



Developing novel video coding model using modified dual-tree wavelet-based multi-resolution technique

S. S. Nithin¹ · L. K. Padma Suresh² · S. H. Krishnaveni³ · P. Muthukumar⁴

Received: 23 August 2021 / Accepted: 16 October 2021 / Published online: 17 January 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

All data are kept on digital platforms in today's digital world, demanding a lot of storage space for images and video, as well as a lot of bandwidth for transmission. Data that have been compressed is highly beneficial for storing more data at the time. The objectives of this work are to examine various compression techniques developed by various researchers and to develop a new video compression method based on multi-resolution techniques. Initially, the video is compressed using wavelet transform and different encoding techniques. As a result, all comparisons will use Empirical Wavelet Transform (EWT). The encoding techniques used here are H.264, Huffman, LZW, SPIHT, and their combinations. With the help of various performance matrices combination of H.264 and modified SPIHT gives better performance. The SPIHT is modified to overcome the limitations of normal SPIHT. The encoding block is constant throughout the next phase, whereas the transform part is variable. The image is transformed using the Biorthogonal Wavelet, Coiflet Wavelet, Demeyer Wavelet, Mexican Hat Wavelet, Dual-Tree Wavelet, Dual-Tree 3d Wavelet, Curvelet, and Modified Dual-Tree Wavelet. This Modified Dual-Tree Wavelet performs better than DTCWT and overcomes its limitations. When comparing the outcomes of several video coding methods, it was discovered that Modified DTCWT with a combination of H.264 and SPIHT provides the best results.

Keywords EWT · Multiresolution · H.264 · SPIHT, SPIHT, DTCWT, video compression

1 Introduction

Nowadays, all data are stored on digital platforms, which necessitates a large amount of storage space for storing photographs and videos, as well as a large amount of bandwidth for transmission. These two disputes resulted in the development of new compression methods. It has been a challenging task to store and transfer large amounts of raw data. Data

compression [1–5] refers to the process of reducing the size of data stored. Video compression is important in the media industry and other related sectors such as video broadcasting, video conferencing, and video streaming.

The capacity to illustrate information acceptably is referred to as compression [1, 6]. The duplication and irrelevancy in the data are exploited and eliminated to get this compact representation. Samuel Morse created a model for data compression using Morse code in the mid-nineteenth century. Dots and dashes are used to encrypt the symbols sent via telegraph. Some letters appear more frequently than others, according to Morse. Shorter sequences are assigned to letters that occur more frequently, and longer sequences are assigned to letters that occur less frequently, to reduce the average time required to transmit a message. Huffman coding and Shannon Fano coding both use this method of employing transient code words for more often occurring features.

Video is one of the most demanding applications due to the large quantity of data it must handle. Due to this reason, compression becomes an essential part of such applications [7, 8]. In general, the goal of most compression systems is to reduce the volume of data by controlling redundancies

Communicated by Y. Zhang.

✉ S. S. Nithin
nithinssme@gmail.com

¹ Department of Electronics Engineering, St Joseph college of Engineering and Technology, Pala, Kerala, India

² Department of Electrical and Electronics Engineering, Baseliros Mathews II College of Engineering, Kollam, Kerala, India

³ Department of Computer Science and Engineering, Baseliros Mathew II College of Engineering, Kollam, Kerala, India

⁴ Department of Electrical and Electronics Engineering, Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

and irrelevancies in the data and eliminating them without causing much distortion in quality. Lossy compression and lossless compression are the two most used techniques of compression. Due to quantization, certain data are lost in lossy compression techniques, resulting in excellent compression ratios but poor reconstruction quality. This decline in quality should be within a certain margin of error. The acceptable level of error depends on computational complexity, memory requirements, data input and output requirements, compression, and decompression delay constraints. As the compression ratio is increased, the reconstructed image becomes distorted and the quality degrades. Lossless compression schemes [9] exploit only the redundancies in the data and do not discard any information. This results in lower compression ratios [10]. The decompressed data are the replica of the original data. Lossless compression schemes are used for medical purposes, wherein the quality of the images is of utmost importance [11–26].

The process of making video coding applications has taken part a crucial position in the progress of digital video conveyance applications over late years by Ghanbari in 1989 [3]. Standardization allows interoperability across disparate creators and is a crucial specification for broadcasting services. The two global standardization organizations are Video Coding Experts Group (VCEG) and Motion Picture Experts Group (MPEG). A typical compression (or coding) system is made up of an encoder and a decoder. The encoder converts the video sequence into a compact representation that is transmitted or stored, while the decoder performs the inverse operation.

The existing technique [27–31] utilized compression rules and coding information but fails to make a dynamic change in the network structure. This paper proposed a novel framework known as Video Coding employing Modified Dual-Tree Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCMDTWHMSE) to enhance the video quality at both the encoder and decoder sides. The major contributions of this paper are delineated as follows:

- The proposed VCMDTWHMSE methodology uses the modified DTCWT transform for decomposition and a combination of H.264 and modified SPIHT for encoding to improve the video quality at both the encoder and decoder sides.
- The Empirical Wavelet Transform (EWT) is performed on the image to enhance the input signal reconstruction using Biorthogonal Wavelet, Coiflet Wavelet, Demeyer Wavelet, Mexican Hat Wavelet, Dual-Tree Wavelet, Dual-Tree 3d Wavelet, Curvelet, and Modified Dual-Tree Wavelet.
- For the encoding process, the combination of H.264 and modified SPIHT is used to improve the processing speed and it is fixed for all comparisons.

- For performing transform function, Biorthogonal Wavelet, Coiflet Wavelet, Demeyer Wavelet, Mexican Hat Wavelet, DTCWT, 3D DTCWT, Curvelet, and Modified DTCWT are utilized. These functions minimize the shift variance problem and overcome the low directional sensitivity.
- Several performance matrices, such as CR, PSNR, MSE, and SSIM, are used in the Modified Dual-Tree Wavelet to improve performance. As a consequence, modified DTCWT is chosen as the best video compression transform.

The rest of this paper is structured accordingly. Section 2 presents the existing literary works conducted in this field along with their drawbacks. Section 3 presents the proposed video coding methodology using different transforms and encoders. The development and implementation of efficient video coding using multi-resolution techniques are discussed in Sect. 4. The extensive experiments conducted to evaluate the efficiency of the proposed methodology are presented in Sects. 5 and 6 concludes the paper.

2 Review of related works

Message transmission among several groups is a basic asset of any community. Speech, audio, and video are used in our daily lives to convey information. In later years, media transmission achieve great significance in various fields including mobile, telemedicine, teleconferencing, and military applications. Here, size of the data is also important for this transmission purpose. As a result, there is a lot of research going on in this area and a literature review emerged as a subdiscipline of this area. It focuses on a variety of compression methods and procedures for limiting the size of the input video.

The various standards used for video compression are presented in the paper [2]. At this, H.264/AVC shows better coding operation development related to previous techniques. To enhance the coding operation better than H.264/AVC, the new technique H.265/HEVC is introduced [32], which provides 64×64 pixels value through the joint collaborating team. The video compression standards are based on the motion compensation standard, which reduces video information linked to motion estimate from one frame to another [3]. The capabilities of DCT-related method for compressing video are explained and this analysis paper presents the development in video compression method with better standard and content of the video. Block matching methods are utilized for motion estimation in video compression [4].

A method for compressing video series using multiwavelet with SPIHT and MW block tree coding is implemented in literature [5]. Wavelet Block Tree Coding (WBTC) standard enhances the compression capabilities of SPIHT at lesser

rates by satisfactorily encoding both inter and intra-scale correlation conditions by utilizing block trees. The embedded coding is enhanced by zero-trees wavelet coding (EZW) [33], which acts as an encoder. It is, however, used in both the wavelet transform and dimensional signals. The improved compression technique employs bit plane slicing, the Huffman algorithm, and the Lempel–Ziv–Welch (LZW) dictionary [6]. The following are the issues with video compression that were discovered during the literature review [34–38].

They use only a few video compression features that do not provide quality compression. The technologies used to compress and transmit each frame result in increased bandwidth requirements while maintaining video quality [39, 40]. The preceding techniques take longer, and the most frequently used wavelet transform algorithms in video compression (Complex Wavelet Transform (CWT), DCT, and DWT) all have significant drawbacks [41–45]. The limitation of DCT is that its vulnerability to block noise and limited scalability, in contrast, DWT reduces the image with less superiority. Due to the presence of a small amount of noise, such constraints in the Wavelet Transform (WT) approach produce an indistinct image and maintain a low PSNR quantity. Despite its widespread use in video compression, the CWT is susceptible to shifts, lacks phase information, and has poor directionality. When an error is introduced, the encoder EWZ takes longer and performs poorly. The classic SPIHT encoding technique [30] requires additional cache space to collect its three lists (List of Significant Pixel (LSP), List of Insignificant Pixel (LIP), and List of Insignificant Set (LIS)). The SPIHT algorithm is a blind encoding approach due to its inefficient coefficient partitioning mechanism, which generates an additional number of comparison operations that constitute scanning redundancy. Furthermore, increasing the number of bits increases the number of unnecessary bits in the output bitstream.

3 Proposed approach

Based on the requirements and accessibility in the existing study, we use DTCWT as the transform and H.264 and SPIHT as the encoders. By overcoming the limitations in this existing technique, modify the Dual-Tree Complex Wavelet Transform (DTCWT) and SPIHT to enhance the performance of the video compression. Therefore, to increase the efficiency of the system, several performance metrics such as PSNR, CR, SSIM are MSE measured and evaluated. The experiments are directed to calculate the efficiency and the process is evaluated by three phases.

EWT is used as the transform in this phase and it is constant for all techniques included in this phase. The encoder component is then varied using different encoders and their

combinations. In this case, a combination of H.264 and SPIHT produces superior results, and SPIHT is modified to provide even better results. In this phase, the encoder is a combination of H.264 and SPIHT, and the transforms are varied. Here, different types of wavelets and Curvelet transform are taken for better video compression. From this, DTCWT shows better results and then modifies them for increasing performance.

3.1 Video coding employ different encoders

This Video Coding employs Wavelet Transform along with Different Encoding techniques (VCWTDE), transform part is fixed but the encoding part is varying. The methods which are used for analyzing are Video Coding employing Wavelet along with H.264 Encoding technique (VCWHE), Video Coding employing Wavelet along with Spiht Encoding technique (VCWSE), Video Coding employing Wavelet and coalescence of H.264 along with Spiht Encoding technique (VCWHSE), Video Coding employing Wavelet and coalescence of Huffman along with Spiht Encoding technique (VCWHUSE), Video Coding employing Wavelet along with LZW Encoding technique (VCWLE), Video Coding employing Wavelet along with Modified Spiht Encoding technique (VCWMSE), Video Coding employing Wavelet and coalescence of H.264 along with Modified Spiht Encoding technique (VCWHMSE) [30–33, 46]. The VCWHMSE approach outperforms the others when compared with PSNR, CR, SSIM, and MSE. The encoding mechanism is altered by combining H.264 and Modified Spiht Encoder to form HMSE. The block diagram of VCWTDE is shown in Fig. 1.

Empirical Mode Decomposition (EMD) is the data examination medium suggested by Huang in 1998 [47]. The EMD decomposes data into specific modes while operating in a relative frequency position. It also functions in a time-related space. It is sensitive, and it is defined by a specific base derived from the data. According to the theory, the data could include multiple identical simple fluctuating modes with significantly different frequencies at any given time, each of which would overlap the other during decomposition.

H.264 is a manufacturing-related video compression technique and the process of converting digital video into an arrangement that takes up less space when saved or transmitted. Digital television, DVD-Video, portable television, broadcasting, and internet video casting all use video compression as a standard. The ability to combine outcomes from multiple developers is made possible by determining video compression yields [44, 45]. An encoder compresses video, whereas a decoder uncompresses it.

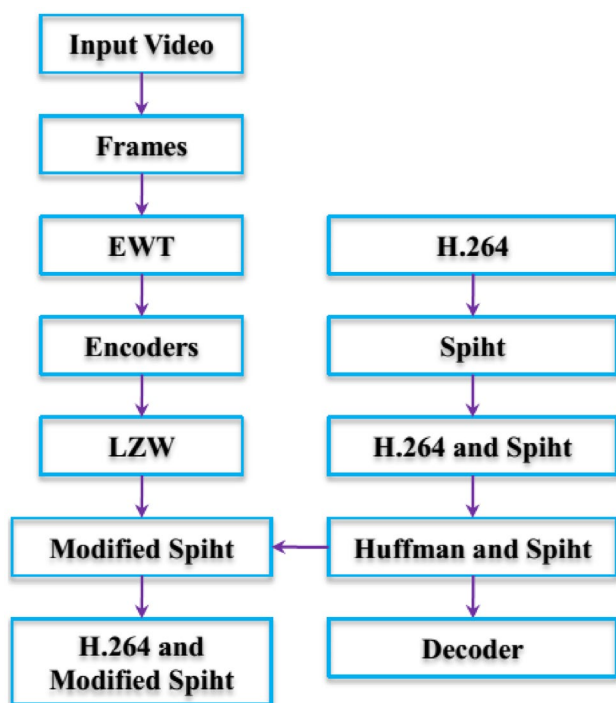


Fig. 1 Block diagram representation of VCWTDE

3.1.1 Encoding procedure

The video is initially uploaded and converted into frames by utilizing Matlab software. For decomposition, the EWT Transform is used. For the encoding process, H.264 is used in VCWHE, SPIHT in VCWSE, H.264 coding will encode low-frequency frames and the SPIHT algorithm will encode high-frequency frames. In VCWHSE, Huffman coding will encode low-frequency frames and the SPIHT algorithm will encode high-frequency frames. In VCWHUSE, LZW in VCWLE, Modified SPIHT in VCWMSE and H.264 coding will encode low-frequency frames and modified SPIHT algorithm will encode high-frequency frames in VCWHMSE. After the compression process is over, the performance of the system is computed with the help of various performance metrics. This video should be stored or transferred in its compressed form.

3.1.2 Decoding procedure

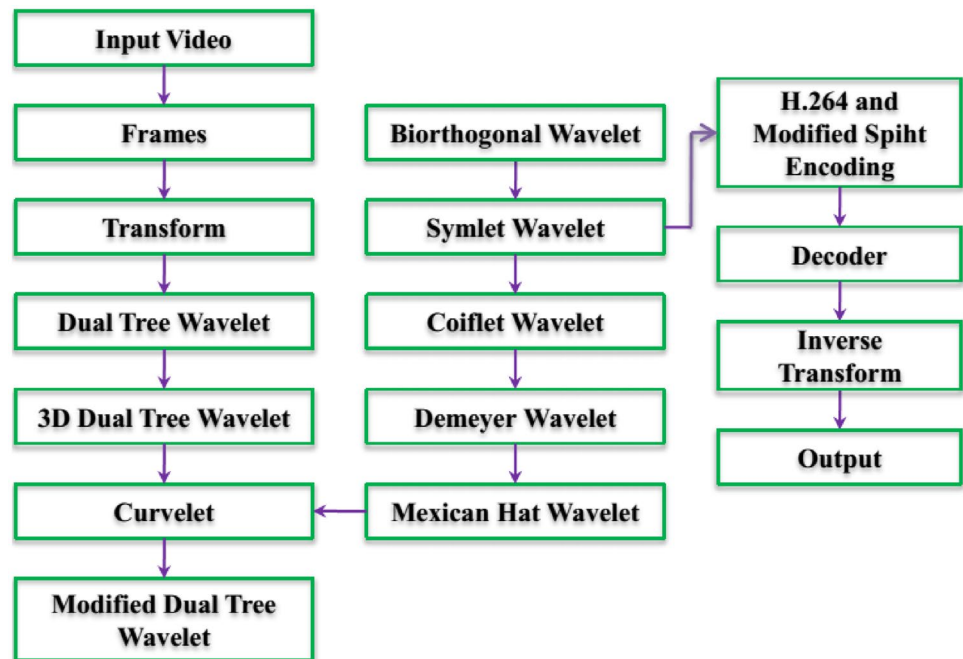
To decode the video, a decoding algorithm is employed. After that, the original video is rebuilt using the inverse EWT Transform and the video is reconstructed. In this phase, different encoding techniques were analyzed with

respect to WT. The transform and encoding step is most important while doing compression. EWT is taken for transforming, and it is fixed for all comparisons. Some of the encoding options are H.264, SPIHT, Huffman coding, LZW, and modified SPIHT. By merging two encoding algorithms, performance can be increased. A number of performance measures, such as CR, PSNR, MSE, and SSIM, show that H.264 with improved SPIHT offers better results. In comparison to previous techniques, VCWHMSE yields superior results. As a result, we conclude that the combination of H.264 and the modified SPIHT encoding approach is the best encoding method and use it as an encoder in future comparisons. In the next phase, the encoding part becomes constant in all techniques and takes different transforms for decomposition.

3.2 Video coding employing different transforms

As part of this phase, the transform part changes, while the encoding part remains the same throughout. Analysis of multiple VCWTDE encoding methods revealed that VCWHMSE produced better results, hence H.264 and modified SPIHT were selected as encoding strategies. The techniques included in this phase are Video Coding employing Biorthogonal Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCBWHMSE), Video Coding employing Symlet Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCSWHMSE), Video Coding employing Coiflet Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCCWHMSE), Video Coding employing Demeyer Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCDWHMSE), Video Coding employing Mexican Hat Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCMHWHMSE), Video Coding employing Dual-Tree Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCDTWHMSE), Video Coding employing Dual-Tree 3d Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCDT3W&HMSE), Video Coding employing Curvelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCCHMSE) and Video Coding employing Modified Dual-Tree Wavelet with coalescence of H.264 along with Modified Spiht Encoding technique (VCMDTWHMSE). Here, these methods are analyzed by utilizing various performance metrics. The block representation of VCDWCHMSE is shown in Fig. 2, which is a general diagram of all the techniques covered in this phase.

Fig. 2 Block representation of VCDWCHMSE



3.2.1 Encoding procedure

The video is initially uploaded and converted into frames by utilizing Matlab software. For decomposition, the Biorthogonal wavelet Transform is used for VCBWHMSE, Symlet Wavelet is used for VCSWHMSE, Coiflet Wavelet is used for VCCWHMSE, Demeyer Wavelet is used for VCDWHMSE, Mexican Hat Wavelet is used for VCMHWHMSE, DTCWT is used for VCDTWHMSE, 3D DTCWT is used for VCDT-3WHMSE, Curvelet is used for VCCHMSE, and modified DTCWT is used for VCMDTWHMSE. For the encoding process, H.264 is used to encode low-frequency frames and modified SPIHT is used to encode high-frequency frames. After the compression process is completed, the performance of this system is evaluated with the help of various parameter metrics. At last, the compressed video is either transferred or saved.

3.2.2 Decoding procedure

The video is decoded using H.264 and a modified SPIHT algorithm. Next, an inverse Transform for reconstructing the original video is applied and the video is reconstructed. In this phase, different encoding techniques were analyzed with respect to WT. Transform and encoding step is most important while doing compression. For doing the encoding process, the combination of H.264 and modified SPIHT is used and it is fixed for all comparisons. For doing transform

function, Biorthogonal Wavelet, Coiflet Wavelet, Demeyer Wavelet, Mexican Hat Wavelet, DTCWT, 3D DTCWT, Curvelet, and Modified DTCWT are utilized. With the help of various performance metrics like CR, PSNR, MSE, and SSIM, Modified DTCWT gives better performance related to other methods. Therefore video coding employing modified dual-tree wavelet with coalescence of H.264 along with modified SPIHT encoding technique gives better performance compared to other techniques. The working principle regarding this new technique is illustrated in the next phase.

4 Development and implementation of efficient video coding using multi-resolution techniques

The transforming part of this VCWTDE technique is stable, but the encoding part varies. When the performance of these approaches is compared using various performance metrics, VCWHMSE produces better results, therefore, the coalescence of H.264 together with the modified SPIHT encoding technique is chosen as the best encoding method. The transforming part of VCWTDE varies, whereas the encoding part remains constant. When evaluating the performance of these approaches using different performance metrics, VCMDTWHMSE produces better results, hence use modified DTCWT as the transform part. VCMDTWHMSE is compared with different methods such as VCWMSE,

VCWHMSE, and VCDTWHMSE. These three methods are related to the proposed work, therefore they are taken for comparison. In VCWMSE, video is compressed using wavelet and modified SPIHT encoding method. In VCWHMSE, video is compressed using wavelet and combination of H.264 and SPIHT encoding method. In VCDTWHMSE, video is compressed using DTCWT and a combination of H.264 and SPIHT encoding methods.

When using VCMDTWHMSE for video coding, the video is first changed by frames, and then those frames are decomposed using a modified dual-tree wavelet transform. H.264 and modified SPIHT are used for encoding, while the reverse technique is used for decoding. In this research, the dual-tree wavelet transform is modified by integrating Dual tree Complex Wavelet Transform along with Discrete Fractional Fourier Transform (DCWTDFFT). The modified DTCWT characterize the advantages of DTCWT and DFFT.

4.1 Development of modified transform DCWTDFFT

The DTCWT is a primary technique in our approach, and it will be derived by merging the DTCWT with the DFFT. DCWTDFFT, which combines the exhilarating mathematical processes of DTCWT and DFFT, is the proposed method. To deconstruct the transform matrix, the discrete Fourier Transform is employed, and Eigenvalues are used to find the DFFT. The DFT is first concentrated, which is a discrete version of the Fourier Transform. The N-point DFT set is described as in Eqs. (1) and (2) as

$$X(K) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) e^{-j2\pi \frac{nk}{N}}, k = 0, 1, \dots \dots N - 1 \tag{1}$$

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(K) e^{j2\pi \frac{nk}{N}}, n = 0, 1, \dots \dots N - 1 \tag{2}$$

Here, $\frac{1}{\sqrt{N}}$ is a normalization parameter and it formulates the DFT and IDFT unique.

The N-point DFT in Eq. (1) can be denoted in a matrix format and it is shown in Eq. (3) as:

$$F_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{-j\frac{2\pi}{N}} & e^{-j\frac{2\pi}{N}2} & \dots & e^{-j\frac{2\pi}{N}(N-1)} \\ 1 & e^{-j\frac{2\pi}{N}2} & e^{-j\frac{2\pi}{N}4} & \dots & e^{-j\frac{2\pi}{N}2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\frac{2\pi}{N}(N-1)} & e^{-j\frac{2\pi}{N}(N-1)2} & \dots & e^{-j\frac{2\pi}{N}(N-1)(N-1)} \end{bmatrix} \tag{3}$$

By utilizing the idea of a complex wavelet then the complex wavelet operation is specified as

$$f(t) = \sum_{n=-\infty}^{\infty} c_{j_0,n}^c \left(\varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) \right) + \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} D_{j,n}^c \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right), \tag{4}$$

where $C_{j_0,n}^c$ and $D_{j,n}^c$ becomes the scaling and wavelet quantities cooperated along with the complex wavelet transform and are stated as in Eqs. (5)–(8)

$$c_{j_0,n}^c = \langle f, \varphi_{j_0} \rangle = \int_{n=-\infty}^{\infty} f(t) \varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) dt \tag{5}$$

$$C_{j_0,n}^c = \langle f, \varphi_{j_0} \rangle = C_{j_0,n}^1 + iC_{j_0,n}^2 \tag{6}$$

$$D_{j,n}^c = \langle f, \psi_j \rangle = \int_{n=-\infty}^{\infty} f(t) \varphi_{j_0}^2(t-n) + i\varphi_{j_0}^1(t-n) dt \tag{7}$$

$$D_{j,n}^c = \langle f, \psi_j \rangle = D_{j,n}^1 + iD_{j,n}^2. \tag{8}$$

Using Eqs. (6) and (8), Eq. (4) can be rewritten as and it is shown in Eq. (9) as

$$f(t) = \sum_{n=-\infty}^{\infty} \left(C_{j_0,n}^1 + iC_{j_0,n}^2 \right) \left(\varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) \right) + \sum_{i=j_0}^{\infty} \sum_{n=-\infty}^{\infty} \left(D_{j,n}^1 + iD_{j,n}^2 \right) \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right). \tag{9}$$

The Eq. (9) can be rearranged as

$$f(t) = \sum_{n=-\infty}^{\infty} c_{j_0,n}^1 \left(\varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) \right) + i \sum_{n=-\infty}^{\infty} c_{j_0,n}^2 \left(\varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) \right) + \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} D_{j,n}^1 \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right) + i \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} D_{j,n}^2 \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right). \tag{10}$$

Then, changing the real and imaginary elements from the equation, we get the equation as

$$f(t) = \left(\sum_{n=-\infty}^{\infty} C_{j_0,n}^1 \left(\varphi_{j_0}^1(t-n) + i\varphi_{j_0}^2(t-n) \right) + \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} D_{j,n}^1 \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right) \right) + i \left(\sum_{n=-\infty}^{\infty} C_{j_0,n}^2 \left(\varphi_{j_0,n}^2(t-n) + i\varphi_{j_0}^2(t-n) \right) + \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} D_{j,n}^1 \left(\psi_j^1(t-n) + i\psi_j^2(t-n) \right) \right). \tag{11}$$

Based on the earlier equation, it was noted that two wavelet tree structures are calculated by utilizing complex-valued scaling and wavelet operations. Accordingly, this transform is basically represented as DTCWT. The DTCWT's most important principle is that real and imaginary divisions of the trees have the ability to efficiently reproduce data. Accordingly, the inverse DTCWT will be considered on the inverse wavelet transform of the real and the imaginary trees.

Because the DTCWT delivers two quantities in comparison to the WT, it is twice as valuable. As the parameters of the operation rise, the computational time will increase. Along with an N -d functions $f(t_k): k=1, 2, \dots, N$, the time difficulty of the DTCWT is provided as shown in Eq. (12).

$$T_{DT-CWT} = 2^N T_{WT}. \tag{12}$$

Here, T_{WT} becomes the time difficulty of the WT in the N -d function. As a result, the DTCWT corrects the WT's shortcomings at the expense of time complexity. A smart way to reduce time difficulties is to choose any real or imaginary tree that corresponds to the requirements for the required basis. When a real tree is chosen for operation, it is referred to as the real DTCWT. When an imagined tree

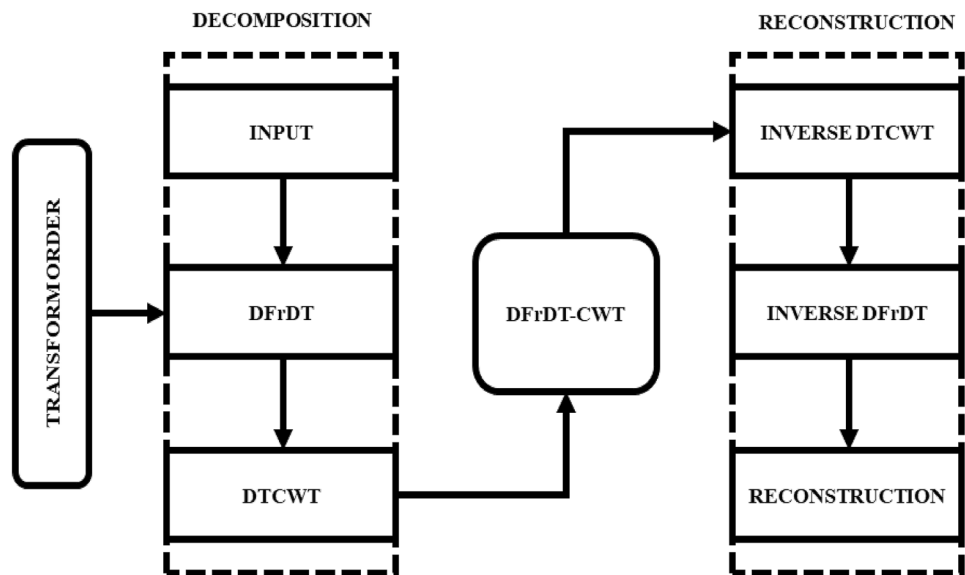
is chosen for operation, it is referred to as the imaginary DTCWT.

Typically, the DFrDT-CWT of any operation is proportional to the FFT combined with the DT-CWT to obtain the DFrDT-CWT of the input operation. The basic concepts of the DFrDTCWT are expressed as:

1. DTCWT function: the DFrDT-CWT of sequence $\alpha = 0$, i.e. utilizing $\alpha = 0$ for obtaining the DTCWT of the input.
2. Operator with Double Frequency: the definition of the operator with double frequency is that the input signal is changed in two separate transforms in succession. The DFrDT-CWT operated $\alpha = \pi/2$ will be the operator with double frequency and the DFrDT-CWT with sequence $\alpha = \pi/2$ provides the double frequency conversion output.
3. Successive program of the DFrDT-CWT: successive functions of the DFrDT-CWT are identical to a single transform and the order is identical to the addition of the separate orders.

A better way to make the DFrDT-CWT easier to execute is also suggested. The DFrDT-CWT was an action on the DTCWT in the discrete fractional Fourier domain, according to the previous description. Because of the rotation of the time-frequency plane across an arbitrary angle, the DFrFT

Fig. 3 Decomposition and reconstruction process for DFrDT-CWT



includes a different concept for representing information in the spatial and frequency domains. The DTCWT has a multi-resolution application in this regard. The DFrDTCWT, which displays the multi-resolution property, is the result of the collaboration of these two domains.

As a result, the forward DFrDT-CWT capability is obtained by computing the DFrFT on the input signal along with the optimal fractional order in conjunction with the DTCWT. Although the regeneration can be completed to return to the input signal's plane by increasing the inverse DTCWT in conjunction with the inverse DFrFT. Hence, the forward DFrDT-CWT is acquired by first computing the DFrFT along the optimal fractional order α on the input signal with the DTCWT, although the regeneration can be completed by enhancing the inverse DTCWT in connection with the inverse DFrFT to go back to the plane of the input signal. Figure 3 represents the decomposition and reconstruction process for DFrDT-CWT.

4.1.1 Decomposition

1. Transform Order (α) Optimization: in this process, the optimal quantity of the transform order (α) is measured.
2. Perform the DFrFT along with the transform.
3. Operate the DTCWT on the attained transformed signal from the earlier process.

4.1.2 Reconstruction

1. Operate the inverse DTCWT on the attained transformed signal.
2. Operate the inverse DFrFT of transform order α along with the transformed signal that is attained.

4.2 Development of modified SPIHT algorithm

The SPIHT is a commonly used compression method for wavelet-transformed images. SPIHT is a much easier, systematic, and fully inserted codec that provides better image conveyance, perfect coordination among distortion defense, and create information on demand, but it does have a few drawbacks that must be overcome in order for it to be used effectively. The slow processing speed is, however, one of the most important disadvantages. According to experimental results, the original SPIHT algorithm has a low coding efficiency since it uses a lot of bits to encode insignificant coefficients. Furthermore, while the energy of an original image is concentrated on the wavelet-transformed image's lost frequency band, the original SPIHT method stores all of

the wavelet coefficients in a comparable manner. As a result, a huge number of redundant 0 bits are produced, which has a significant impact on its coding efficiency. Scanning wavelet coefficients from largest to smallest during the encoding process, on the other hand, consumes a lot of bits for insignificant coefficients.

This work proposes a development approach for encoding the LIS and LIP to address the aforementioned issues. The experimental results show that the modified SPIHT method outperforms the original SPIHT algorithm in terms of PSNR values and visual quality by drastically reducing the number of output 0 bits. In this stage, we'll reduce the redundancy in the original SPIHT-based compression technique. In this paper, we improve the SPIHT (Modified SPIHT) algorithm by presenting a new scheme for encoding the LIS and LIP to control the redundancies that remain in the traditional SPIHT. The improved algorithm's approach is detailed below.

4.2.1 Modified scheme in LIS coding

Step 1

- (a) Set threshold value as $T0 = 2n$, for all the coefficient 'n' is the largest integer in the logarithm of the highest merit.
- (b) Starting value of LIS is set as the child node of the root node.

Step 2

- (a) Obtain the set in LIS. If LIS is D-set, then proceed to step 3 if LIS is L-set, then proceed to step 4.
- (b) If not proceed to step 6.

Step 3

- (a) Result will be "0" when coefficient in the D-set is not vital and step 6 will be performed. If this is not the case, the outcome will be "1" for the coefficient in the D-set.
- (b) If the coefficient is more than 4, the result will be "1" again, and the coefficient will be linked to L-set. If it has a child node, step 2 will be performed.
- (c) When the coefficient is less than 4, the O-set is processed, and the coefficient is placed in the O-set; when the coefficient is more than $T0$, the second result is "1," and the coefficient is placed in the LIS.
- (d) When the coefficient is less than $T0$ and placed in LIS, the result will be "0."
- (e) Then upgrade LIS.

Step 4

- (a) When the coefficient in the L-set is not important, the result will be “0,” and step 6 will be performed.
- (b) If not, when the coefficient in the L-set is critical, the results will be “1” and the L-set will be divided into four D-sets.
- (c) Upgrade LIS.

Step 5

- (a) When the coefficient in O-set is not important, the output is “0,” and the process moves on to step 6.
- (b) The coefficient in the O-set is crucial, because the output is “1” and the L-set is divided into four D-sets.
- (c) Upgrade LIS.

Step 6

- (a) The operation is finished when the LIS is clear.
- (b) If the LIS is not clear, it proceeds to step 2.

4.2.2 Modified scheme in LIP coding

For the original SPIHT algorithm, wavelet coefficients of LLn, LHn, HLn and HHn bands are arranged in LIP for *n* level DWT. These coefficients in LLn are relatively larger than those in other bands, while all of them in LIP are processed simultaneously. The original SPIHT first traverses through LLn and then other bands (LHn, HLn, and HHn). Due to the weakness of LIP, a large quantity of 0 bits unnecessary for accurate reconstruction is encoded in the bitstream.

It is worth noting that there is no significant coefficient in LLn in the case of the threshold. The optimization of LIP is proposed in the aforementioned study [42, 43] to stop encoding inconsequential coefficients once all significant coefficients have been encoded. The main idea of the proposed method is to add an additional bit *L* representing the number of significant coefficients in LIP encoded. In LIP, the sum of significant values is designated as *S*, and it is calculated, after which the coefficients in LIP are encoded until *L* is similar to *M* (maximum coefficient). The proposed optimization approach for the original SPIHT can save a significant number of bits that would otherwise be needed to encode insignificant coefficients.

5 Results and discussion

In this phase, the results of the experiments are discussed, and a comparative analysis of the result is presented. The optimum and efficient solutions are proposed to overcome

Table 1 Details of the input video set

Input videos	Length (s)	Frame height	Frame width	Format	Size
VIP traffic	8	120	160	Avi	6.59 MB
Video	59	240	320	Avi	1.43 MB
Dance	30	240	320	Avi	880 kb
Earth	30	360	640	Avi	809 kb
Foreman	30	–	–	Avi	466 kb

the identified problems. Three different phases with classification schemes were re-explained and evaluated. The proposed methods were simulated and tested using Mat lab R2014a. The results of the existing and proposed algorithms are checked using various performance analyses. The performance metrics such as PSNR, SSIM, CR, and MSE are obtained based on the following equations.

Mean Square Error: the distortion between the original and reconstructed image is represented as Mean Square Error (MSE).

$$MSE = \frac{1}{S} \sum_{i=1}^a \sum_{j=1}^b [x(a, b) - y(a, b)]^2 \tag{13}$$

Peak Signal to Noise Ratio (PSNR): the ratio between the maximum power value of a signal to the distortion noise which affects the quality of the video. A maximum PSNR value represents an improved video quality [48, 49]. The PSNR value is computed using the below equation:

$$PSNR = 10 \log \left(\frac{250^2}{MSE} \right). \tag{14}$$

Structural Similarity Index (SSIM): this process involves two videos/images and checks the similarity between the original and reconstructed image. Different factors such as local contrast sensitivity, local structure similarity, and local luminance sensitivity are taken into account.

$$SSIM = \frac{(2\mu_a \mu_b + K1) (2\sigma_{ab} + K2)}{(\mu_{a2} \mu_{b2} + K1) (\sigma_{a2} + \sigma_{b2} + K2)} \tag{15}$$

The mean values of the local luminance are represented as μ_a and μ_b . The standard deviation of the actual and reconstructed images is represented as σ_a and σ_b based on the luminance similarity. The local contrast sensitivity of the original and reconstructed image is represented as *K1* and *K2*.

Compression Ratio (CR): it measures the capability of the data compression technique by comparing the reconstructed image size to the actual image size.

Fig. 4 Various input videos used in this video coding **a** VIP traffic. **b** Video. **c** Dance. **d** Earth. **e** Foreman

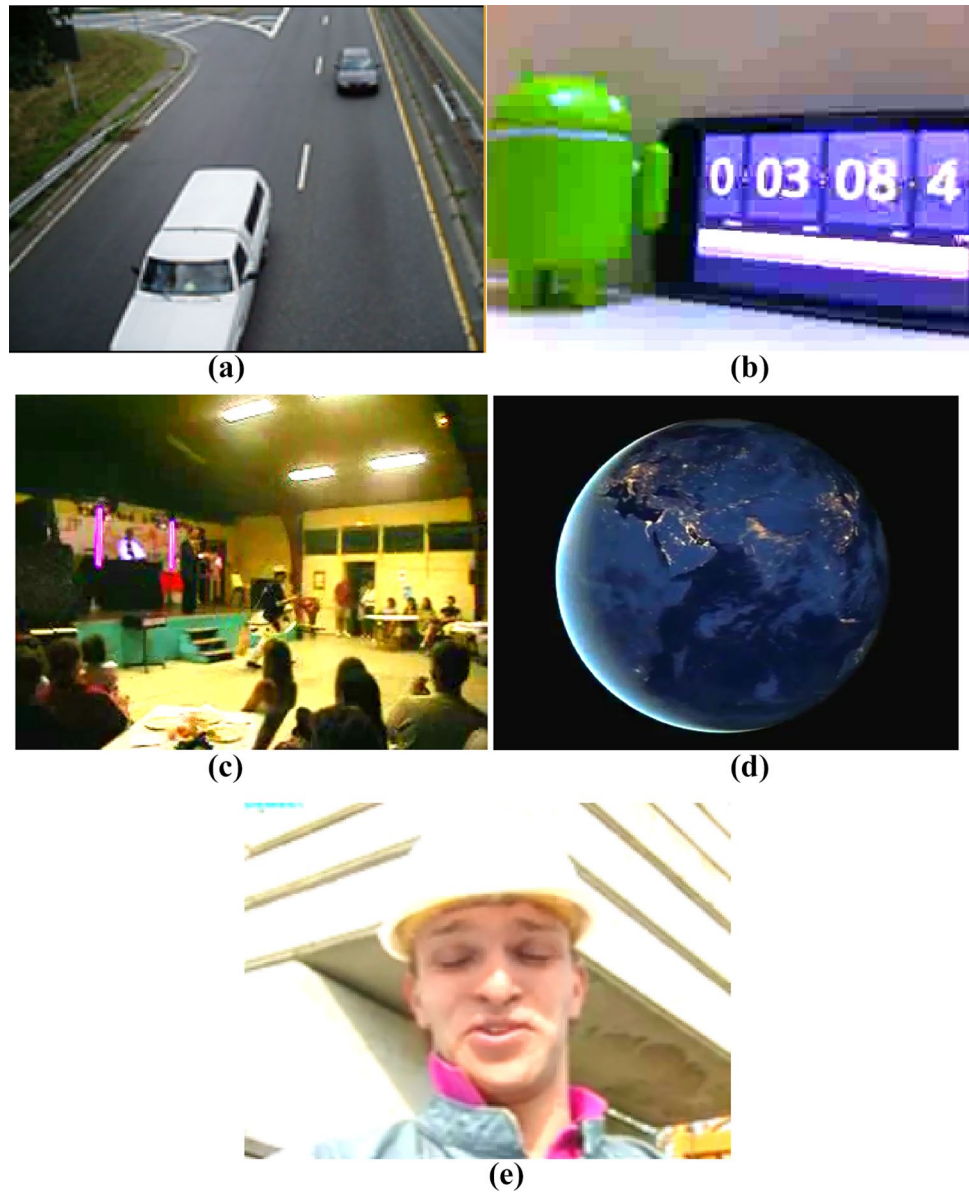


Table 2 PSNR, CR, SSIM, and MSE representation of VCWTDE by using VIP traffic input video

Techniques	Input video	PSNR	MSE	SSIM	CR
VCWHE	VIP traffic	39.28	7.67	0.26	0.72
VCWSE	VIP traffic	43.58	2.85	0.53	1.64
VCWHSE	VIP traffic	47.09	1.27	0.8	2.18
VCWHUSE	VIP traffic	45.60	1.79	0.75	2.05
VCWLE	VIP traffic	43.16	3.14	0.52	1.61
VCWMSE	VIP traffic	43.91	2.64	0.62	1.69
VCWHMSE	VIP traffic	50.49	0.58	0.82	2.21

$$CR = \frac{i_{Size}}{C_{size}}, \quad (16)$$

where i_{Size} is the input image size and C_{size} is the compressed output image size. Table 1 represents the details of input video sets that are used in video compression for evaluating the performance of existing and proposed techniques.

Figure 4 represents various input videos. (a) represents of VIP traffic, (b) represents video, (c) represents Dance, (d) represents Earth and (e) represents Foreman.

Fig. 5 Performance characteristics of VCWTDE by using VIP traffic input video

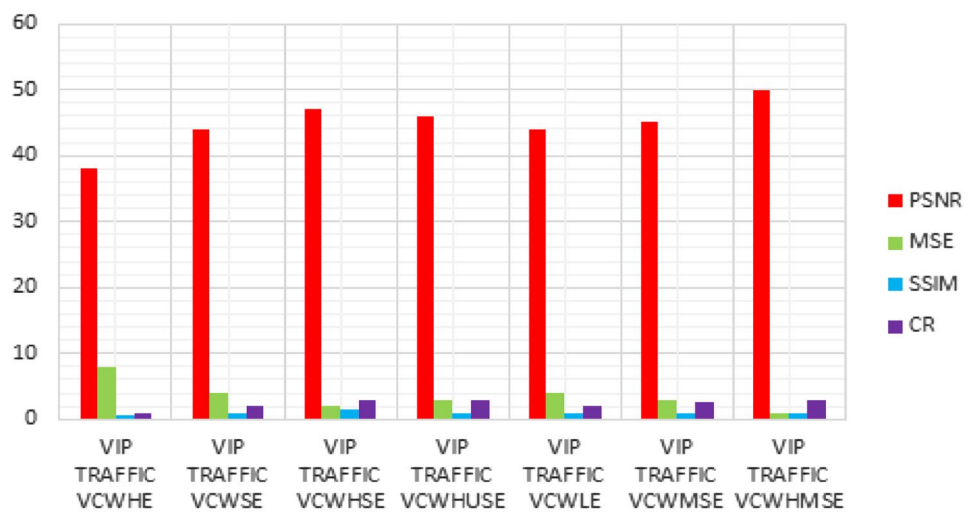


Table 3 PSNR, CR, SSIM, and MSE representation of VCDWCHMSE by using VIP traffic input video

Techniques	Input video	PSNR	MSE	SSIM	CR
VCBWHMSE	VIP traffic	39.28	7.67	0.26	0.53
VCSWHMSE	VIP traffic	47.44	1.17	0.53	1.64
VCCWHMSE	VIP traffic	48.26	0.97	0.81	2.18
VCDWHMSE	VIP traffic	48.35	0.95	0.75	2.05
VMHWHMSE	VIP traffic	48.89	0.84	0.72	1.61
VCDTWHMSE	VIP traffic	48.17	0.99	0.71	1.69
VCDT3WHMSE	VIP traffic	49.05	0.81	0.77	2.21
VCCHMSE	VIP traffic	47.13	1.26	0.62	0.76
VCMDTWHMSE	VIP traffic	52.96	0.33	0.82	2.64

5.1 Performance analysis of VCWTDE

The video compression methods are tested using different input video sequences. The results are checked using various performance analyses. PSNR, SSIM, CR, and MSE are used to check the efficiency of the techniques, and Mat lab R2014a is used for execution. The input videos used here are VIP traffic, foreman, video, dance, and earth. The methods which are used for analyzing VCWTDE are VCWHE, VCWSE, VCWHSE, VCWHUSE, VCWLE, VCWMSE, and VCWHMSE. For every method, efficiency is checked using various quality metrics and corresponding results are noted.

For doing the transform part in this process, take EWT is for decomposition and it is constant for all comparisons. The encoding part is varied and chooses H.264, SPIHT, Huffman

Fig. 6 Performance characteristics of VCDWCHMSE by using VIP traffic input video

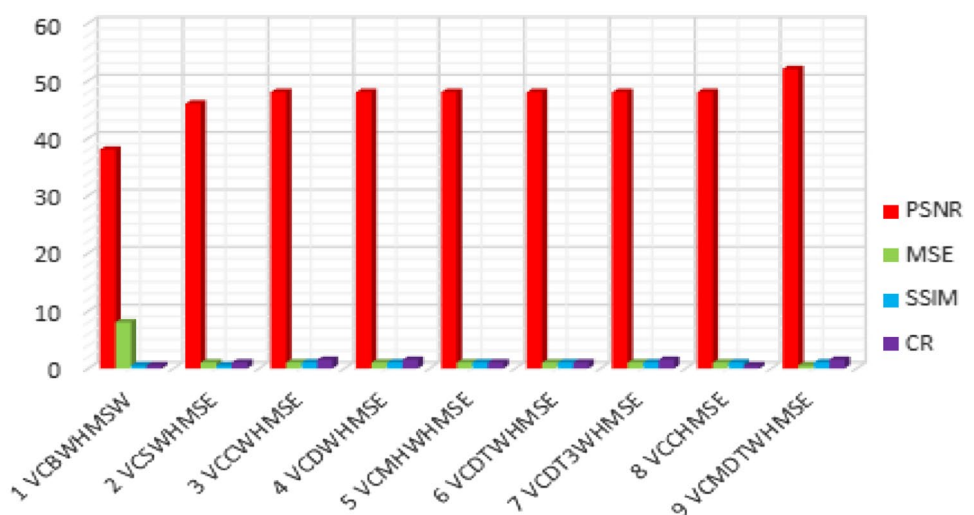


Table 4 PSNR, CR, SSIM and MSE representation of efficient video coding using multi-resolution techniques

Techniques	Input video	PSNR	MSE	SSIM	CR
VCWMSE	VIP traffic	43.91	2.64	0.62	1.69
VCWHMSE	VIP traffic	50.49	0.58	0.82	2.21
VCDTWHMSE	VIP traffic	48.17	0.99	0.71	1.69
DES [34]	VIP traffic	48.65	0.98	0.69	1.58
ES-KFE technique [35]	VIP traffic	48.21	0.98	0.68	1.72
Equal partition-based clustering approach [36]	VIP traffic	39.65	3.01	0.59	0.99
Event bagging [37]	VIP traffic	38.65	3.21	0.58	0.88
Deep event learning boost-up approach [38]	VIP traffic	47.54	0.82	0.65	1.15
SOMES [39]	VIP traffic	48.69	0.95	0.69	1.24
ESUMM [40]	VIP traffic	48.65	0.98	0.68	1.21
EVS-DK [41]	VIP traffic	48.17	0.95	0.71	1.69
Key-lectures [42]	VIP traffic	49.32	0.74	0.69	1.78
Proposed VCMDTWHMSE	VIP traffic	52.96	0.33	0.82	2.64

coding, LZW, and modified SPIHT. Here, combinations of two encoding techniques are used for getting better performance. With the help of various performance matrices like CR, PSNR, MSE, and SSIM, the combination of H.264 and modified SPIHT gives better performance. The experimental outcomes display that VCWHMSE shows better values by comparing other techniques.

Table 2 represents various performance comparisons of VCWTDE using VIP traffic input video and Fig. 5 denotes the performance characteristics of VCWTDE using VIP traffic input video.

5.2 Performance analysis of VCDWCHMSE

The results are checked by using various performance analyses. Here, PSNR, SSIM, CR, and MSE are utilized for checking the efficiency of the process. Matlab R2014a is used for executing the program. The input videos used for this evaluation are VIP traffic, foreman, video, dance, and earth. The methods which are included for analyzing

the VCDWCHMSE are VCBWHMSE, VCSWHMSE, VCCWHMSE, VCDWHMSE, VCMHWHMSE, VCDTWHMSE, VCDT3WHMSE, VCCHMSE, and VCMDTWHMSE.

From the previous analysis, VCWHMSE gives better performance, therefore H.264 and modified SPIHT are selected as an encoding technique for improving the efficiency of video compression. For every method included in this evaluation, efficiency is checked by using various quality metrics and corresponding results are noted. In this technique, different encoding techniques are analyzed with respect to a constant encoder. The combination of H.264 and SPIHT is employed for the encoding process, and it is consistent across all methods included in this strategy. For doing transform function, Biorthogonal Wavelet, Coiflet Wavelet, Demeyer Wavelet, Mexican Hat Wavelet, DTCWT, 3D DTCWT, Curvelet, and Modified DTCWT are used. Modified Dual-Tree Wavelet improves performance by utilizing several performance matrices such as CR, PSNR, MSE, and SSIM. As a result, modified DTCWT is selected

Fig. 7 Performance characteristics of efficient video coding using multi-resolution techniques

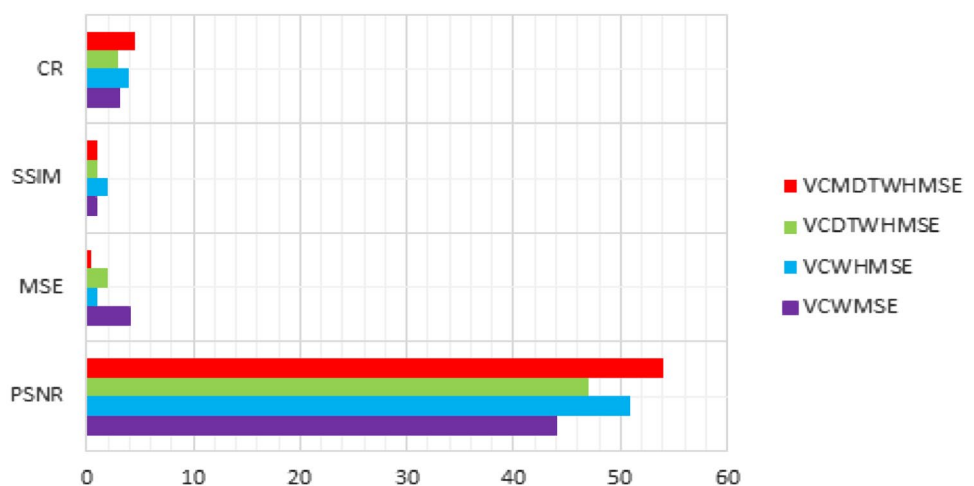
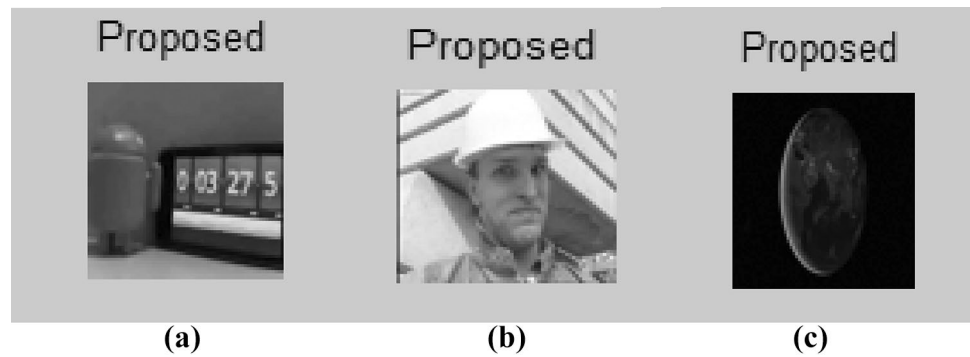


Fig. 8 Reconstructed output images of the proposed method using the input video **a** video, **b** foreman, and **c** Earth



as the optimal transform for video compression. Therefore video coding employing modified dual-tree wavelet with coalescence of H.264 along with modified SPIHT encoding technique gives better performance compared to other techniques.

Table 3 represents PSNR, CR, SSIM, and MSE representation of VCDWCHMSE by using VIP traffic input video and Fig. 6 represents the performance characteristics of VCDWCHMSE by using VIP traffic input video.

From Table 3 and Fig. 6, it is clear that the proposed transform modified dual-tree wavelet gives better results compared to the other techniques. As a result, this transform is chosen for the next comparison analysis, and video coding with modified DTCWT and modified encoder outperforms other techniques.

5.3 Performance analysis of efficient video coding using multi-resolution techniques

When evaluating performance using different performance metrics, the VCMDTWHMSE methodology produces better results to fix modified DTCWT as the transform component. Next, the proposed VCMDTWHMSE technique is compared with the state-of-art techniques such as DES: Fast and deep event summarization (DES) [34], Eratosthenes sieve-based keyframe extraction (ES-KFE) technique [35], Equal partition-based clustering approach [36], Event bagging [37], Deep event learning boost-up approach [38], Self-Organizing Map (SOM) technique for event summarization (SOMES) [39], Event summarization on scale-free networks (ESUMM) [40], Event video skimming using deep keyframe (EVS-DK) [41], and keyframes extraction in video lectures (Key-lectures) [42]. The other three methods already shown in the previous section are also used for comparison here. Here, the comparison is used for verifying that the proposed method performance is better than other methods.

Table 4 represents various performance comparisons of efficient video coding using multi-resolution techniques using VIP traffic input video. The performance of certain

techniques such as DES [34], ES-KFE technique [35], Equal partition-based clustering approach [36], Event bagging [37], Deep event learning boost-up approach [38], and SOMES [39] are low due to the large overlap between the objects and disordered motions. Techniques like ESUMM [40], EVS-DK [41], and Key-lectures [42] suffer from a lack of summarization capacity and treat each keyframe in the video identically. The performance characteristics of effective video coding utilizing multi-resolution methods with input video VIP traffic are depicted in Fig. 7.

This analysis shows that the proposed VCMDTWHMSE method (Fig. 8) achieves better results and comparable performance than other strategies, as well as overcoming their limitations. The performance factors PSNR, CR, SSIM, and MSE get better values compared to other methods in the table.

6 Conclusion

This work aims to develop an efficient video compression technique with the help of multi-resolution techniques and it is suitable for multimedia applications. The proposed VCMDTWHMSE methodology uses the modified DTCWT transform for decomposition and a combination of H.264 and modified SPIHT for encoding. In the first phase, different encoding techniques are analyzed and the best encoder is selected for doing compression. The combination of H.264 along with the modified SPIHT encoding technique is said to be the best encoding method based on the results obtained. In the second phase, different transform techniques are evaluated and the best transform is selected for doing compression. For the encoding process, the combination of H.264 and SPIHT is used and it is constant for all comparisons. Therefore, video coding employing modified DTCWT with coalescence of H.264 along with modified SPIHT encoding technique gives better performance. In the third phase, the development and implementation of efficient video coding using multi-resolution techniques were explained and

implemented. By comparing the performance using various performance metrics like PSNR, CR, SSIM the proposed method gives better results.

Funding Not applicable.

Availability of data and materials Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights This article does not contain any studies with human or animal subjects performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

- Sayood, K.: Introduction to data compression, 2nd edn. Morgan Kaufmann, Burlington (2000)
- Dass Member, R., Singh, L., Kaushik, S.: Video compression technique. *Int. J. Sci. Technol. Res.* **1**(10), 114–119 (2012)
- Shaikh, M.A., Badnerkar, S.S.: Video compression algorithm using motion compensation technique. *Int. J. Adv. Res. Electron. Commun. Eng.* **3**, 625–629 (2014)
- Wadd, S.S., Patil, S.B.: Video compression using DCT. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **4**(9), 394–398 (2014)
- Meiyazhagan, M., Sundaravadeivel, P.: High efficiency video compression using multiwavelet block coding. *UJEAS* **2**(2), 179–183 (2014)
- Taleb, S.A., Musafa, H.M., Khtoom, A., Gharaybih, K.: Improving LZW image compression. *Eur. J. Sci. Res.* **44**(3), 502–509 (2010)
- Suganya, G., Mahesh, K.: A survey: various techniques of video compression. *Int. J. Eng. Trends Technol.* **7**(1), 10–12 (2014)
- Gohil, R., Pandya, V.: A comparative study of different video compression technique. *Int. J. Adv. Res. Comput. Sci. Manag. Stud.* **3**(6), 39–43 (2015)
- Stobaugh, J.D.: Novel use of video and image analysis in a video compression system, pp. 1–61. University of Iowa Research Online, Iowa (2015)
- Yang, X., Fan, J., Wu, C., Zhou, D., Li, T.: NasmamSR: a fast image super-resolution network based on neural architecture search and multiple attention mechanism. *Multimed. Syst.* (2021). <https://doi.org/10.1007/s00530-021-00841-2>
- Sundararaj, V., Muthukumar, S., Kumar, R.S.: An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. *Comput. Secur.* **77**, 277–288 (2018)
- Vinu, S.: An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. *Int. J. Intell. Eng. Syst.* **9**(3), 117–126 (2016)
- Sundararaj, V.: Optimised denoising scheme via opposition-based self-adaptive learning PSO algorithm for wavelet-based ECG signal noise reduction. *Int. J. Biomed. Eng. Technol.* **31**(4), 325 (2019)
- Sundararaj, V., Selvi, M.: Opposition grasshopper optimizer based multimedia data distribution using user evaluation strategy. *Multimedia Tools Appl.* **80**(19), 29875–29891 (2021)
- Ravikumar, S., Kavitha, D.: CNN-OHGS: CNN-oppositional-based Henry gas solubility optimization model for autonomous vehicle control system. *J. Field Robot.* (2021). <https://doi.org/10.1002/rob.22020>
- Ravikumar, S., Kavitha, D.: IoT based home monitoring system with secure data storage by Keccak-Chaotic sequence in cloud server. *J. Ambient Intell. Human. Comput.* (2020). <https://doi.org/10.1007/s12652-020-02424-x>
- Rejeesh, M.R., Thejaswini, P.: MOTF: Multi-objective Optimal Trilateral Filtering based partial moving frame algorithm for image denoising. *Multimedia Tools Appl.* **79**(37), 28411–28430 (2020)
- Kavitha, D., Ravikumar, S.: IOT and context-aware learning-based optimal neural network model for real-time health monitoring. *Transa. Emerg. Telecommun. Technol.* **32**(1), e4132 (2021)
- Jose, J., Gautam, N., Tiwari, M., Tiwari, T., Suresh, A., Sundararaj, V. and Rejeesh, M.R.: An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion. *Biomed. Signal Process. Control* **66**, 102480 (2021)
- Edwin, A.C. and Madheswari, A.N.: Job scheduling and VM provisioning in clouds. In 2013 Third International Conference on Advances in Computing and Communications pp. 261–264. IEEE (2013)
- Hassan, B.A., Rashid, T.A., Mirjalili, S.: Formal context reduction in deriving concept hierarchies from corpora using adaptive evolutionary clustering algorithm star. *Complex Intell. Syst.* (2021). <https://doi.org/10.1007/s40747-021-00422-w>
- Azath, M., Banu, R.W. and Madheswari, A.N.: Improving fairness in network traffic by controlling congestion and unresponsive flows. In International Conference on Network Security and Applications, pp. 356–363. Springer, Berlin, Heidelberg (2011)
- Nirmal Kumar, S.J., Ravimaran, S., Alam, M.M.: An Effective Non-Commutative Encryption Approach with Optimized Genetic Algorithm for Ensuring Data Protection in Cloud Computing. *Comput. Model. Eng. Sci.* **125**(2), 671–697 (2020)
- Gowthul Alam, M.M., Baulkani, S.: Reformulated query-based document retrieval using optimised kernel fuzzy clustering algorithm. *Int. J. Bus. Intell. Data Min.* **12**(3), 299 (2017)
- Alam, M.G., Baulkani, S.: A hybrid approach for web document clustering using K-means and artificial bee colony algorithm. *Int. J. Intell. Eng. Syst.* **9**(4), 11–20 (2016)
- Nisha, S., Madheswari, A.N.: Secured authentication for internet voting in corporate companies to prevent phishing attacks. *IJETCSE* **22**(1), 45–49 (2016)
- Ambalgi, S.: A comparative study of video compression techniques. *IJSRSET* **1**(3), 174–177 (2015)
- Jha, S., Ranga, K.K.: Advanced video compression—a breakthrough to speed up internet downloads. *IJERT* **3**, 1–6 (2015)
- Lee, J.-H., Jeong, S., Kim, B.-G., Jang, K.-S., Choi, J.S.: Fast video encoding algorithm for the internet of things environment based on high efficiency video coding. *Int. J. Distrib. Sensor Netw.* **1**, 1–10 (2015)
- Anugrace Rani, R., Muthulakshmi, K.: Performance and analysis of video compression using SPIHT algorithm. *Int. J. Innov. Res. Comput. Commun. Eng.* **3**(3), 86–94 (2015)
- Ginesu, G., Giusto, D.D., Pearlman, W.A.: Lossy to lossless SPIHT-based volumetric image compression. *Int. Conf. ICASSP* **3**, 1–5 (2004)

32. Shang, X., Zhao, H., Wang, G., Zhao, X., Zuo, Y.: A novel objective quality assessment method for transcoded videos from H.264/AVC to H.265/HEVC utilizing probability theory. *IEEE Trans. Broadcast.* **65**, 777–781 (2019)
33. Feng, W., Ju, W., Li, A., Bao, W., Zhang, J.: High-efficiency progressive transmission and automatic recognition of wildlife monitoring images with WISNs. *IEEE Access* **7**, 161412–161423 (2019)
34. Sivaranjani, J., Madheswari, A.N.: A novel technique of motif discovery for medical big data using hadoop. In: 2017 Conference on Emerging Devices and Smart Systems (ICEDSS) (pp. 214–217). IEEE (2017)
35. Kumar, K., Shrimankar, D.D., Singh, N.: Eratosthenes sieve based key-frame extraction technique for event summarization in videos. *Multimed. Tools Appl.* **77**(6), 7383–7404 (2018)
36. Kumar, K., Shrimankar, D.D., Singh, N.: Equal partition based clustering approach for event summarization in videos. In: 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems IEEE (SITIS), pp. 119–126 (2016)
37. Kumar, K., Shrimankar, D.D., Singh, N.: Event bagging: a novel event summarization approach in multiview surveillance videos. In: 2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC), pp. 106–111 (2017)
38. Kumar, K., Shrimankar, D.D.: Deep event learning boost-up approach: delta. *Multimed. Tools Appl.* **77**(20), 26635–26655 (2018)
39. Pal, T., Bit, S.D.: An energy-saving video compression targeting face recognition of disaster victim. *Multimed. Syst.* 1–21 (2021)
40. Kumar, K., Shrimankar, D.D.: ESUMM: event summarization on scale-free networks. *IETE Tech. Rev.* (2018). <https://doi.org/10.1080/02564602.2018.1454347>
41. Kumar, K.: EVS-DK: event video skimming using deep keyframe. *J. Vis. Commun. Image Represent.* **58**, 345–352 (2019)
42. Zhang, Q., Wang, Y., Huang, L., Jiang, B., Wang, X.: Fast CU partition decision for H. 266/VVC based on the improved DAG-SVM classifier model. *Multimedia Sys.* **27**(1), 1–14 (2021)
43. Kumar, K.: Text query based summarized event searching interface system using deep learning over cloud. *Multimed. Tools Appl.* **80**(7), 11079–11094 (2021)
44. Yan, C., Gong, B., Wei, Y., Gao, Y.: Deep multi-view enhancement hashing for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(4), 1445–1451 (2020)
45. Yan, C., Li, Z., Zhang, Y., Liu, Y., Ji, X., Zhang, Y.: Depth image denoising using nuclear norm and learning graph model. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **16**(4), 1–17 (2020)
46. Singh, P., Singh, P.: Design and implementation of EZW & SPIHT image coder for virtual images. *Int. J. Comput. Sci. Secur. IJCSS* **5**(5), 433–442 (2011)
47. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 454(1971), pp.903-995. (1998)
48. Yan, C., Hao, Y., Li, L., Yin, J., Liu, A., Mao, Z., Chen, Z., Gao, X.: Task-adaptive attention for image captioning. *IEEE Trans. Circuits Syst. Video Technol.* (2021). <https://doi.org/10.1109/TCSVT.2021.3067449>
49. Yan, C., Teng, T., Liu, Y., Zhang, Y., Wang, H., Ji, X.: Precise no-reference image quality evaluation based on distortion identification. *ACM Trans. Multimed. Comput. Commun. Appl.* (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.