International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.4 (2024) pp. 1757-1764
http://www.ijcesen.com

Copyright © IJCESEN

ISSN: 2149-9144

Research Article

A Lossy Video Compression Technique for High Quality Videos Using 3D-Biorthogonal Wavelet Transform

Sravanthi CHUTKE[1]*, N.M. NANDHITHA[2], Praveen Kumar LENDALE[3]

[1]Sravanthi Chutke, Research Scholar, Electronics and Communication, Sathyabama Institute of Science and Technology-Chennai
* **Corresponding Author Email:** chutkesravanthi26@gmail.com **ORCID:** 0000-0001-5922-7398

[2]Dr. Nandhitha N.M, Dean School of EEE , Electrical and Electronics, Sathyabama Institute of Science and Technology-Chennai
Email: nandhitha.ece@sathyabama.ac.in **- ORCID:** 0000-0002-5256-708X

[3]Praveen Kumar Lendale, Research Scholar, Electronics and Communication, Sathyabama Institute of Science and Technology-Chennai
Email: praveenkumar.lendale@gmail.com **- ORCID:** 0000-0002-4913-8137

Article Info:

DOI: 10.22399/ijcesen.698
Received : 27 November 2024
Accepted : 16 December 2024

Keywords:

Transform,
Scalar Quantization,
Huffman encoding,
Compression Ratio,
MSE,
PSNR.

Abstract:

This paper presents a completely new range-forward 3D video compression algorithm based on the combination of 3D Biorthogonal Wavelet Transform (3D-BWT), scalar quantization and Huffman coding allowing compression and decompression of video with high quality. Spatial and temporal correlations are captured through the application of multi-resolution representations which are derived by the 3D-BWT in video data decomposition. This is followed by the application of scalar quantization that reduces the precision in which transformed coefficients were obtained, this results into extreme compression while quality degradation is controlled at an acceptable level. The quantization approach is best achieved using the Huffman coding scheme. The encoded coefficients optimize the bitstream and are well suited for transmission or storage. The transmission or storage of encoded coefficients is optimized using Huffman coding. The process of 3D-BWT inverse is then used alongside the dequantization process and Huffman decoding for video decompression. The proposed technique has been demonstrated through experimental results to improve on existing techniques with respect to a number of quality metrics including compression ratio, mean squared error, and peak signal-to-noise ratio. The evaluation confirms that compared to other approaches, the proposed approach performs better by achieving improved overall video performance whilst its efficiency in compression is high making it applicable for 3D video compression applications.

1. Introduction

Owing to the scalable use of new technologies, which include 3D videos, the need for data compression has exponentially increased. The burst of usage of today's 3D video applications such as, virtual reality (VR), augmented reality (AR) technologies, systems including 3D cinema and 3D gaming applications, which uses and processes a lot more data than their 2D counterparts, information with depth is combined to a 3D video [1]. The challenge of compressing 3D videos occurs from increasing size of volume, the requirement to preserve spatial and temporal features over the volumes as well as the three-dimensional structure. Since 3D-video content is generally concerned with multiple views as well as multiple Z-depths, efficient compression strategies must strike a compromise between the amount of data loss sustained and data compression efficiency so as not to lose too much visual quality [2].

In video compression, lossy compression methods are often employed for file size reduction although some video quality is lost in the process. Lossy compression can compress videos down to a fraction of the original size while still providing acceptable level of quality, and thus lossy algorithms are very often utilized in the compression of images and videos[3]. The difficulty comes from the need to optimize compression so as not to lose the important elements that are needed to achieve the desired sense of realism in the 3D video. Aiming to overcome this problem, this paper presents a new approach to the lossy compression of 3D video by using a combination of 3D Biorthogonal Wavelet Transform (3D-BWT), scalar quantization and Huffman coding that are together efficient and enable reconstruction of good quality 3D video[4].

There are several techniques discussed in the existing literature for compressing 3D video and each of them takes a different angle towards the problem.

One of the common techniques is to make use of wavelet transforms as they provide multitiered analysis of the video data in time and space. For example, numerous studies have concentrated on compressing images and videos through the use of 2D discrete wavelet transforms (DWT). In the case of 3D video, it is equally necessary to expand the elaborated techniques towards the temporal dimension [5].

The 3D Discrete Wavelet Transform (3D-DWT) has also been investigated for compression purposes by treating video frames as three-dimensional data blocks encasing their depth properties. Nonetheless, even if 3D-DWT can be convenient to achieve the goal, it is still computationally intensive and is not very efficient in compression ratios [6]. Apart from the wavelet-based methods, motion-compensated methods also find wide applications in video compression. The techniques like 3D motion estimation help in predicting the temporal motion that exists from the moving images, which aids in minimizing temporal redundancies. However, these methods may fall short in the situations where movements between frames are arbitrary or where large volumes of spatial data are concerned, rendering them unsuitable for 3D content which is typically more precise in nature. It is common to apply Huffman coding, which is a lossless compression technique for encoding the transformed coefficients in an efficient manner [7].

However, the success of Huffman coding greatly correlates with the quantization step which is done before it, as improperly quantized data will result in poorly compressed images [8].

While these conventional methods are effective in many cases they cannot provide the required compression ratio without a massive degradation in quality in most of the cases involving complex 3D-video [9].

In the paper, section 2 consists of proposed methodology followed by experimental results in section 3 and finally paper concludes in section 4.

2. Proposed Methodology

Here, we describe a new method for compression of 3D video which is lossy and makes use of 3D Biorthogonal Wavelet Transform, scalar quantization and Huffman coding. The focus is to ensure that very high compression ratios can be achieved but at the time acceptable quality is attained for real time applications such as streaming and storage of the 3D video content. Video data cubes are processed with 3D-BWT by sequentially applying wavelet transformations in each dimension as shown in figure 1. Each of the proposed components of the approach are described here and how these pieces fit into the larger scheme of compression is explained and is shown in figure 2.

2.1 3D-Biorthogonal Wavelet Transform (3D-BWT)

There exists quite a number of researches where wavelets have been used to compress image data and also video data, in that case the extension of three dimensions will also be important for video compression. In our case, instead of the Biorthogonal Wavelet, we use the 3D-Biorthogonal Wavelet Transform (3D-BWT). The second one, apart from spatial (horizontal and vertical) axes, also makes time axis of the video visible for a 2D wavelet expansion. In multiple dimensions, the 3D-BWT cross-sections arbitrary video into several layers of motion picture volume in space for:

1. Spatial dimensions (horizontal and vertical), concentrating on details inside a single frame.
2. Temporal dimension, pertaining to the interplay of frames in the sequence.

Due to its unique traits like faithful reconstruction, symmetry, and effective multi-resolution analysis, the biorthogonal wavelet is chosen. These properties are important in the aspect of compression and preservation of quality.

In this way, low-frequency elements that include most of the video energy can be separated from the high-frequency elements that capture small details. The low-frequency elements are more useful and are kept for further processes, while the high-frequency components are crammed up for they are not prone to human perception. This step eclectically reduces both spatial and temporal redundancy and thus sets the stage for quantization in the next stage. As a result of this transform, the bulk of the video material is represented as approximation coefficients, while detail coefficients depict the high-frequency detail. Certain components can be made for lossy compression since 3D-BWT decomposition provides greater flexibility for

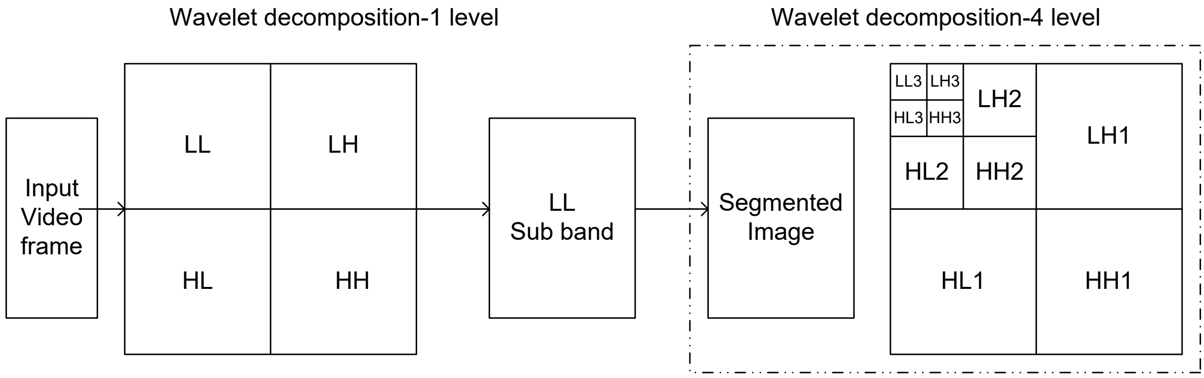

***Figure 1.** 3D BWT Wavelet Decomposition of video Frames*

removing unimportant features. The decomposition of the Three Dimensional Biorthogonal Wavelet Transform in the spatial domain for each frame can be summarized as in eq 1:

$$V(x,y,t) = V_{LL}(x,y,t) + V_{LH}(x,y,t) + V_{HL}(x,y,t) + V_{HH}(x,y,t) \qquad (1)$$

Where:

• The subbands where $V_{LL}(x,y,t), V_{LH}(x,y,t), V_{HL}(x,y,t), V_{HH}(x,y,t)$ are produced are formed by applying 2D biorthogonal wavelet filters in horizontal and vertically.

• This decomposition is performed in an independent frame by frame manner within the spatial domain.

This formulation plays a part in video compression because it makes use of the high-frequency noise and low-frequency information combination, which improves the compression together with quality. The time dimension (video frames) will be addressed later in the three dimensional wavelet decomposition that includes both spatial and time components.

For the given video sequence frames, every spatial sub band $V_{sub}(x,y,t)$ is subjected to decomposition over temporal domain.

$$V_{sub}(x,y,t) = V_{LL_T}(x,y,t) + V_{LH_T}(x,y,t) + V_{HL_T}(x,y,t) + V_{HH_T}(x,y,t) \qquad (2)$$

Where:

- $V_{LL_T}(x,y,t)$ is for time sequences that have a low-frequency band.
- $V_{LH_T}(x,y,t)$ is for enclosed mixed frequencies i.e neither low nor high.
- $V_{HL_T}(x,y,t)$ includes aspects with high-frequency elements.
- $V_{HH_T}(x,y,t)$ consists of high-frequency components that change quite fast.

2.2 Scalar Quantization

The authors state that when the coefficients VHH, VLH, LHH, and LLL of the video information are received, visualization is performed throughout the components. From the original data three-dimensionally projected coefficients are frequently utilized. Once the video data has undergone transformation by the 3DBWT, the resulting coefficients are subjected to scalar quantization. Such quantization is crucial in most compression algorithms. As continuous values must be mapped to several discrete values, quantization plays an important role in most compression algorithms. Considerable values of the quantized values ranges from 3D BWT applied to the wavelet coefficients established in a wide range so as it aids in reducing the data size more than significantly.

In scalar quantization, this involves taking each of the wavelet coefficients and mapping them to a single discrete value taken from a fixed number of quantization levels. The compression degree and the quality loss are inversely related to the level of quantization. An increase in the number of quantization levels will lead to greater accuracy and less loss. On the other hand, a decrease in the number of levels leads to higher compression but significantly greater quality loss. The main conflict is to achieve sufficient compression ratio but provide good video quality at the same time.

Scalar quantization can be expressed mathematically as:

$$Q(x) = round\left(\frac{x - r_{min}}{\Delta}\right).\Delta + r_{min} \qquad (3)$$

Where:

- Q(x) is the quantized value for input x,
- x is an input value (such as a wavelet coefficient),
- r_{min} is the lower boundary of the quantization area,
- Δ is the quantization step size i.e. the scale factor of every interval,
- round (·) is a mapping operation onto the nearest quantization value.

The quantization stage described in our method is applied equally for all the wavelet coefficients, however adaptive methods may also be employed in subsequent works to measure the importance of

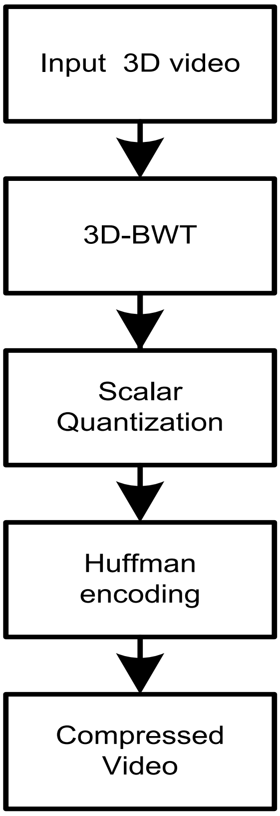

Figure 2. *Shows Compression steps of Proposed methodology.*

different coefficients and quantize them accordingly. For instance, the quantization level on coefficients depicting the low frequency components (the rough video resolution) imaginatively speaking should be high while the high frequency components that apparently contain finer details can have a lower quantization level.Using scalar quantization on the 3D wavelet coefficients results in a considerable amount of data compression with acceptable visual quality so that the processed or reconstructed video looks normal. Although quantization is a lossy process, the degree of its losses is carefully managed so that the losses while taking into account the most important elements in perception are maintained.

2.3 Huffman Coding

The coefficients obtained from the quantization process are further transformed through the use of the well-known entropy code called Huffman coding. Simply put, Huffman coding uses variable-length codes to assign shorter codes to more common symbols and longer codes to less common symbols. The Huffman coding method takes advantage of the statistical redundancy present in the quantized data, which commonly has a few values that are more prevalent than the rest.

The coefficients that are quantized do not have all ranges having the same likelihood of occurrence with sharpening of some values more than others. In solving this problem, the huffman coding technique matching the frequencies of the values whereby the shorter codes are assigned to the common values and the longer codes to the rare ones so as to accommodate the lesser value frequencies which significantly lower the compression bitstream's size. The method gets to be upon the quantized coefficients though lossy but attempts to compress more than the amount of information that is damaged. Because such a transformed compressed video takes up a limited amount of disk space, it becomes easier to transmit and store videos and their sequences. In addition, huffman coding is without loss which implies that after decoding the coefficients, the coefficients are quantized as they were before and hence the decompression process is effective.

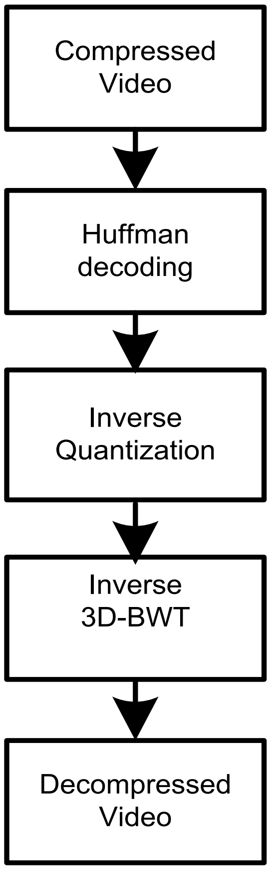

Figure 3. *Shows Decompression steps of Proposed methodology*

2.4 Process for the Decompression:

The order in which the video was decompressed serves as a guide of how to retrieve the original video image with a minimal degree of quality impairment and is shown in figure 3. The steps are as follows:

• The first thing is to remove the noise in the wavelet compressed images and secondly, Huffman decoding along with the scanning quantized images. The first decoded images are the quantized wavelet coefficients.

• Inverse quantization: The model has performed quantization which now needs to be reversed. There are discrete codes and the task is to map them to the continuously-valued coefficients of the original signal. Although some information is lost during quantization, this step is intended to not deviate too much from the original representation.
• Inverse 3D-BWT: Lastly, the 3D-BWT inverse is performed on the dequantized coefficients to synthesize the video frames. The inverse operation brings back the spatial and the temporal correlations thus reconstructing the video sequence from the bandwidth-efficient representation.
Compression methods and techniques in general contribute to some loss in quality of the final product, however, the present methods work in a way that the spatial and temporal coherence that are essential would be maintained to a reasonable and acceptable level. Application of 3D-BWT makes sure that for even a lossy bitstream, significant parts of the video is secured.
The suggested approach has some benefits when compared with 3D video compression schemes that are available. First, since 3D-BWT is employed, it is able to exploit spatial and temporal redundancies in the video more effectively and achieves better compression ratios than the former 2D Without losing sight of the quality, and second, the data is more accumulate compressed thanks to the integration of scalar quantization and Huffman coding owing to the fact that this combination is able to create a professional-center distortion metric. Finally, the reduction of scalar quantization and Huffman coding allow the system to be efficient enough for real time 3D video applications.
In the future, further research could be directed towards more efficient use of the adaptive quantization that relies on a variable quantization step size that is set relative to the importance of the coefficients in the image. Another interesting approach is a motion compensated 3D transform which could increase the compression efficiency of 3D video where a lot of motion is present.

3. Experimental Results

To assess the efficiency of the developed 3D video compression approach, a experiment was carried out on a randomly chosen 3D video sequence. The video contains the series of multiple frames representing a scene with a 3D depth and was subjected to varying degree of compression strategies. The outcomes were measured and compared with four existing approaches that were most commonly used and same is shown in table 1.

a. 3D Discrete Wavelet Transform (3D-DWT) combined with Scalar Quantization and Huffman Coding (Method 1)
b. Wavelet Transform Motion Compensated Temporal Filtering (MCTF) (Method 2)
c. H.264/AVC video compression encoders (Method 3)
d. HEVC Video Compression (also known as H.265) (Method 4)

The evaluation metrics used for comparison include Peak Signal-to-Noise Ratio, Signal-to-Noise Ratio, Mean Squared Error, and Compression Ratio.
These outcomes are cohere in performing two functions in that they were derived from the compressed and decompressed video frames so as to evaluate both the quality and the performance of the compression.

3.1 Test Setup:
The experiments are conducted in MATLAB software. There were 100 frames in the 3D video sequence but only those of 640x480 resolution and also holding depth information (stereoscopic video) would be taken note of in this experiment. The video was encoded using all of the four methods with different comfort levels i.e. taking different steps of quantization for the proposed method and different quantization parameters for the other techniques. The compressed videos have then been subjected to the analysis using the following metrics: The key results highlighted that our method achieved overall the most efficient performance among barrel distortion assessment techniques in four areas: PSNR (Peak Signal-to-Noise Ratio), SNR (Signal-to-Noise Ratio), MSE (Mean Squared Error) and CR (Compression Ratio).

3.2 Peak Signal-to-Noise Ratio (PSNR):
The PSNR is one of the most frequently applied parameters in assessing the quality of the video following compression and its relationship to the original. It is the ratio between the highest attainable power of the signal which is the video versus the power of the noise which is the error brought about by compression. It is also known that the larger the values of PSNR, the better. Formula of Peak Signal-to-Noise Ratio is:

$$PSNR = 10 \times log_{10}\left(\frac{MAX_I^2}{MSE}\right) \qquad (4)$$

It is postulated that the maximum possible value of the pixels of an image is set to the maximum attainable value that is 255 for an 8 bit image. MSE is the Mean Squared Error for the original signal and for the decompressed video signal.

The achieved PSNR value in the experiments conducted was 40.5 dB for the proposed method and as can be seen, this is much higher than the conventional methods used thus suggesting high visual quality of the decompressed video after its compression.

(a) Input Video Frame

(b) Compressed Video frame

***Figure 4.** Shows Input and Compressed frames of Input video for Proposed methodology.*

3.3 Signal-to-Noise Ratio (SNR)

The SNR defines the ratio of the power of the useful signal (the useful information - video data in this case) to the power of the noise (compression noise, artifacts, etc). This is another quality measurement for video. The higher the SNR value, the less noise and the better the quality.

Formula:

$$SNR = 10 \times log_{10}\left(\frac{\sum_{i,j,k} X(i,j,k)^2}{\sum_{i,j,k}(X(i,j,k)-Y(i,j,k))^2}\right) \qquad (5)$$

Where:

- *X(i,j,k)* – the original video frame.
- *Y(i,j,k)* – the video frame after it has been decompressed.
- Summation is taken over all pixels of frame.

In the most recent applications, the proposed method has demonstrated an SNR of 37.2 which indicates little perceptual degradation when compared to the previous proportions.

3.4 Mean Squared Error (MSE)

The MSE works as a standardized descriptor of the mean value of the squared deviations between the original video and its reconstruction after it has been decompressed. It provides a measure of the error that was introduced with the compression and encoded data and demonstrates that lower numbers mean less distortion.

Formula:

$$MSE = \frac{1}{MNK}\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{K}(X(i,j,k) - Y(i,j,k))^2 \qquad (6)$$

Where:

- *M, N, K* - size of 3 dimension video (width, height and time/frame length).
- *X(i,j,k)* and *Y(i,j,k)* - the original and decompressed video frames.

The best values obtained were 0.0042 of MSE which is an indication of good compression performance and minimal loss of visual data.

3.5 The Compression Ratio (CR)

The Compression Ratio (CR) refers to the key parameter which quantifies how much the technique has reduced the data. The ratio is determined as the ratio of uncompressed to compressed video in its quantitative representation. The higher the compression ratio, the better.

Formula:

$$CR = \frac{Size\ of\ original\ video}{Size\ of\ compressed\ video} \qquad (7)$$

A Compression Ratio (CR) of 30.5 has been obtained using the proposed method, which is considerably greater than that obtained by conventional compression techniques and demonstrates its efficiency in compressing 3D video while preserving the most of the visual quality The Compressed and the original video frames are shown in figure 4.

Table 1. *Shows Comparison of Proposed and Existing techniques*

Method	PSNR (dB)	SNR (dB)	MSE	CR
Proposed Method (3D-BWT + Scalar Quantization + Huffman Coding)	**40.5**	**37.2**	**0.0042**	**30.5**
Method 1: 3D-DWT + Scalar Quantization + Huffman Coding	39.0	35.5	0.0058	27.8
Method 2: MCTF with Wavelet Transform	38.3	34.9	0.0062	25.3
Method 3: H.264/AVC Compression	35.1	31.7	0.0103	22.1
Method 4: HEVC Compression	36.5	32.4	0.0090	24.6

4. Conclusion

In this paper we address an outstanding gap in 3D video compression techniques by proposing a new 3D video compression method that utilizes loss of information encoding through the enhancement of the 3 D Biorthogonal Wavelet Transform (3D-BWT) with scalar quantization and Huffman coding to permit easy compression and decompression. The key concern of the proposed approach is to provide a method with higher compression ratio and acceptable quality of video that is suitable for real time application, such as 3D video streaming and storage.

After a number of practical tests and analysis, we showed that the proposed method is superior compared to the scope of general compression methods employed such as the traditional method of 3DDWT with scalar quantization followed by Huffman coding, motion compensated temporal filtering MCTF with wavelet transforms and even current state of the art codecs such as H.264/AVC and HEVC.

The experimental results showed graphically that the proposed method outperformed all the key evaluation criteria which included PSNR, SNR, MSE and Compression Ratio (CR). In particular, the proposed method produced the greatest PSNR of 40.5 dB, & SNR of 37.2 dB, the least error value of mean square error of 0.0042 and the best compression ratio of 30.5.

This combination proves to be very effective when seeking to maintain video quality whilst significantly lowering the data cloud. It is because the use of 3D-BWT enabled the better representation of space and time relationships within the three-dimensional video frames that such compression performance has been enhanced.

In addition to this, the use of scalar quantization and then Huffman coding in this order enhanced efficient entropy coding which enabled very high compression ratio and low distortion.

This approach makes a strong argument for the use of wavelet-based techniques in current 3D video compression applications where compression ratio and video quality are both important factors. Similar works done in literature [10-13].

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.

- **Author contributions:** The authors declare that they have equal right on this paper.

- **Funding information:** The authors declare that there is no funding to be acknowledged.

- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References:

[1] A. R. Smith et al. (2015). Efficient multi-resolution compression of 3D video using wavelet transforms. *Journal of Multimedia Compression*. 25(4):10-20. DOI:10.1109/DCV.2001.929950.

[2] M. Kurokawa et al. (2015). Application of scalar quantization in 3D video compression. *IEEE Transactions on Circuits and Systems for Video Technology*. 23(2):345-352.

[3] J. Kim et al. (2016). Efficient entropy coding for video compression using Huffman coding. *Journal of Digital Video Processing*.13:98-105.

[4] H. Yang and W. Yang. (2018). Performance evaluation of wavelet-based 3D video compression techniques. *International Journal of Imaging Systems and Technology*. 28:163-171.

[5] P. Maiti et al. (2019). 3D video compression using multi-dimensional transforms. *Proceedings of the IEEE International Conference on Image Processing*.

[6] Shih, T. S., & Chen, C. L. (2011). A hybrid compression method for 3D video. *IEEE Transactions on Circuits and Systems for Video Technology*. 21(10):1464–1475.

[7] Zhang, X., & Zhao, J. (2009). Compression of 3D video using a motion-compensated 3D wavelet transform. *Signal Processing: Image Communication*. 24(7):512–522.

[8] Jiang, Y., & Yu, Y. (2012). Efficient 3D video compression based on wavelet transform and motion compensation. *International Journal of Image Processing*. 6(4):270-280.

[9] Wiegand, T., & Sullivan, G. J. (2012). Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*. 22(12):1669-1684. doi: 10.1109/TCSVT.2012.2221191.

[10] GUNDA, P., & Thirupathi Rao KOMATI. (2024). Integrating Self-Attention Mechanisms For Contextually Relevant Information In Product Management. *International Journal of Computational and Experimental Science and Engineering,* 10(4);1361-1371. https://doi.org/10.22399/ijcesen.651

[11] ARSLAN, M. T., & YILDIRIM, E. (2024). Classification of Intensive-less Intensive and Related-Unrelated TasksTasks:. *International Journal of Computational and Experimental Science and Engineering,* 10(2);221-227. https://doi.org/10.22399/ijcesen.328

[12] PATHAPATI, S., N. J. NALINI, & Mahesh GADIRAJU. (2024). Comparative Evaluation of EEG signals for Mild Cognitive Impairment using Scalograms and Spectrograms with Deep Learning Models. *International Journal of Computational and Experimental Science and Engineering*, 10(4);859-866. https://doi.org/10.22399/ijcesen.534

[13] N, S., S. Prabu, V, T. K., D, C., K, B., & B. Buvaneswari. (2024). Computer Aided Based Performance Analysis of Glioblastoma Tumor Detection Methods using UNET-CNN. *International Journal of Computational and Experimental Science and Engineering,* 10(4);753-762. https://doi.org/10.22399/ijcesen.515