024 | 25 | 10 | 256 |
| Video 6 | 1080x1720 | 20 | 6 | 128 |
| Video 7 | 144x176 | 30 | 13 | 392 |
| Video 8 | 240x320 | 14 | 37 | 512 |

Table-3 spatial level-1 and temporal level-1

| | *Time & MSER* | *Mechanism -1* | *Mechanism -2* | *Mechanism -3* | *Mechanis m-4* |
|---|---|---|---|---|---|
| **Vide o-1** | DWT | 0.143261 | 0.170923 | 0.161118 | 0.173486 |
| | IDWT | 0.107609 | 0.142283 | 0.133975 | 0.130619 |
| | MSER | 3.529e-31 | 3.529e-31 | 3.529e-31 | 3.529e-31 |
| **Vide o-2** | DWT | 2.416103 | 2.967789 | 2.578908 | 2.751292 |
| | IDWT | 1.853454 | 2.425089 | 2.214391 | 2.339750 |
| | MSER | 3.479e-31 | 3.479e-31 | 3.479e-31 | 3.479e-31 |
| **Vide o-3** | DWT | 2.321458 | 2.749909 | 2.422879 | 2.611065 |
| | IDWT | 1.98654 | 2.304075 | 2.043316 | 2.204896 |
| | MSER | 3.413e-31 | 3.413e-31 | 3.413e-31 | 3.413e-31 |
| **Vide o-4** | DWT | 1.965847 | 3.012332 | 2.654446 | 0.691840 |
| | IDWT | 1.956475 | 2.493011 | 2.233568 | 0.576039 |
| | MSER | 5.407e-31 | 5.407e-31 | 5.407e-31 | 5.407e-31 |
| **Vide o-5** | DWT | 9.43214 | 10.733898 | 9.547196 | 9.924238 |
| | IDWT | 8.21456 | 9.117467 | 8.148449 | 8.387920 |
| | MSER | 4.841e-31 | 4.841e-31 | 4.841e-31 | 4.841e-31 |
| **Vide o-6** | DWT | 15.02134 | 18.458442 | 15.843601 | 17.91986 |
| | IDWT | 14.23456 | 15.496419 | 14.336402 | 15.30472 |
| | MSER | 6.887e-33 | 6.887e-33 | 6.887e-33 | 6.887e-33 |
| **Vide o-7** | DWT | 0.558697 | 0.824465 | 0.669313 | 0.734134 |
| | IDWT | 0.499871 | 0.606323 | 0.549009 | 0.604860 |
| | MSER | 1.525e-30 | 1.525e-30 | 1.525e-30 | 1.525e-30 |
| **Vide o-8** | DWT | 1.27456 | 1.417694 | 1.281702 | 1.418934 |
| | IDWT | 1.024578 | 1.155950 | 1.074151 | 1.144363 |
| | MSER | 5.364e-31 | 5.364e-31 | 5.364e-31 | 5.364e-31 |

Table-4 spatial level-2 and temporal level-2

| | *Time & MSER* | *Mechan ism-1* | *Mechanism -2* | *Mechanism -3* | *Mechanis m-4* |
|---|---|---|---|---|---|
| **Video-1** | DWT | NA | 0.246506 | 0.173512 | 0.215467 |
| | IDWT | NA | 0.198421 | 0.136978 | 0.172915 |
| | MSER | NA | 8.025e-31 | 8.104e-31 | 8.025e-31 |
| **Video-2** | DWT | NA | 4.279791 | 3.054225 | 3.841134 |
| | IDWT | NA | 3.482500 | 2.487691 | 3.127823 |
| | MSER | NA | 9.340e-31 | 9.941e-31 | 9.340e-31 |
| **Video-3** | DWT | NA | 3.984854 | 2.969299 | 3.784854 |
| | IDWT | NA | 3.246693 | 2.289465 | 2.909998 |
| | MSER | NA | 9.336e-31 | 9.413e-31 | 9.336e-31 |
| **Video-4** | DWT | NA | 4.438583 | 3.082685 | 0.973292 |
| | IDWT | NA | 3.573375 | 2.507201 | 0.753774 |

| | *Time & MSER* | *Mechanism-1* | *Mechanism-2* | *Mechanism-3* | *Mechanism-4* |
|---|---|---|---|---|---|
| | MSER | NA | 1.604e-30 | 1.656e-30 | 1.604e-30 |
| **Video-5** | DWT | NA | 15.457898 | 11.185289 | 14.649464 |
| | IDWT | NA | 13.124028 | 8.850935 | 11.923820 |
| | MSER | NA | 1.077e-30 | 1.172e-30 | 1.077e-30 |
| **Video-6** | DWT | NA | 28.691505 | 20.432294 | 26.134293 |
| | IDWT | NA | 22.641633 | 16.061131 | 0.178058 |
| | MSER | NA | 1.508e-32 | 6.887e-33 | 1.508e-32 |
| **Video-7** | DWT | NA | 1.160209 | 0.777375 | 0.993300 |
| | IDWT | NA | 0.864553 | 0.618469 | 0.788771 |
| | MSER | NA | 3.474e-30 | 3.487e-30 | 3.474e-30 |
| **Video-8** | DWT | NA | 2.049021 | 1.493861 | 1.910256 |
| | IDWT | NA | 1.669597 | 1.213583 | 1.534415 |
| | MSER | NA | 1.084e-30 | 1.107e-30 | 1.084e-30 |

## V. Results

Results for various videos are discussed in this section. Figure 5,6,7,8 and 9 (part b,c,d) shows results for video-1( Frame # 35) with, (Xt=2, Xs=2), video-3 (Frame # 20) with (Xt=2, Xs=2), video-5( Frame # 51) with, (Xt=4, Xs=4), video-7 (Frame # 61) with (Xt=3, Xs=3), video-8 (Frame # 121) with (Xt=1, Xs=1), respectively. For all these cases, Figure 5(a), 6(a), 7(a), 8(a), 9(a) is for (Xt=1, Xs= as specified for part (b,c,d) of respective figure).

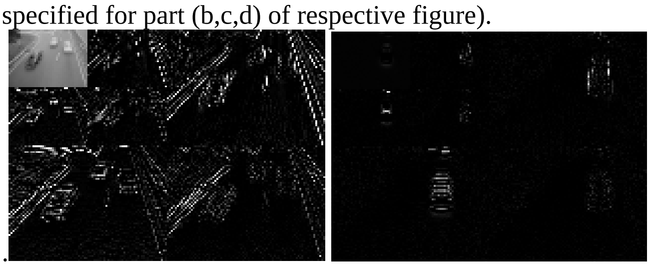


Fig 5(a): Mechanism-1 Fig. 5(b): Mechanism-2

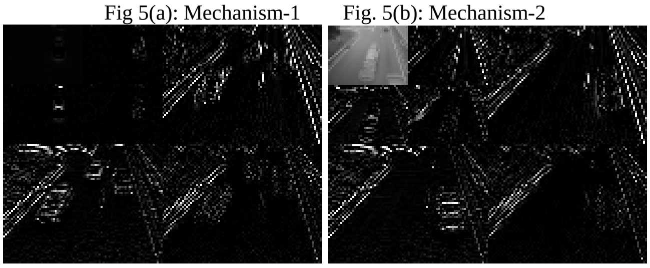


Fig 5(c): Mechanism-3 Fig. 5(d): Mechanism-4

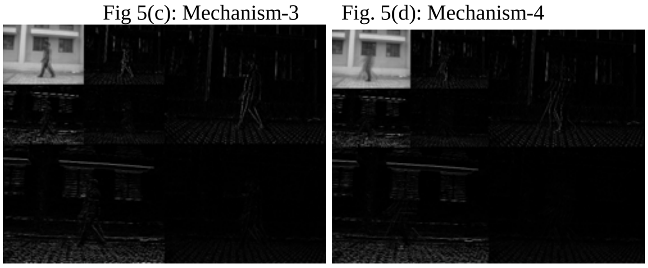


Fig 6(a): Mechanism-1 Fig. 6(b): Mechanism-2

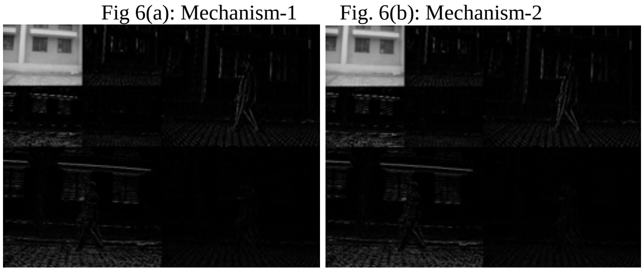


Fig 6(c): Mechanism-3 Fig. 6(d): Mechanism-4

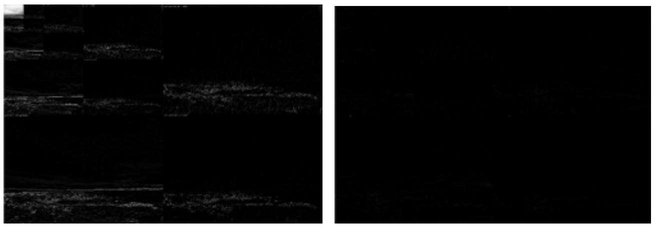


Fig 7(a): Mechanism-1 Fig. 7(b): Mechanism-2

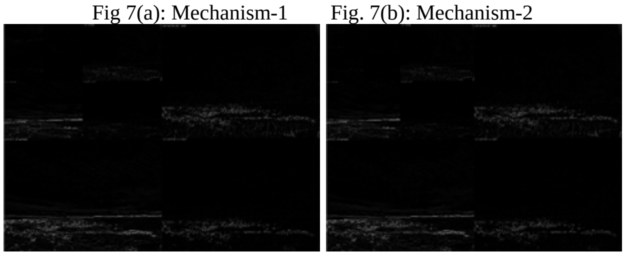
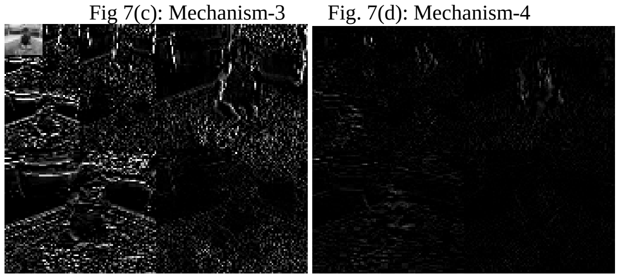


Fig 7(c): Mechanism-3 Fig. 7(d): Mechanism-4



8(a): Mechanism-1 Fig. 8(b): Mechanism-2

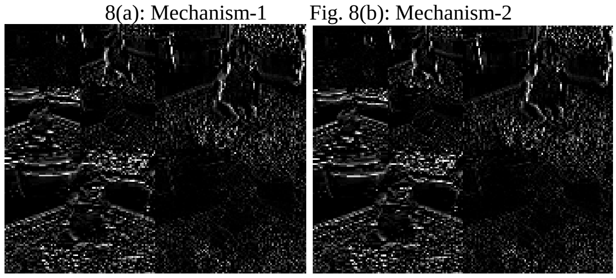


Fig 8(c): Mechanism-3 Fig. 8(d): Mechanism-4

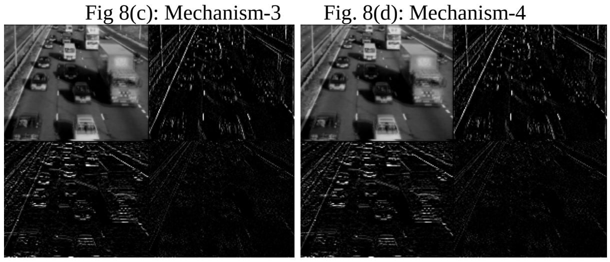


Fig 9(a): Mechanism-1 Fig. 9(b): Mechanism-2

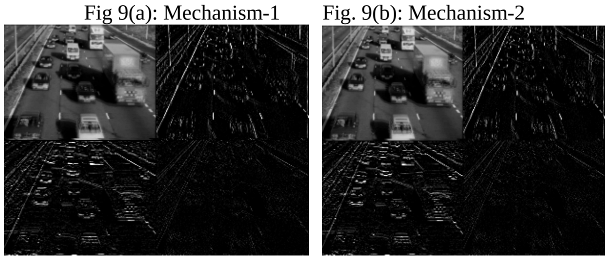


Fig 9(c): Mechanism-3 Fig. 9(d): Mechanism-4

## VI. Conclusion

In this paper, the four video pre/post processing mechanisms which can yield high compression ratios using the existing compression techniques have been compared and presented. From figure (3) one can observe the MRA on video in both spatial and temporal axis i.e., both on frames and in each frame. This principle of generating MRA will also be much beneficial and helpful when we are applying 3-D based video compression techniques such as 3D-SPIHT, EZW, EBCOT, etc.

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

## REFERENCES

[1] Ch. Naveen, V. R. Satpute, K. D. Kulat, A. G. Keskar,” Video Encoding Techniques Based on 3D-DWT”, 2014 IEEE Students’ Conference on Electrical, Electronics and Computer Science, SCEECS 2014, March 1-2, 2014.

[2] Ch. Naveen, V. R. Satpute, A. G. Keskar, K. D. Kulat,” Fast and Memory Efficient 3D-DWT Based Video Encoding Techniques”, Proceedings of the International Multi Conference of Engineers and Computer Scientists 2014 Vol. I, IMECS 2014, March 12 - 14, 2014, Hong Kong, pp.- 427-433.

[3] K. Sureshraju, V. R .Satpute, Dr. A. G. Keskar, Dr. K. D. Kulat, “Image Compression using wavelet transform compression ratio and PSNR calculations”, Proceedings of the National Conference on Computer society and informatics- NCCSI’12, 23rd &24th April 2012

[4] Nagita Mehrseresht and David Taubam, “An Efficient content adaptive motion-compensated 3D-DWT with enhanced spatial and temporal scalability”, IEEE Transactions on Image processing, VOL.15, No.6, JUNE 2006

[5] Shadi Al Zu’bi, Naveev Islam and Maysam Abbod, “3D Multi resolution analysis for reduced features segementation of medical volumes using PCA”, 978-1-4244-7456-1/10/$26.00 © 2010 IEEE.

[6] Jerome M. Shapiro,”Embedded Image Coding Using Zero Trees of Wavelet Coefficients”, IEEE Transactions on signal processing VOL. 41 No.12 December 1993.

[7] Michael weeks and Magdy Bayoumi, “3-D Discrete Wavelet Transform Architectures”,0-7803-4455-3/98/$10.00 ©1998 IEEE.

[8] G.Liu and F.Zaho, “Efficient compression algorithms for Hyperspecral Images based on correlation coefficients adaptive 3D zero tree coding”, published in IET Image Processing, doi:10.1049/ict-ipr:20070139.

[9] Richard M.Jiang, Danny Crookes,”FPGA implementaiotn of 3D Discrete Wavelet Transform for real-time medical imaging”, 1-4244-1342-7/07/$25.00 © 2007 IEEE.

[10] O.Lopez, M.Martinez-Rach, P.Pinol, M.P.Malumbers, J.Oliver, “A fast 3D DWT video encoder with reduces memory usage suitable for IPTV”, 978-1-4244-7493-6/10/$26.00 © 2010 IEEE.

[11] Anirban Das, Anindya Hazra, and Swapna Banerjee,“An Efficient Architecture for 3-D Discrete Wavelet Transform”, IEEE transactions on circuits and systems for video technology.

[12] Wikipedia: http://en.wikipedia.org/wiki/Mean_squared_error. Dated, 15th December, 2013.

[13] K. P. Soman, K. I. Ramchandran, N. G. Resmi, "Insight into wavelets: from theory to practice", Third Edition, Prentice Hall of India.

[14] Hui Li Tan, Zhengguo Li, Yih Han Tan, Susanto Rahardja, Chuohuo Yeo, "A Perceptually Relevant MSE-Based Image Quality Metric", IEEE Transactions on Image Processing, Vol. 22, No. 11, November 2013, pp- 4447 - 4459.

[15] Beong-Jo Kim, Zixiang Xiong, William A. Pearlman, "Low Bit-Rate Scalable Video Coding with 3-D Set Partitioning in Hierarchical Trees (3-D SPIHT)", IEEE Transactions on Circuits And Systems For Video Technology, Vol. 10, No. 8, December 2000, pp-1374 - 1387.

[16] Sungdae Cho, William A. Pearlman, "A Full-Featured, Error-Resilient, Scalable Wavelet Video Codec Based on the Set Partitioning in Hierarchical Trees (SPIHT) Algorithm", IEEE Transactions on Circuits And Systems For Video Technology, Vol. 12, No. 3, March 2002, pp-157 - 171.

ISBN: 978-988-19252-0-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)