

An Efficient DWT and Tucker Decomposition with H.264 Video Compression for Multimedia Applications

N. Sardar Basha, A. Rajesh

Abstract--- In last thirty years, there has been so much of intensive research has been carried out on video compression techniques and now it has become mature and used in a large number of applications. In this paper, we are trying to present video compression using H.264 compression with Tucker decomposition. The largest Kn sub-tensors and their eigenvectors with run length encoding to compress the frames in the video was obtained by implementing tucker decomposition of tensor. DWT is used to separate each frames into sub-images and TD on DWT coefficient to compact the energy of sub-images. The obtained experimental results supported that our proposed method yields higher compression ratio with good PSNR.

Index Terms--- DWT, Tensors, Tucker Decomposition, H.264 and Multimedia Applications.

1. INTRODUCTION

Huge demand for the videos has been created due to the arrival of multimedia computing, due to this not only the manipulation, but storage of these videos in their raw form which is very expensive and it significantly increases the transmission time and makes storage costly. Most of the surveillance applications need to share their network with other data intensive applications infrequently and uncompressed videos has been shared over digital networks which needs to have very high bandwidth. In order to overcome this problem, video compression techniques were derived so that one can reduce the number of bits actually required to represent a digital video data along with the maintenance of video quality as well as with acceptable fidelity. The comparison and analysing of the video coding standards such as MPEG and the H.264 family is the main objective here. An important and difficult problem is real time video compression which attracts to the researcher to deal with it. However, the evolution is going on and day by day we encounter with new technologies and video codec standard as H.264/AVC. This research paper explores the possibility of modified real time video compression algorithm and quality enhancement by using tensor decomposition with H.264 compression. Based on few parameters, the approximation of tensor data sets is possible using the tensor decomposition which allow us to have less parameters when compared to the number of total entries of the tensor. The capturing of the essential structures in multidimensional data sets was made possible due to reduction in the degree of freedom and as tensor decomposition is known to take two forms called Tucker decomposition and Canonical Polyadic (CP) decomposition

[1]. As we know Tensor methods were proved to be a competent technique for data analysis [4] and it has been applied successfully in many applications like computer vision, graph analysis, data mining, signal processing etc. We have selected the Tucker decomposition which is a well-known algorithm to achieve TD. Tucker has introduced this classical Tucker decomposition [5].

2. LITERATURE SURVEY

Mahalakshmi *et. al* [3] has summarized not only the prevailing new video coding techniques but also next generation video compression existing algorithms in the paper. Here in this work, we tried to organize the various basic techniques which are used for video compression along with various generations, as well as the implementation strategies used for video compression techniques with an idea on video compression latest trends giving an insight on emerging technologies in this field and different issues related to it.

Lei Chen *et.al* [6] used precompression algorithm for compression of hyperspectral image. Depending on the high spectral correlation, the general relation between bands has to find in this algorithm. In this algorithm they first find the relationship existing between the bands with respect to high spectral correlation. Basing on the relationship obtained, precompression algorithm has been designed in such a way that it can segment the bands into group, in order to achieve the removal of both spectral and spatial correlation of each group of images. After this 3D-DWT is applied for further compression. At low bit rate, this method produces significantly improved performance. Based on the three dimensional discrete cosine transform as well as on tucker decomposition, Karami *et. al.*, [7] has proposed a hyper spectral image compression, which applies Tucker decomposition on DCT coefficient of image. It reduces the size of every dimensions of tensor and thus reduces the quantization depth and dimension.

G. A. Licciardi *et.al* [8] has proposed a method for the compression of hyperspectral data to preserve original information. Proposed method is based on non-linear generalization of standard Principal Component Analysis for each image the compression rate depends only on nodes of bottleneck layer. Hence it by operating on number of nodes, one can have the possibility of increasing the spectral information. By this method, achieved compression rate is around 95%.

Revised Version Manuscript Received on August 14, 2019.

N. Sardar Basha, Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya University, Kanchipuram, Tamil Nadu, India. (e-mail: sardarbashame@gmail.com)

A. Rajesh, C Abdul Hakeem College of Engineering & Technology, Vellore, Tamil Nadu, India. (e-mail: amrajesh73@gmail.com)



By comparing both H.264/AVC and LWC, the Hyun Kim et al [9] has proposed a low-power hybrid recoding system and LWC has been used for the temporary storage and for permanent storage, the H.264/AVC has been made use of. LWC is known to consume very less power in comparison with H.264/AVC, thus it is used for temporary storage and using the down sampling operation in the video recording system, one can achieve a further power reduction. The reduction in the computational complexity and power consumption for the video compression in comparison with the rate of scaling down can be achieved by scaling down the video frame by down sampling operation.

An improved algorithm was proposed by Hao Chen et al [10] by using traditional temporal frame interpolation method with the use of bidirectional motion vectors in order to classify the motional relationship to determine the weighting ration existing between adjacent framed and this can be used to construct the missing frame with the help of interpolation. For a given input video sequence, the temporal resolution can be enhanced by diving into classes the movement in adjacent frames suing the improved temporal frame interpolation algorithm (I-TFI).For moving images in the various video processing applications, a macro block classification methods has been proposed by Saravanan[11]. in the compressed video, the Motion Vector field has been analysed to propose a methods where we have classified into classes the Macro blocks of each video frame and used this information of classes to provide description of the frame content. In our algorithm which has been proposed, it can observe both H.264 results and macro block classification can work more effectively on video processing applications. For scalable video compression coding, ZHU Xiao-dong *et. al* [12] has proposed both motion estimation as well as compensation technology. About H.264/AVC motion estimation technology, we can see a detailed discussion which includes module decision making,

principles of motion estimation, rate distortion optimization, macro block division, standards of evaluation of the quality of coding as well as best matching algorithm.

3. EXISTING METHOD

In this paper, we have used very potent video compression system and originally with the aid of Listless SPECK (LSK) algorithm, the resulting coefficients are quantized after applying wavelet decomposition to the I-frame. Secondly, the motion estimation is done by spatio-temporal correlation technique (ARS-ST) with an adaptive rood search to find from the I-frame, the displaced blocks of P-frame [2].

4. PROPOSED METHOD

The H.264 is considered as a broad open standard as it comes with many suggested tools which can give us resulting compression rate with wide difference from one implementation to another. For different markets as well as industries, profiles (various sets of capabilities) are made use. A very high data compression resultls can be achieved using the H.264 and even at lower bit rates, it provides better image quality in comparison to the previous standards. One can achieve a image reduction of approximately 50% by using H.264 compression standard in comparision with MPEG-4, thus you can achieve a reduction in the use of network bandwidth and storage and mostly advantageous for high motion video. In order to improve the efficiency, we have combined tucker decomposition with H.264 standard. In this input video is converted into frames. The frames are pre-processed and the transformed using 2DWT.The transformed image is decomposed using tucker decomposition. Then the decomposed image is compressed using H.264 standard. Then the compressed frames are converted into the video.

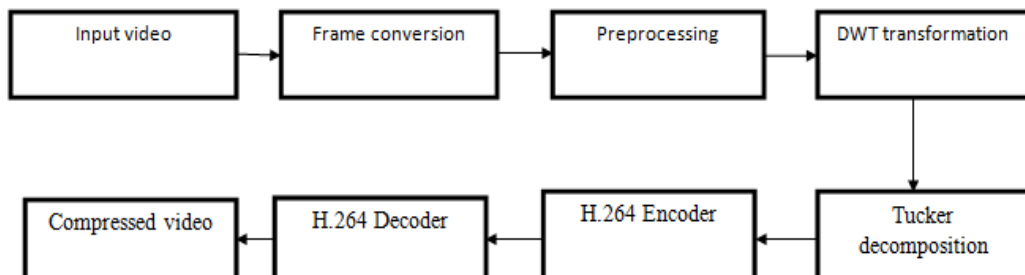


Fig. 1: Block diagram of proposed method

4.1 Tensor Construction

In general, a tensor is denoted by $A \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ an N-order tensor, and denote by $a_{i_1 \dots i_N}$ its elements, where $1 \leq i_n \leq I_n$. Let $B \in \mathbb{R}^{J_n \times I_n}$ denote a matrix. The n-mode product of a tensor A and a matrix B can be defined as

$$D = A \times_n B \quad (1)$$

where the symbol \times_n denotes the tensor-times matrix operation, and the mode-n product output $D \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N}$ is a tensor of order N. The element wise representation of Eq. (1) can be written as:

$$D(i_1 \dots i_{n-1}, j_n, j_{n+1}, \dots, i_N) = \sum_{k=1}^{I_n} A(i_1 \dots, i_{n-1}, k, i_{n+1}, \dots, i_N) B(j_n, k) \quad (2)$$

Using matrix multiplication, we can calculate the mode-n product:

$$D(n) = BA(n) \quad (3)$$

where $A(n) = \text{unfold}_n(A) \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$ and $D(n) = \text{unfold}_n(D) \in \mathbb{R}^{J_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$, $1 \leq n \leq N$, are called mode-n matrices. The reordering of the elements of an n-way data array into a matrix can be done using a process called operator $\text{unfold}_n(\cdot)$. Conversely, we can achieve the transformation of unfolding matrices to the tensor along the nth mode using the $\text{fold}_n(\cdot)$ operation. For convenience, we define

$$\begin{aligned}
 & A X_{-n} \{B^{(j)}\}^N j = 1 \\
 & = A x_1 B^{(1)} x_2 \dots x_{n-1} B^{(n-1)} x_{n+1} B^{n+1} \dots x_N B^{(N)} \\
 & = A_{(n)} (B^{(N)} \otimes \dots \otimes B^{(n+1)} \otimes B^{(n-1)} \otimes \dots \otimes B^{(1)}) \quad (4)
 \end{aligned}$$

where \otimes denotes the Kronecker product. $A(n)$ is defined as:

$$A(n) = \Psi(n) \Sigma(n) V(n)^T \quad (5)$$

and the leading R_n -dimensional left singular subspace of $A(n)$ is defined as

$$\Psi(n)_{rn} = \Psi(n) (:, 1 : R_n).$$

4.1.1 Tucker Decomposition

Tucker decomposition aims to decompose an N -order tensor $X \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ as an affiliation of N orthogonal bases $\{U(n) \in \mathbb{R}^{I_n \times R_n} \mid N=1\}$ and the associated core tensor $Y \in \mathbb{R}^{R_1 \times R_2 \times \dots \times R_N}$, where $R_n \leq I_n$. It can be formulated as:

$$Y = X x_1 U^{(1)T} x_2 U^{(2)T} x_3 \dots x_N U^{(N)T} \quad (6)$$

To find the optimal orthogonal matrices $\{U(n) \mid N=1\}$ and the core tensor Y , we could minimize the error between the original data tensor X and its approximation \hat{X} leading to the following optimization problem:

$$\underset{U^{(1)} U^{(2)} \dots U^{(N)}}{\operatorname{argmin}} \|X - x_1 U^{(1)} x_2 U^{(2)} x_3 \dots x_N U^{(N)}\|_{F^2} \quad (7)$$

Eq. (6) can be recast as an optimization problem

$$\underset{U^{(1)} U^{(2)} \dots U^{(N)}}{\operatorname{argmin}} \|X x_1 U^{(1)T} x_2 U^{(2)T} x_3 \dots x_N U^{(N)T}\|_{F^2} \quad (8)$$

to maximize $\|Y\|_{F^2}$

we can easily obtain the corresponding core tensor Y by Eq. (5).

4.2 H.264 Compression

H.264 known to make use of transform based residual coding as well as block based motion compensation to achieve the output and also additionally you can with adjustable quantization step size due to scalar quantization to get output bit rate control along with run length coding of quantized transform coefficient and zigzag scanning and in detail if we observe we can see significant differences. For motion compression, the H.264 is known to use efficient and more flexible model and uses seven different block sizes along with multiple reference pictures to achieve the motion compensation. 1/8-pixel accuracy is used to specify the motion vectors and within the motion compensation loop, the improved de-blocking filter is also used to achieve improved prediction with reduced visual artifacts.

4.2.1 H.264 Encoder

Through this work, we came to know that H.264 provides us a higher coding efficiency and at both low bit rate and low bandwidth, we can get high quality video images. This H.264 allows us to work for real time communication in low latency mode like video conferencing.

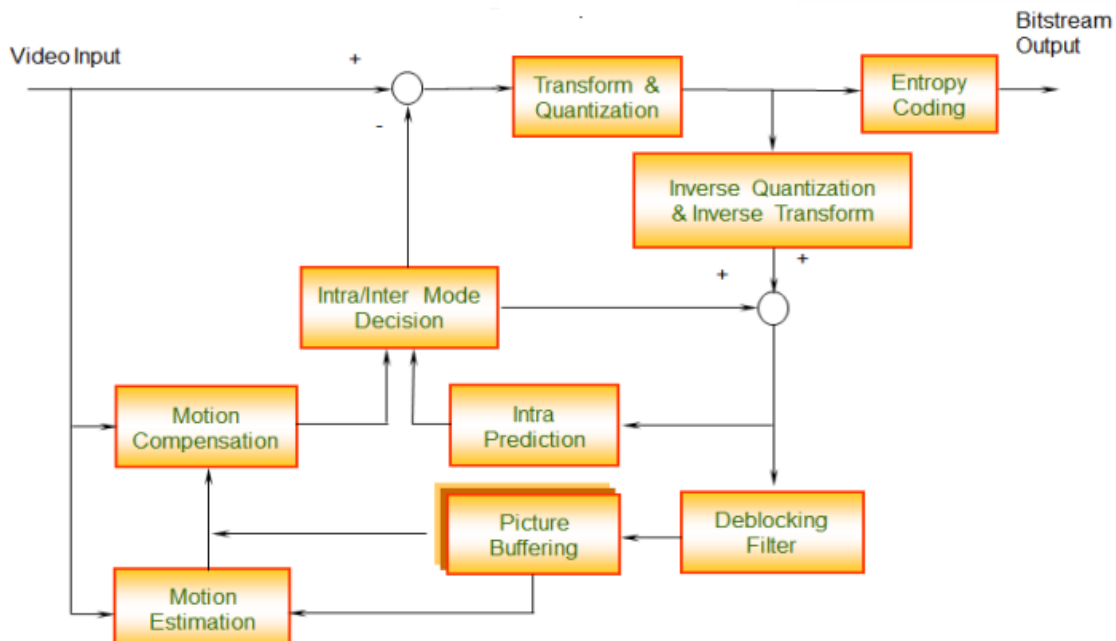


Fig. 2: Architecture of H.264 Encoder

4.2.2 Intra Prediction and Coding

In intra coding efficiency, a significant improvement can be achieved by prediction of the current block in adjacent block from the neighbouring pixels in a given defined set of directions using the intra prediction attempts and intra coding. Instead of taking the actual block into consideration, the difference between the best resulting prediction and the block is coded.

4.2.3 Inter Prediction and Coding

Quarter-pel motion compensation, motion compensation, adaptive loop de-blocking and multiple reference frames can

be achieved for various block sizes by leveraging the Inter-frame coding in H.264 with most of the key features of earlier standards which is known to add both functionality and flexibility.

4.2.4 Motion compensation

We can perform motion compensation using a number of block sizes of different sizes and transmission of individual motion vectors can be achieved as small as 4x4, even upto



32 motion vectors for a single macroblock in bi-directional prediction case. This will support 16x8, 8x16, 8x8, 8x4, and 4x8 block sizes and we have the advantage to improve the ability in order to handle fine motion detail by having option for smaller motion compensation, which gives a better subjective quality which includes the absence of large blocking artifacts.

4.2.5 Motion Estimation

With the help of quarter-pel and half-pel motion vector resolution, we can achieve improved motion compensations.

4.2.6 Adaptive Loop Deblocking Filter

Within the prediction loop, the H.264 operates on the horizontal as well as vertical block edges using the adaptive deblocking filter in order to remove artifacts which are caused by the block prediction errors. On 4x4 block boundaries, the filtering is generally based and with the help

of 3-tap filter the two pixels on either side of the boundary may be updates. Loop de-blocking filter rule application is intricate and even it is quite complex.

4.2.7 Transform and Quantization

Without any widened dead-zone, the scalar quantization can be used for quantization of transform coefficients. On a macro block basis, one can select thirty two different quantization step sizes can be chosen which are similar to prior standards, at 12.5% the compounding rate is seems to be increased with increasing step sizes. The finer quantization step sizes can be using to improve the fidelity of chrominance components in comparison with luminance coefficients, in particular for a condition where the luminance coefficients are coarsely quantized.

4.2.8 H.264 decoder

H.264 decoder is the inverse process of encoder.

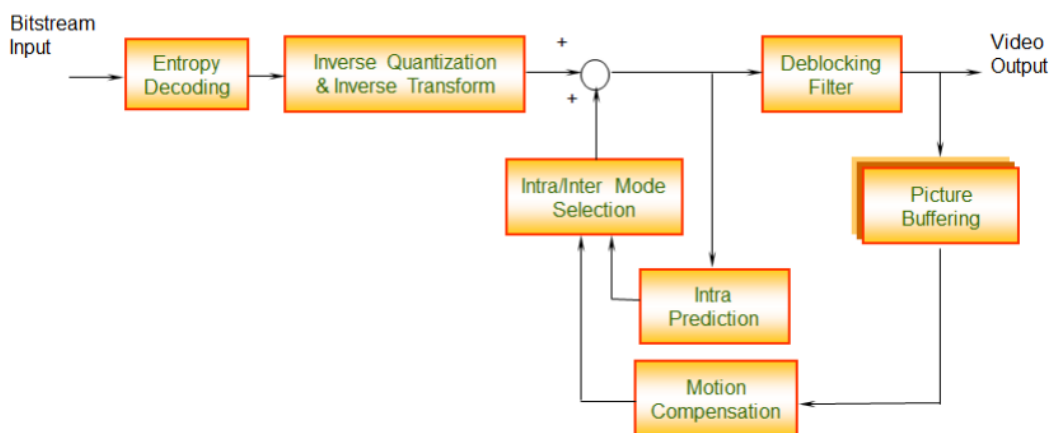


Fig. 3: Architecture of H.264 Decoder

4.3 DWT Algorithm

Discrete wavelet transform has been successfully used for image processing application such as edge detection, image compression. Wavelet transform produces coefficients of image. These coefficients are then compressed more easily, because the information is focused on few coefficients. Wavelet transform has an advantage to separate the fine details in the signals. At each level of compression, high pass filter produces the detailed information and low pass filter produces the approximate information. We apply 2DWT on each frame in the video. If each frame has M row size and N column size then after applying 2DWT, we obtained four sub-images (A, H, V, D) each band having size M/2 rows and N/2 column. The approximation(A) sub-band having highest energy amongst all other three sub-bands. The 2DWT of function f(x, y) for size MxN is given by

$$W_{\phi}(m, n) = \frac{1}{\sqrt{MN}} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} f(x, y) \phi_{m,n}(x, y) \quad (9)$$

$$W_{\psi}^k(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} f(x, y) \Psi^k_{j,m,n}(x, y) \quad (10)$$

$$K = \{H, D, V\}$$

ϕ is a scaling function. $W_{\phi}(m,n)$ is the approximation of function $f(x,y)$. For wavelet filtering, the biorthogonal3.7

wavelets is used for better compression. Inverse 2DWT is given by

$$f(x, y) = \frac{1}{\sqrt{MN}} \sum_{i2} \sum_{i1} W_{\phi}(m, n) \phi_{i1,i2}(x, y) + \frac{1}{\sqrt{MN}} \sum_{k=\{H,D,V\}} \sum_{j=0}^{J-1} \sum_n \sum_m W_{\psi}^k(m, n) \psi^k_{j,m,n}(x, y) \quad (11)$$

5. EXPERIMENTAL RESULTS

The tests are performed on standard video sequence which are Akiyo and Hall video.

The performance is compared in terms of peak signal to noise ratio PSNR and compression ratio(CR) using following relationship

$$PSNR = 10 \log_{10} \left(\frac{255}{MSE} \right)^2 \quad (12)$$

$$MSE = \frac{1}{xy} \sum_{x=1}^x \sum_{y=1}^y [i(x, y) - e(x, y)]^2 \quad (13)$$

Where $i(x,y)$ - intensity of input pixel. $e(x,y)$ - intensity of output pixel. CR represents the ratio between the number of bits required to store the pre-compression image (I) and the number of bits which are required to store the compression image (O).

$$CR = I/O$$





Fig. 4: Comparison of input and encoded frame for news reader Hall Monitor

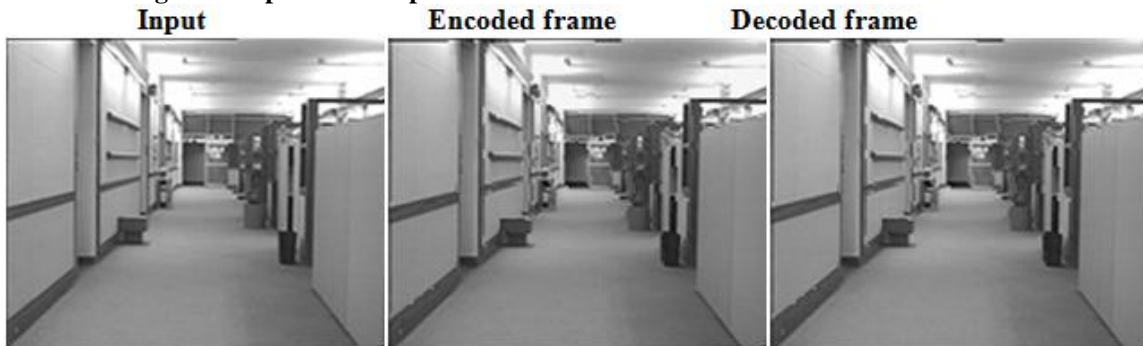


Fig. 5: Comparison of input, encoded and decoded frame for hall monitor

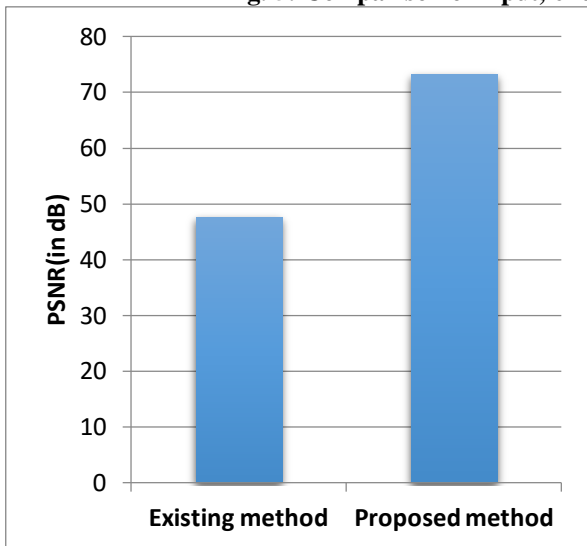


Fig. 6: PSNR comparison for Akiyo video

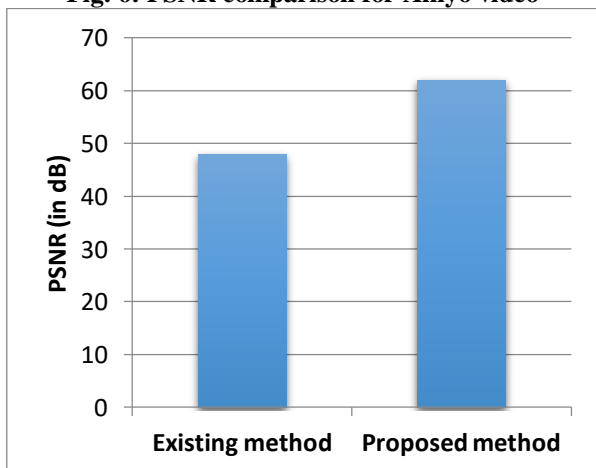


Fig. 7: PSNR comparison for Hall video

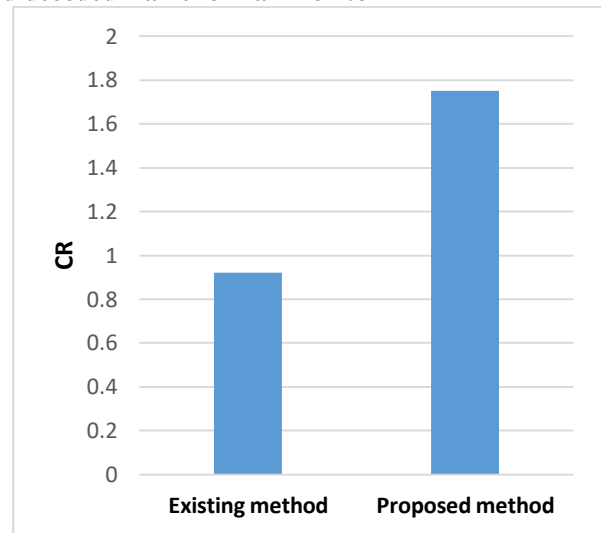


Fig. 8: CR comparison for Akiyo video

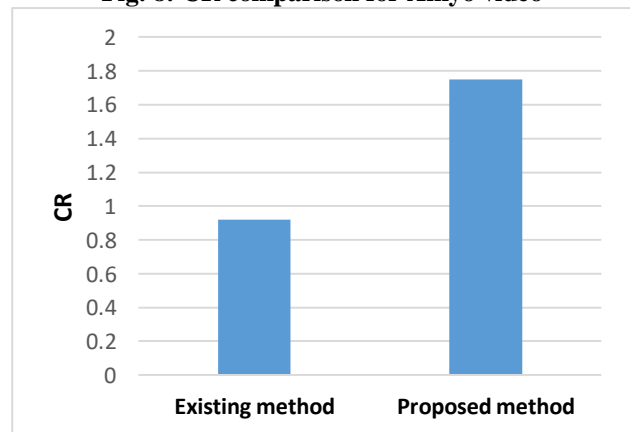


Fig. 9: CR comparison for Hall video

AN EFFICIENT DWT AND TUCKER DECOMPOSITION WITH H.264 VIDEO COMPRESSION FOR MULTIMEDIA APPLICATIONS

We can see the comparison of PSNR for the proposed method with the existing technique as shown in Fig6 and Fig 7 for akiyo and hall video. The proposed method as shown in this graph has high PSNR ratio in comparison with the existing technique which is represented in Table 1. From Fig8 and Fig 9 we found that the proposed method has higher CR ratio in comparison with video compression technique which already exist as shown in Table. 2.

Table 1: PSNR Comparison

Method	Akiyo video	Hall video
Existing method[2]	47.59	48.02
Proposed method	73.10	62.02

Table 2: CR Comparison

Method	Akiyo video	Hall video
Existing method[2]	1.37	1.74
Proposed method	0.92	0.93

6. CONCLUSION

Though this work, we tried to put forward a novel method of image compression, by encoding images into combined H.264 with tucker decomposition. Not only we tried to keep the spatial intrinsic structures of given high dimensional data but also made use of the traditional method's handling vectored data. In comparison to SPECK-LSK, we obtained high quality reconstructed images using our experimental results. The efficiency of our proposed method was shown using the PSNR and CR. Thus, as we can see in this paper, in terms of PSNR, the proposed algorithm has better performance and we can propose this as a very suitable algorithm with respect to high quality video sequence compression.

REFERENCES

1. T. G Kolda and B.W Bader, "Tensor decompositions and applications," SIAM review, vol. 51, no. 3, pp. 455–500, 2009.
2. S. A. Padmanabhan and S. Chandramathi," A Wavelet based video compression using Adaptive Rood Search with Spatio-Temporal Correlation and LSK", International Journal of Computers and Applications, Vol. 35, No. 1, 2013.
3. R.Mahalakshmi, S.K.Mahendran," Analysis of Video Compression Technique", International Journal of Advanced Research in Computer Science & Technology (IJARCST 2017).
4. X. Liu, S. Ji, W. Glänzel and B. D. Moor, "Multiview partitioning via tensor methods,"IEEE Transactions on Knowledge and Data Engineering, vol.25, no.5, pp.1056–1069, 2013.
5. X. Cao, X. Wei, Y. Han and D. Lin," Robust face clustering via tensor decomposition", IEEE Transactions on Cybernetics, vol.45, no.11, pp.2546–2557, 2015.
6. Lei Chen, Chao Bei, YujiaZhai, Jing Wu "Efficient Wavelet-based Compression for Hyperspectral Images Using Precompression", 2011 International Conference on Multimedia and Signal Processing
7. A. Karami, M. Yazdi and A.Zolghadre "Hyperspectral Image Compression Based on Tucker Decomposition and Discrete Cosine Transform", 2010,

8. G. A. Licciardi, J. Chanussot, A. Piscin "Spectral Compression of Hyperspectral Images by means of Nonlinear Principal Component Analysis Decorrelation",2014 International Conference On Image Processing(ICIP).
9. H. Kim, C. ERhee and H. J Lee," A Low-Power Hybrid Video Recording System with H.264/AVC and Light-Weight Compression", 2014.
10. H. Chen, Y. Zhang, Member, IEEE, Yu Tao, Bin Zou, Member, IEEE, Wenyan Tang," An Improved Temporal Frame Interpolation Algorithm for H.264 Video Compression", 2011
11. S.Saravanan," Enhancing the performance of Video compression on display devices using Macro block classification and H.264 AVC", 978-1-4799, 2013.
12. Y.U Song-nian," Research on Scalable Coding Technology Application in Video Compression Based on H.264 Standard", 978-1-61284, 2011
13. Macías, T. M. D., Meza, A. K. T., Garcia, B. B. B., &Bozada, M. A. T. (2018). Characterization of Physical and Motor Disability at the Technical University of Manabí. International Research Journal of Management, IT and Social Sciences, 5(2), 1-8.
14. Briones, MFB, Intriago, HAM, & Ruiz, MSL (2018). Motor Disability and Resilience in University Students. International Research Journal of Management, IT and Social Sciences, 5 (2), 47-54.