



Discrete transformation technique for video compression

Gabriel Babatunde Iwasokun¹ · Monday Olutayo Olaoye²

Received: 4 November 2020 / Accepted: 9 March 2021 / Published online: 24 March 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

A large composition of image, sound, motion, video and other signal sequences is always a rich source of useful and reliable information. With the proliferating cost and complexities of mass storage media, the need for redundancy and complex transmission-free signals is on the rise. Furthermore, the growing increase in the number of video-specific applications has heightened the need for digital multimedia technologies that emphasized bit compression. This paper presents a Discrete Transform Technique (DTT) that leveraged the integration of Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) for higher scale video compression. The experimental study of the new technique was based on some online videos with varying sizes, sources and formats. The analysis of the experimental values obtained for standard performance metrics such as maximum difference, mean square error and peak signal to noise ratio for the pre- and post-compression videos portrayed a very significant and encouraging impact of the new technique. Comparative analysis based on compression ratios derived using the new technique and some recent video compression techniques also established a very impressive performance of the new technique.

Keywords Video compression · DWT · DCT · DTT · JPEG · Lossy and lossless compression

1 Introduction

Compression is a fundamental requirement for modern-day efficient and effective management of signals, images, videos and other related data. It is primarily performed as means of eliminating unwanted or redundant components from data with a view to reducing storage requirement and transmission time as well as endearing compact representation and optimal bandwidth utilization [1]. The rudiment of data compression is presented in Fig. 1 [2].

The encoder applies discrete transform on the original data to obtain transform coefficients that are subsequently quantized and entropy coded to obtain a code stream (bit stream). The decoder is used to reverse the operation of the encoder [3]. Video or image compression is classified as

lossless or lossy. Lossy compression is conceptualized in Fig. 2 with a predictor that interfaces with a lossless encoder before releasing its final output to the entry encoder. If the entry encoded stream passes the lossless test, it is passed to the final stage as the compressed image otherwise it is passed to the lossless encoder for further encoding. Lossless compression encodes all information from a still or motion image based on Huffman, run length, arithmetic, bit plane, dictionary or any other encoding technique and it is most suitable for archival, medical and other related images [4, 5].

The basic flow of lossy compression is presented in Fig. 3. The source encoder encodes the input video and passes its output for quantization. The quantized image is subsequently passed to the entropy encoder for transformation into the final output. Lossy compression often results in loss of little information and it is specifically suitable for compression of natural images in applications where minor loss of conformity is tolerable for achieving a sizeable diminution in bit rate [6–8].

Compression offers optimized and cost-effective representation, storage and transmission of still or motion images [9]. Retrospectively, video coding standards like Joint Photographic Expert Group (JPEG), Moving Picture Experts Group (MPEG) and H.26× had been widely explored but

✉ Gabriel Babatunde Iwasokun
gbiwasokun@futa.edu.ng
Monday Olutayo Olaoye
montagdavid777@gmail.com

¹ Department of Software Engineering, Federal University of Technology, Akure, Nigeria

² Department of Computer Science, Federal University of Technology, Akure, Nigeria

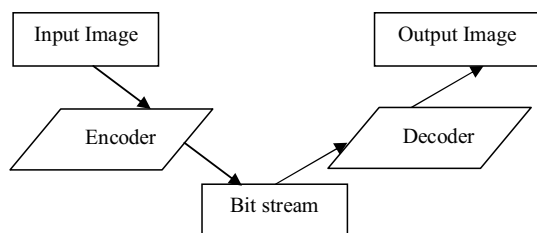


Fig. 1 Basic flow of image compression

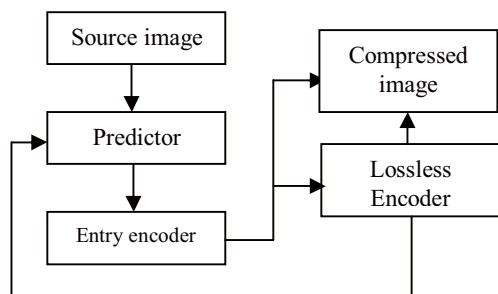


Fig. 2 Lossless image compression

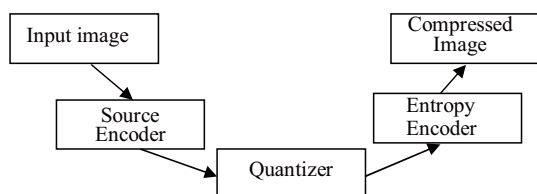


Fig. 3 Lossy image compression

with no provision for effortless and swift compression and decompression due to the complex nature of the encoders and decoders [10]. These conventional coding techniques require reassessment when adopted for video surveillance, telemedicine, space and satellite imaging, live broadcasting and other related applications. The escalating and multidimensional form of present-day activities results in enormous video data sizes which have heightened the demand for more flexible and result-oriented compression techniques.

Some of the established video compression techniques include motion JPEG, motion JPEG 2000, MPEG-1, MPEG-2, MPEG-4, H.261, H.263 and H.264. A motion JPEG is often coded as a string of JPEG pictures displayed over time with flexibility in terms of quality as well as compression ratio. However, its compression ratio is often lower than what obtains for other video compression techniques. JPEG 2000 offers a slightly better compression ratio compared to JPEG but with greater complexity. MPEG-1 video compression standard is based on the same technique in JPEG as well

as methods for efficient coding and limiting the bandwidth consumption. MPEG-2 is primarily targeted at attaining very high picture quality in Television transmission and other applications capable of 4 Mbps and higher data rates. It is compatible with MPEG-1 and supports interlaced video formats, increased image quality and other features for promoting high definition video transmission. It is however not suitable for real-time surveillance applications. MPEG-4 video compression technique is useful to establishing applications with minimal bandwidth consumption and high photograph quality. However, it is susceptible to high encoding complexities coupled with very weighty computational load for motion evaluation. H.261 is also a video coding standard based on DCT and with specific applications in video conferencing. It promotes full encoding of certain video frames and encoding of the differences between others with prediction, block transformation, quantization and entropy coding being its main elements. H.263 is a video coding technique for achieving low bit-rate visual telephone services in multimedia terminals. Though structurally similar and with backward compatibility with H.261, it offers superior picture quality. H.264 video compression technique is used to deliver optimal bit rate reduction, tolerate transmission errors over various networks, support low latency capabilities, offer simpler implementation and exact match decoding [1, 9]. Transform coding constitutes an integral component of contemporary image/video processing applications and relies on the premise that pixels in an image exhibit a certain level of correlation with neighboring pixels [11]. It is often used to map the spatial (correlated) data into transformed (uncorrelated) coefficients [12]. By transformation, an image/video is changed into a format with highly reduced inter-pixel redundancies based on a reversible, linear mathematical mapping of the pixel values onto a set of coefficients, which are then quantized and encoded [13]. The key success factor of video transformation is its small magnitude as well as quantized and distortion-free outputs for most natural images [14–16].

Majority of the existing knowledge on video compression adopts single-mode approach that is based on lossless, lossy, DCT, DWT, Huffman coding, frame skipping, run length coding or any other technique. The single-mode approach has experienced some limitations which include error of parallax, decrease in quality measurement, failure to handle blocking effects and artifacts, non-applicable to 3D and web stressing applications, failure with images exhibiting full mathematical precision and coarse quantization of frequency coefficients resulting in reconstructed image with poor edge quality. Other limitations include the inability to handle distortions arising from higher compression and abnormality arising from large pixel blocks and false contouring due to poor quantization of the transform coefficients. The need for addressing the aforementioned limitations buttressed the importance of this study. The proposed technique uses discrete transformation to

achieve higher scale compression of videos. The underlying difference between the new technique and other existing techniques is the mode of operation. While existing techniques adopt single or serial mode of operation, the new technique adopts an integrative mode that is based on DCT and DWT. The objective is to establish a video compression technique that takes the storage, processing and transmission of videos to optimized levels. The new technique made four significant contributions. The first is that it established that DCT and DWT can be combined for video compression. The second contribution is that it established a platform that outperforms some other existing techniques in terms of compression ratios. The third contribution is that it established a platform for the compression of videos of varying formats with no adversity or side effect on key attributes. The fourth contribution is that it established a video compression technique that is suitable for optimizing storage requirement as well as higher speed processing and archiving. The higher compression ratio is attributed to the increase in the bit per pixel strength of the video while retaining the key attributes was achieved using the segregation and decomposition method. The reduction in the storage requirements as well as increase in the speed of processing, retrieval and processing are on the basis of the significant reduction in the bit size. The research has therefore resolved the hypothesis that non-compressed video is confronted with problems of high processing time, expensive memory requirement, slow speed query, inefficient load distribution among others. The dataset used in the study comprises twenty videos downloaded from YouTube between January and April 2020. The preference for the selected videos was based on quality, application and format multiplicity criteria which are key metrics for evaluating the performance of any video compression technique. The general results from the research include a platform for obtaining bit-optimized videos as well as bit-optimized versions of the experimental videos. Notably, the study established that the integration of DCT and DWT for video compression provides advancement over some existing ones. Section 2 presents the review of relevant literature, Sects. 3 and 4 focus on discrete cosine transform and discrete wavelet transform respectively. The discrete transform technique, case study of some selected videos and the conclusion drawn from the research are presented in Sects. 5, 6 and 7 respectively. Section 8 presents the future research focus.

2 Synopsis of some related works

In [1], a compressive sensing model is presented for accurate acquisition and reconstruction of signals, images and video sequences. The model uses minimization and orthogonal matching based on some performance parameters like compression ratio, peak signal to noise ratio (PNSR) and

structural similarity index for its reconstruction schemes. PSNR however depends on pixel intensity and human eye perception for similarity which can lead to error of parallax. The authors in [4] present a DCT and DWT model for video compression. The model effectively reduced the number of bits required for quality video presentation but with decrease in quality measurement due to lossy approach. In [17], a lossless and run length coding model for digital image compression is presented. The model supports compression of grayscale images and JPEG with data loss exerting minimal effect on image clarity. The implementation of the model demonstrated its usefulness for dividing an image into approximation and detailed sub-signal prior to compression as well as established a scheme for lossless transformation. The model is however susceptible to blocking effect and artifacts. The authors in [18] present a DCT and DWT technique for video compression. Lossy compression is adopted for achieving reconstructed video quality as well as mass screening of video images. The technique is, however, limited by its inability to handle distortions arising from higher compression and abnormality arising from large pixel blocks and false contouring due to poor quantization of the transform coefficients. A model for lossy compression of poor contrast images is presented in [19]. The model effectively uses DCT and DWT to suppress redundancies and achieve increased clarity but it is unfit for compression of images with full mathematical precision.

In [20], a model for video compression using DWT, DCT and Huffman encoding techniques is presented. The model uses the basic idea behind these techniques to reduce the average number of bits per pixel for better representation and presentation. The model is however limited by coarse quantization of its frequency coefficients resulting in reconstructed image with poor edge quality. A vector quantization and k-means clustering model for video compression are presented in [21]. The model computes the difference between two input frames using motion compensation and estimation techniques and generates its compressed image using DCT-DWT transform and arithmetic encoding followed by arithmetic decoding of residual reconstruction and compression. The model achieves encouraging compression ratio with improved video quality based on hybridization of DCT but not applicable to 3D and web stressing applications. The authors in [22] present a hybridization of DWT, DCT and encoding model for biomedical video application. The model relies on DCT to achieve less computation and higher energy compaction. It also uses DWT and Huffman coding to achieve higher compression ratios and represent a string of symbols with lesser number of bits respectively. The model is however susceptible to high blocking artifacts due to heavy compression and false contouring arising from distortion of smoothly graded areas.

A sensing model for medical image compression and reconstruction is presented in [23]. The model is anchored on non-linear programming and involves marginalization of compressive sensing and recovering of signal or images from far fewer samples than traditionally required. It also involves the selection of a continuous sub-group of coefficients with the largest energy. A model for mobile application video decoding is presented in [24]. The model uses up-down sampling, frame skipping and DCT for high scale video compression and web based streaming with scalable vector. The results from the model however demonstrate some inconsistencies. Rate distortion optimized motion estimation and quad tree model for video compression are presented in [25]. Quad tree algorithm was used to analyze the natural ecology and estimate the rate of distortion and motion optimization. Block matching algorithm was applied on the natural ecology protection system to locate video regions most prone to pollution. It is, however, noted that the model underperforms with edge detection and compression. The authors in [26] established a model that combines some simple functions with DCT to establish a 2-D DCT platform for video compression. With this platform, video compression is via 8-by-8 transformation and retransformation. Results from the model displayed some goodness but not without some noticeable differences between the original and reconstructed versions. A run-length encoding video compression model is presented in [27]. The model eliminates redundant data but with no usefulness to images with a limited number of runs. In [28], a DCT and DWT motion estimation and compensation model for video compression are presented. The model achieves sufficiently high compression ratios but with computational complexities and diminishing performances with poor quality videos. Table 1 presents the summary of the objectives, methods, advantages and disadvantages of the reviewed works.

3 Discrete cosine transform (DCT)

The DCT is based on the mathematical function, δ which defines a 1-D sequence of length N and is presented as follows:

$$\delta(l) = d(l) \sum_{i=1}^{\epsilon} \rho(i) \cos(((2i + 1)w\pi)^{-2\epsilon}), \tag{1}$$

$$\rho(i) = \sum_{w=1}^{\epsilon} \beta(l) \delta(l) \cos(((2i + 1)w\pi)^{-2\epsilon}), \tag{2}$$

$$\beta(l) = \begin{cases} \sqrt{\frac{1}{\epsilon}}, & \text{for } l = 0 \\ \sqrt{\frac{1}{\epsilon}}, & \text{for } l = 1, 2, \dots, \epsilon \end{cases}, \tag{3}$$

$\beta(l)$ is the inverse transformation. The 2-D DCT represented by δ is a direct extension of the 1-D case and is given by:

$$\delta(\varphi, l) = \beta(\varphi) \beta(l) \sum_{s=1}^{\epsilon} \sum_{i=1}^{\epsilon} \rho(s, i) \cos((2s\varphi\pi + \varphi\pi)^{-2\epsilon}) \cos((2i\varphi\pi + l\pi)^{-2\epsilon}), \tag{4}$$

$$\rho(s, i) = \sum_{\varphi=1}^{\epsilon} \sum_{l=1}^{\epsilon} \beta(\varphi) \beta(l) \delta(\varphi, l) \cos((2s\varphi\pi + \varphi\pi)^{-2\epsilon}) \cos((2i\varphi\pi + l\pi)^{-2\epsilon}), \tag{5}$$

$\varphi, w = 1, 2, \dots, \epsilon$ and $\rho(s, i)$ is the inverse transform. ϵ is the count of partitions defined in the input video ρ having i and s pixels, $\rho(s, i)$ is the intensity of the pixel in row s and column i of the image, $\delta(\varphi, i)$ gives the DCT coefficient in row φ and column i of the DCT matrix. The corresponding 2-D DCT function is generated based on the product of the horizontally oriented 1-D DCT function and the vertically oriented set of the same functions. The functions for $\epsilon = 6$ are shown in Fig. 4 which portrays a progressive increase in frequency both in the vertical and horizontal directions. As shown in Fig. 4, the number of horizontal bars increases from 2 for row 1 to 7 for row 6 while the number of vertical bars increases from 3 for column 1 to 8 for column 6.

4 Discrete wavelet transform (DWT)

Discrete wavelet transform is a multi-resolution transform technique used to achieve higher image or video compression ratio. It compounds series of transformation methods as a means of compensating for their respective weaknesses [28]. For a 2-D DWT, while the input data are passed through a set of low and high pass filters in the rows and columns directions, the outputs are down-sampled by factor 2 in each bearing in a manner related to 1-D DWT as shown in Fig. 5. The output represents a set of quadruple coefficients LL, HL, LH and HH where the first and second correspondences signify the transform in row and column respectively. The correspondent L means low pass signal while H means high pass signal. Thus, LL represents low pass signal in row direction and low pass signal in column direction while HL implies high pass signal in row direction and low pass signal in column direction.

The LH signal contains horizontal elements while HL and HH signals contain vertical and diagonal elements, respectively. In DWT reconstruction, input data is obtained with more than a few resolutions by further splitting the LL coefficient into multiple levels as shown in Fig. 6. The output is

Table 1 Summary of the objectives, methods, advantages, and disadvantages of the reviewed works

Research	Objectives	Method	Advantages	Disadvantages
Charu and Meenakshi [1]	To formulate a video compressive sensing model	Minimization Orthogonal matching Performance parameters	Accurate acquisition and reconstruction of signals, images and video sequences	Performance depends on pixel intensity and human eye perception for similarity Susceptible to error of parallax
Kumar et al. [4]	To formulate a model for video compression	DCT DWT	Effectively reduces the number of bits required for quality video presentation	Prone to reduction in quality measurement
Dhande and Gulhane [17]	To formulate model for digital image compression	Lossless coding Run length coding	Supports compression with minimal effect of data loss. Useful for approximation and detailed sub-signal	Susceptible to blocking effect and artifacts
Phakade and Harish [18]	To formulate a model for video compression	Lossy compression	Support lossless transformation	Unable to handle distortions and abnormalities
Bheemeswara et al. [19]	To formulate a model for compression of poor contrast images	Lossy compression	Achieves reconstructed video quality as well as mass screening of video images	Prone to false contouring
Singh et al. [20]	To formulate model for video compression	DWT DCT Huffman encoding	Suitable for suppressing redundancies Achieve increased clarity	Unfit for compression of images with full mathematical precision
Rami and Chitra [21]	To formulate model for video compression	Vector quantization K-means clustering	Effectively reduces the average number of bits per pixel	Limited by coarse quantization of its frequency coefficients Produced reconstructed image with poor edge quality
Bharath and Padmajadevi [22]	To formulate an encoding model for biomedical video applications	Hybridization of DWT, DCT and encoding model	Achieves encouraging compression ratio Led to improved video quality	Not applicable to 3D and web stressing applications
Bhatt and Banniya [23]	To formulate a sensing model for medical image compression and reconstruction	Linear programming	Achieve less computation Achieve higher compression ratios	Susceptible to high blocking artifacts
Ronik et al. [24]	To formulate model for mobile application video decoding	Up-down sampling Frame skipping DCT	Supports signal or image recovery	Computationally expensive
Kumar et al. [25]	Formulate model for video compression	Rate distortion Optimized motion estimation Quad tree	Achieves high scale video compression and web based streaming with scalable vector	The results demonstrate some inconsistencies
Gupta and Garg [26]	Formulate model for video compression	Simple functions DCT	Effectively locate video regions most prone to pollution	Underperforms with edge detection and compression
Dai et al. [27]	Formulate a video compression model	Run-length encoding	It establishes a 2-D DCT platform for video compression	It exhibits some noticeable differences between the original and reconstructed versions
Shanthi and Somasundaram [28]	Formulate a video compression model	DCT DWT motion estimation Compensation model	Eliminates redundant data	Not useful for images with limited number of runs
			Achieves sufficiently high compression ratios	Experiences computational complexities and diminishing performances with poor quality videos

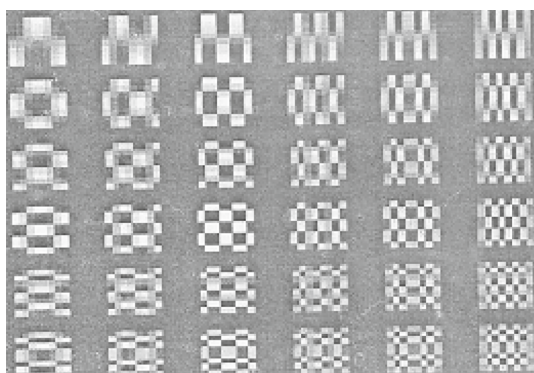


Fig. 4 Two dimensional DCT basic functions with $\epsilon = 6$

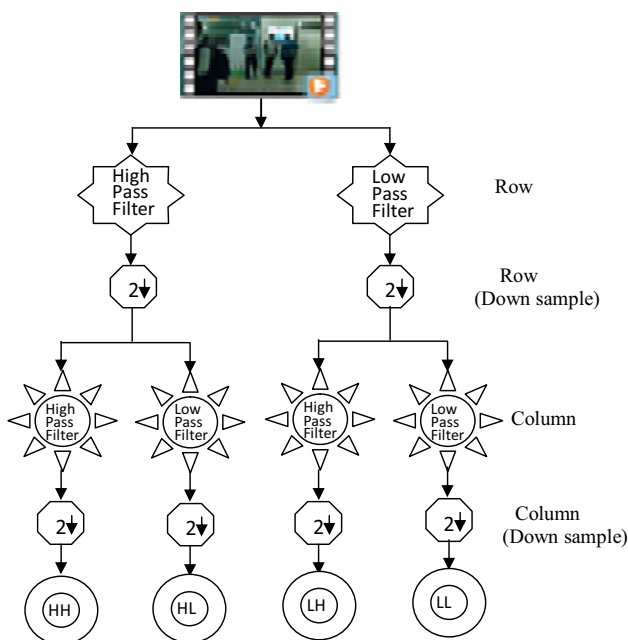


Fig. 5 Block diagram of 2-D forward DWT row

derived based on factor 2 up-sampling compression using high pass and low pass filters in both row and column.

The poly-phase network of the wavelet filter is factorized into an order of substituting upper and lower triangular grids and a corner to corner lattice, leading to joined grid duplications. The DWT scheme comprises of some processing set-up, integer-to-integer number wavelet transform (IWT), symmetric forward and converse transform among others. It is used to separate the high pass and low pass filters into a banding of upper and lower triangular grids and deliver the filter execution into united network duplications. A wavelet transform with separate taps in the high and low pass investigation filters which represent the shortest and least complex symmetrical bi-orthogonal wavelet is used to break an image into high and low-frequency segments. The mathematical

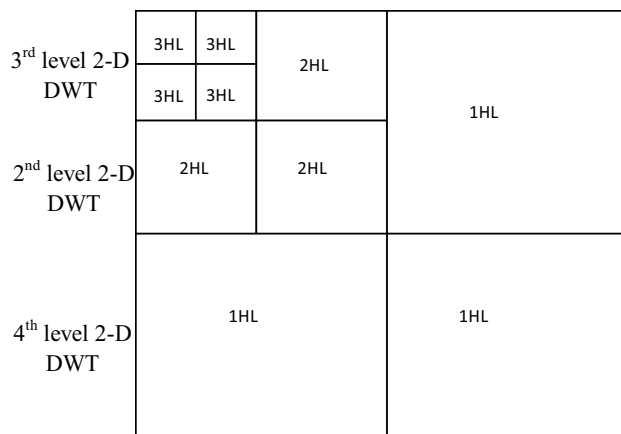


Fig. 6 Splitting of LL coefficient

formula for the fundamental lifting schemes low pass (*l*) and high pass (*h*) comparisons are presented as follow:

$$l = (2i + 1)(\delta - 2\rho), \tag{6}$$

$$h = 2i(3\delta - \rho), \tag{7}$$

i, δ and ρ represent the pixel size, DCT coefficient and intensity respectively.

5 Discrete transform technique (DTT) for video compression

The methods proposed in the literature for video compression are noted for problems of error of parallax [1], reduction in quality measurement [4], blocking effect and artifacts [17, 22], failure to handle distortions, abnormalities and false contouring [18], coarse quantization of frequency coefficients [20], failure to handle 3D and web stressing applications [21] and computationally expensiveness [23, 28]. The need to address these problems is partly responsible for the formulation of the proposed DTT. The adoption of the new technique is also based on the need to provide answers to some questions bothering on the impact of non-optimization of bits on video management vis-a-vis memory requirement, processing time, query servicing, load distribution among others. As conceptualized in Fig. 7, DTT is a combination of the DCT and DWT techniques presented in Sect. 3 and Sect. 4 respectively. DCT is a fast orthogonal technique with usefulness for the transformation of matrices with a large number of (near) zero values. Its major limitation is the compelling quantization step required for taking decisions on integer-valued output. DWT on its own does not offer specific information on the output but provides a better way to detect any anomaly in the transformation. The

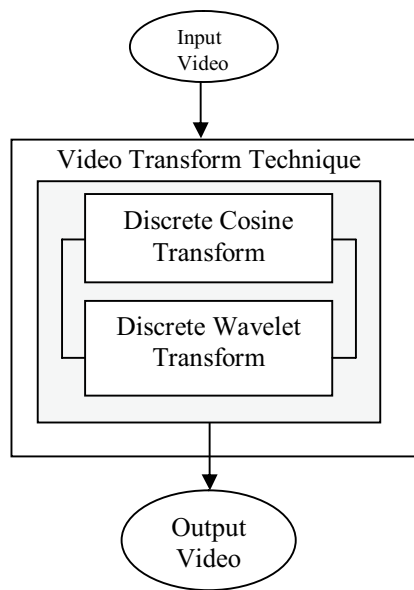


Fig. 7 Digital video compression techniques

motive behind the combination of DCT and DWT is to cancel out the weaknesses of each technique using the strength of the other.

The new technique performs image segregation using 2-D frame of the model presented in Fig. 6 as well as the decomposition of each segment using a 2-D DWT based on quantization and elimination of zeros and compression of each sub-band by Arithmetic Coding as illustrated in Fig. 8. The quantization process takes the LL coefficients as input to form a [0,1] element matrix which is passed to the later stage for discarding of the zeros. The zero-free matrix is then passed to the final stage for arithmetic coding. The new technique also performs the compression of the sub-bands (HL, LH, HH) using T – Matrix Coding based on the order shown in Fig. 9. The discrete transform technique uses 2-D DWT to decompose each component to obtain coefficients which are further subjected to 8-point DCT. For improved compression, the preponderance of the high coefficients is subsequently discarded based on a JPEG-like quantization in which the higher frequency components are subjected to

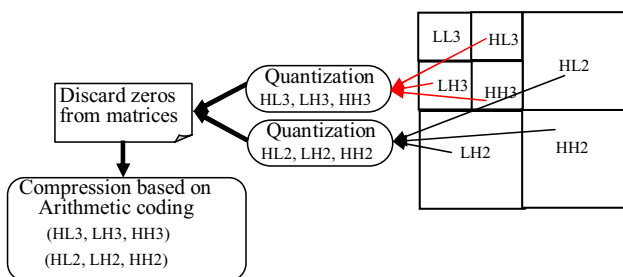


Fig. 8 Quantization and elimination of zeros form each sub-band

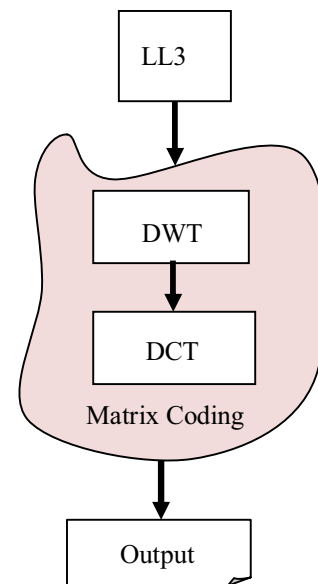


Fig. 9 Sub-bands compression matrix coding

scaling. A 1-D DCT quantization is also used to obtain the array of each sub-band. The combination of DWT technique with the DCT technique is to achieve higher video compression ratio as the former reduces false contouring and artifacts effects while the later reduces redundancies using frequency-based classifications.

6 Case study

A case study of some randomly selected online videos (shown in Fig. 10) which were downloaded from YouTube between January and April 2020 was carried out on Pentium 3 CPU with 256 MB RAM and 128 MB HDD and runs Windows XP Operating System. Visual Basic.NET was used to implement the programming logic while Access database system provided the auxiliary storage unit for the information on the videos and system. The experimental videos shown in Fig. 10 exhibit varying sizes, formats and attributes. Table 2 presents the pre- and post-compression attributes for all the experimental videos while Fig. 11 presents a section of the post-compression report on video link, pre- and post-compression sizes as well as the date of the operation. As shown in Table 1, the accumulated bytes of spaces released after the compression operations on all the twenty videos is 1.11196 GB (4326.76–3214.80) which is 25.70% of the total pre-compression sizes. The compression operation on all the videos also led to an average bit per pixel of 98 and average compression time of 58.8 s. These results indicate good performances of the discrete transform technique and established its ability to successfully compress images of

Fig. 10 Experimental videos



Table 2 Pre- and post-compression attributes of the experimental videos

Video	Dimension	Format	Pre-compression Size (MB)	Post-compression size (MB)	Compression ratio	Space saving (100%)	Bits per pixel (BPP)	Compression time (seconds)
VID-01	1280 × 720	MP4	161.00	108.00	1.49	32.92	083	64.8
VID-02	1280 × 720	MP4	170.63	134.05	1.27	21.44	120	60.2
VID-03	320 × 174	3GP	252.10	162.86	1.55	35.4	081	66.8
VID-04	1280 × 720	3GP	155.57	097.53	1.60	37.31	123	63.3
VID-05	1280 × 720	MPEG	164.48	128.89	1.28	21.64	121	64.2
VID-06	1280 × 720	MP4	167.86	127.66	1.31	23.95	118	60.5
VID-07	1280 × 720	MPEG	173.89	119.53	1.45	31.26	115	66.7
VID-08	1280 × 720	MP4	252.10	162.86	1.55	35.4	080	65.8
VID-09	1280 × 720	MP4	211.44	148.23	1.43	29.9	081	65.2
VID-10	1280 × 720	MP4	167.86	107.66	1.56	35.86	121	64.2
VID-11	960 × 720	MPEG	167.86	107.66	1.56	35.86	121	64.2
VID-12	960 × 720	MP4	181.45	124.10	1.46	31.61	086	65.0
VID-13	960 × 720	MP4	129.62	084.43	1.54	34.86	081	67.5
VID-14	1280 × 720	MPEG	702.39	642.24	1.09	8.56	120	77.6
VID-15	960 × 720	MPEG	013.59	003.43	3.96	74.76	029	01.7
VID-16	1280 × 720	FLV	300.71	231.88	1.30	22.89	121	63.3
VID-17	1280 × 720	JPEG	513.52	430.83	1.19	16.10	110	70.5
VID-18	960 × 720	MP4	129.62	084.43	1.54	34.86	081	64.2
VID-19	1280 × 720	MPEG	181.45	124.10	1.46	31.61	086	67.7
VID-20	320 × 240	JPEG	129.62	084.43	1.54	34.86	081	69.2
Total/average			4326.76	3214.80	1.56	31.55	98	58.8

all types and formats. It is equally proved that the combination of DCT and DWT for video compression promotes storage conservation and management efficiency. The

compressions were based on low pass (h) and high pass (g) filter coefficients of the conventional 5/3 filter given as:

Fig. 11 The post-compression report

Video	Initial Size	New Size	Date
C:\Users\HP PC\Desktop\compress\VD001.flv	161.00 MB	108.00 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD002.flv	170.63 MB	134.05 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD003.flv	252.10 MB	162.86 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD004.flv	155.57 MB	097.53 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD005.flv	164.48 MB	128.89 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD006.flv	167.86 MB	127.66 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD007.flv	173.89 MB	119.53 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD008.flv	252.10 MB	162.86 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD009.flv	211.44 MB	148.23 MB	9/24/2020
C:\Users\HP PC\Desktop\compress\VD010.flv	167.86 MB	107.66 MB	9/24/2020

$$h : \{-0.125, 0.25, 0.75, 0.25, -0.125\}$$

$$g : \{-0.25, 1, -0.25\}$$

Based on this co-efficient, the filter outputs, f_1, f_2 and f_3 were computed based on the inverse transform, x and 1-D sequence of length n as follows:

$$f_1 = -0.25x(n - 1) + 0.5x(n - 2) - 0.25x(n - 3), \tag{8}$$

$$f_2 = 0.125x(n) + 0.25x(n - 1) + 0.75x(n - 2) + 0.25x(n - 3) - 0.125x(n - 4), \tag{9}$$

$$f_3 = 0.125[-x(n) + 2x(n - 1) + 6x(n - 2) + 2x(n - 3) - x(n - 4)] \tag{10}$$

Results obtained for some standard metrics such as compression ratio (∂), bits per pixel (γ) and relative redundancy (ϵ) served as the basis of the justification for the good performance claim. ∂ and γ are obtained from:

$$\partial = \frac{S_o}{S_c}, \tag{11}$$

$$\gamma = \frac{1}{8\partial}, \tag{12}$$

S_o and S_c are the pre- and post-compression bit sizes, respectively. The space conservation index, μ is obtained from

$$\mu = \left(1 - \frac{S_c}{S_o}\right) \times 100. \tag{13}$$

The frequently used objective measurements based on numerical and statistical terms for performance evaluation are peak signal noise ratio (ϑ), mean square error (\emptyset) and maximum difference (τ). These measurements are defined as follows:

$$\vartheta = 10 \log\left(\frac{m^2}{\emptyset}\right), \tag{14}$$

$$\emptyset = \frac{1}{ab} \sum_{x=1}^a \sum_{y=1}^b (O_{(x,y)} - C_{(x,y)})^2 \tag{15}$$

$$\tau = \max \{|O_{(x,y)} - C_{(x,y)}|\}, \tag{16}$$

m is the bit value, a and b are the image dimensions while $O_{(x,y)}$ and $C_{(x,y)}$ are the original and compressed images, respectively. ϑ is based on \emptyset and the higher its value, the higher the level of compression. \emptyset represents the cumulative squared error between the compressed and the original image and the lower its value, the lower the error. Similarly, the higher the value of τ , the higher the compression. The plots of the obtained pre and post-compression ϑ , \emptyset and τ values are presented in Figs. 12, 13 and 14. Visual inspection of the plots reveals higher post-compression peak to noise signal ratios and maximum differences as well as lower

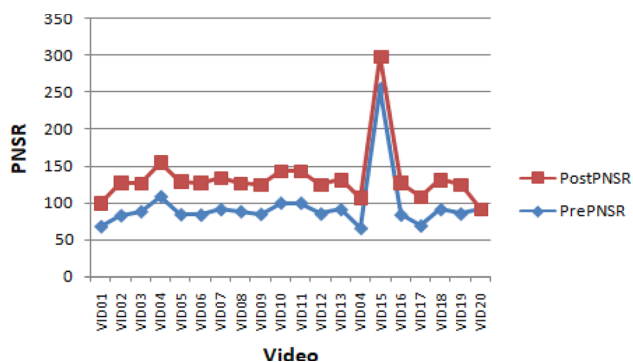


Fig. 12 Plot of pre and post compression PSNR

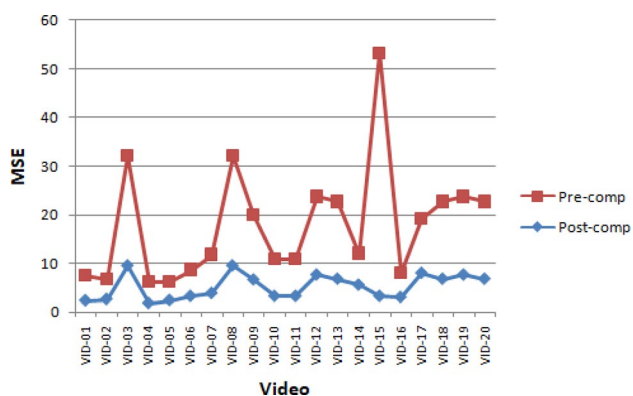


Fig. 13 Plot of pre and post compression MSE

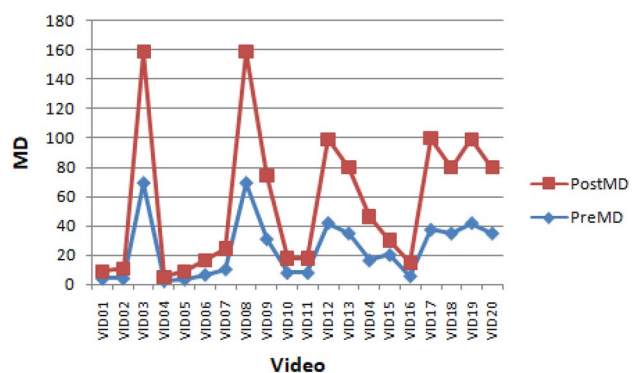


Fig. 14 Plot of pre and post compression MD

mean square errors in all corresponding cases. These figures conveyed the integration of DCT and DWT as suitable for achieving size reduction, higher bit density and improved quality of display for the videos.

A comparative analysis based on the experimental study of ten other techniques; namely Compressive Sensing [1], Convolutional Neural Network [7], Hybrid DCT-DWT [18],

Residual Vector Quantization [21], DCT [25], Differential Detection and Run-Length Encoding [27], Huffman [29, 30], Lossless JPEG-Huffman [31], Run Length [32, 33] and Motion Compensation [34] techniques was carried out. The choice of these techniques is based on their recentness, popularity and applications. The comparative analysis adopted the same environment, engines and dataset used for the new technique and its purpose is to provide the basis for comparing the new technique with existing ones. The compression ratios obtained from the experimental study of the ten techniques on the twenty research videos are presented in Table 3 along with the ones obtained for the new technique. Visual inspection of the figures presented in Table 3 reveals that higher figures are recorded for the current (new) technique in most cases. This presents the fact that the new technique stands the ground of preference over the ten other techniques when optimal bit compression of videos of all types and formats is required. Few cases where the older algorithms outperformed the new algorithm are highlighted.

7 Conclusion

Videos are noted for size explosiveness which largely accounts for their large download/upload and storage complexities as well as lengthened times for transmission and processing. Compression operation on videos is therefore often required for bit reduction to more economic friendly storage, transmission and playback. Previous works had adopted single-mode approaches that are based on lossless, lossy, DCT, DWT, Huffman coding, frame skipping, run length coding or any other technique for video compression with attendant limitations which include error of parallax, decrease in quality measurement, failure to handle blocking effect and artifacts, non-applicable to three dimensional (3D) and web stressing applications, failure with image containing full mathematical precision and coarse quantization of frequency coefficients resulting in reconstructed image with poor edge quality. Other limitations of some existing techniques include inability to handle distortions arising from higher compression and abnormality due to large pixel blocks and false contouring emanating from poor quantization of the transform coefficients. Specifically, the research is based on the hypothesis that non-bit-optimized videos are faced with the challenges of high processing time, expensive memory requirement, slow speed query, inefficient load distribution among others. The experimental study of the DTT for resolving the hypothesis was based on some online videos with varying sizes, sources and formats. The figures recorded from the experimental study using some standard metrics such as maximum difference, mean square error and peak signal to noise ratios for the original and compressed videos formed the basis of performance analysis. The

Table 3 Compression ratios for different video compression algorithms

Video	Size (MB)	Algorithm										
		Ref [1]	Ref [7]	Ref [18]	Ref [21]	Ref [25]	Ref [27]	Refs [29, 30]	Ref [31]	Refs [32, 33]	Ref [34]	Current
VID-01	161.00	1.03	1.39	1.46	1.17	1.44	1.39	0.94	1.26	1.42	1.06	1.49
VID-02	170.63	1.07	1.32	1.17	1.20	1.20	1.25	0.97	1.29	1.45	1.09	1.27
VID-03	252.10	1.13	1.29	1.44	1.28	1.51	1.52	1.03	1.17	1.31	1.16	1.55
VID-04	155.57	1.17	1.34	1.51	1.32	1.56	1.49	1.06	1.22	1.37	1.20	1.60
VID-05	164.48	1.16	1.33	1.20	1.30	1.21	1.26	1.05	1.21	1.36	1.18	1.28
VID-06	167.86	1.05	1.53	1.23	1.18	1.34	1.30	0.95	1.39	1.57	1.07	1.31
VID-07	173.89	1.24	1.61	1.40	1.40	1.50	1.41	1.13	1.46	1.64	1.27	1.45
VID-08	252.10	1.08	1.74	1.45	1.22	1.46	1.55	0.98	1.58	1.77	1.11	1.55
VID-09	211.44	1.30	1.25	1.41	1.46	1.47	1.41	1.18	1.14	1.28	1.33	1.43
VID-10	167.86	1.03	1.22	1.38	1.16	1.42	1.41	0.94	1.11	1.25	1.05	1.56
VID-11	167.86	0.99	1.25	1.41	1.12	1.42	1.45	0.90	1.14	1.28	1.02	1.56
VID-12	181.45	1.05	1.31	1.47	1.18	1.41	1.23	0.95	1.19	1.34	1.07	1.46
VID-13	129.62	1.16	1.42	1.40	1.30	1.59	1.51	1.05	1.29	1.45	1.18	1.54
VID-14	702.39	1.12	1.40	1.07	1.27	1.05	1.01	1.02	1.27	1.43	1.15	1.09
VID-15	013.59	2.25	2.22	2.11	21.12	2.26	2.25	1.90	1.11	1.25	1.02	3.96
VID-16	300.71	1.09	1.29	1.24	1.23	1.23	1.19	0.99	1.17	1.31	1.12	1.30
VID-17	513.52	1.12	1.29	1.14	1.25	1.21	1.11	1.02	1.17	1.31	1.14	1.19
VID-18	129.62	1.07	1.24	1.40	1.20	1.49	1.49	0.97	1.13	1.27	1.09	1.54
VID-19	181.45	1.01	1.25	1.42	1.14	1.41	1.42	0.92	1.14	1.29	1.04	1.46
VID-20	129.62	0.99	1.21	1.36	1.12	1.49	1.50	0.90	1.10	1.24	1.02	1.54

analysis showed that compression based on the integration of DCT and DWT achieved significant reduction in bit size for all the videos. It was also revealed that the new technique is applicable to 3D and web stressing applications as well as suitable for full mathematical precision and coarse quantization of frequency coefficients. Specifically, the post-compression bit figures recorded for the research led to 25.70% reduction in the aggregate storage requirements for all the experimental videos. Judging from the superior compression ratios presented in Table 3, this cumulative reduction figure will unarguably exceed equivalent figures that may be derived using the algorithms proposed in [1, 7, 18, 21, 25, 27] and [29–34]. Ultimately, the reduction in the storage requirements will leads to lower processing and querying times as well as more efficient load distribution.

8 Future works

Thirty-three (33) of the two hundred (200) cases of the compression ratios reported and highlighted in Table 3 for the ten older but related algorithms exceed the values reported for the new technique. This implies a 16.5% below performance figure which is blamed on some complexities that the new technique could not handle for the specific images. Further research, therefore, focuses on strengthening the new

technique with the addition of other compression algorithms such as Huffman coding, frame skipping and run length coding for attaining near-zero below performance figure when compared with other existing and recent techniques. Consideration will also be given to the compression of higher dimension images.

Declarations

Conflict of interest As far as the content of this article is concerned, the authors have no conflicts of interest to declare.

References

1. Charu, B.U., Meenakshi, S.: Performance evaluation of compression for images and videos using compressive sensing technique. In: Proceedings of the 4th International Conference on Computing for Sustainable Global Development, BVICAM, New Delhi, pp. 6253–6257 (2017).
2. Shriram, P.H.: Algorithms and Architecture for Discrete Wavelet Transform Based Video encoder. PhD thesis submitted to Department of Electrical and Electronics Engineering, Vinaya Missions University (2015)
3. Anuja, R., Sridevi, S., Vijayakumar, V.R.: A survey on various compression methods or medical images. *Int. J. Intell. Syst. Appl.* **3**, 13–19 (2012)

4. Kumar, P.R., Murty, P.S., Babu, P.N.: The novel lossless text compression technique using ambigram logic and Huffman coding. *Inf. Knowl. Manag.* **2**(2) (2016)
5. Balle, J., Laparra, V., Simoncelli, E.P.: End-to-end optimization of nonlinear transform codes for perceptual quality. In: *IEE Picture Coding Symposium*. <https://arxiv.org/abs/1607.05006>. Accessed 25 Sep 2017 (2016).
6. Dosovitskiy, A., Brox, T.D.: Generating Images with Perceptual Similarity Metrics Based on Deep Networks. <https://arXiv.org/abs/1602.02644>. Accessed 25 Sep 2017 (2016)
7. Gatys, L.A., Ecker, A.S., Bethge, M. Image style transfer using convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. www.semanticscholar.org/paper/image-style-transfer-Using-Convolutional-Neural-Network-Gatys-Ecker/7568d13a82f7afa4be79f09c295940e48ec6db89, Accessed 12 Nov 2018 (2016)
8. Hall, K., Gregor, I., Danihelka, A.G., Wierstra, D.: Towards Conceptual Compression. <https://arXiv.org/abs/1601.06759>. Accessed 14 Sep 2018 (2016)
9. Zhao, C., Siwei, M., Zhang, J., Xiong, R., Gao, W.: Video compressive sensing reconstruction via reweighted residual sparsity. In: *IEEE Transaction on Circuits and Systems for Video Technology*, pp. 1–14 (2015)
10. Zhang, F., Duan, Y.: A highly effective impulse noise detection algorithm for switching median filters. *IEEE Signal Process. Lett.* **17**(7), 647–650 (2015)
11. Gao, C., Cao, Y., Zhou, Z., Sun, X.: Coverless information hiding based on the molecular structure images of material. *Comput. Mater. Contin.* **54**(2), 197–207 (2015)
12. Devore, J.L.: *Probability and Statistics*. Thomson Asia Pte Ltd, Singapore (2002)
13. Hong, S.W., Bao, P.: Hybrid image compression model based on sub band coding and edge preserving regularization. *IEEE Proc. Vis. Image Signal Process.* **147**(1), 16–22 (2005)
14. Ying, L., Dimitris, A.P.: Compressed sensed domain L1 PCA video surveillance. *IEEE Trans. Multimed.* **18**(3), 351–363 (2016)
15. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 3rd edn. Pearson Prentice Hall (2005)
16. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn. Prentice Hall (1992)
17. Dhande, J.D., Gulhane, S.M.: Daubechies wavelet based neural network classification system for biomedical signal. *Int. Conf. Inf. Process. Pune* **99**, 188–191 (2015)
18. Phakade, S.V., Harish, P.: Video compression using hybrid DCT-DWT algorithm. *Int. Res. J. Eng. Technol.* **3**, 2758–2761 (2016)
19. Bheemeswara, G.R., Kavitha, T.S., Yedukondalu, U.: FPGA Implementation of JPEG2000 image compression using modified DA based DWT and lifting DWT-IDWT technique. *Int. J. Eng. Res. Appl.* **2** (2012)
20. Singh, L., Rajeshwar, D., Sandeep, K.: Video compression techniques. *Int. J. Sci. Technol. Res.* **1**(10), 114–119 (2012)
21. Rani, S., Chitra, P.: An improved fractal image compression using residual vector quantization for compressing medical images. *Proc. Int. Conf. Front. Eng. Appl. Sci. Technol.* **3**, 86–90 (2017)
22. Bharath, K.N., Padmajadevi, G.K.: Hybrid compression using DWT-DCT and Huffman encoding techniques for biomedical image and video applications. *Int. J. Comput. Sci. Mobile Comput.* **2**(5), 255–261 (2013)
23. Bhatt, M.U., Bamniya, K.: *Medical Image Compression and Reconstruction Using Compressive Sensing*. Academic Press (2015)
24. Ronik, D., Poorva, W., Sangeeta, J.: Web based modified video decoding for mobile application. *Int. J. Comput. Appl.* **126**(15), 19–23 (2015)
25. Kumar, S.S., Mohan, B.C., Chatterji, B.N.: Video compression system for online usage using DCT. *Int. J. Trend Res. Dev.* **2**(4), 67–72 (2019)
26. Gupta, M., Garg, K.: Analysis of image compression algorithm using DCT. *Int. J. Eng. Res. Appl. (IJERA)* **2**(1), 515–521 (2016)
27. Dai, B., Yin, S., Gao, Z., Wang, K., Zhang, D., Zhuang, S., Wang, X.: Data compression for time-stretch imaging based on differential detection and run-length encoding. *J. Light Wave Technol.* **35**(23), 5098–5104 (2017)
28. Shanthi, R.M., Somasundaram, K.: Mode based K-means algorithm with residual vector quantization for compressing images. In: *International Conference on Control Computation and Information System*, pp. 105–112 (2017)
29. Huffman, D.: A method for the construction of minimum-redundancy codes. *Proc. IRE* **40**(10), 1098–1101 (1952)
30. Shikhar, M.: Greedy Algorithms Set 3 Huffman Coding. <http://www.geeksforgeeks.org/greedy-algorithms-set-3-huffman-coding/>. Accessed 15 Dec 2016 (2017)
31. Pennebaker, W.B., Mitchell, J.L.: *JPEG Still Image Data Compression Standard*, 3rd edn. Springer (2003)
32. Hussein, A.H., Mahmud, S.S., Mohammed, R.J.: Image compression using proposed enhanced run length encoding algorithm, *IBN AL-Haitham. J. Pure Appl. Sci.* **24**(1) (2011)
33. Bandyopadhyay, S.K., Paul, T.U., Raychoudhury, A. Image compression using approximate matching and run length. *Int. J. Adv. Comput. Sci. Appl.* **2**(6) (2011)
34. Zaid, H., Maher, K.: Video compression based on motion compensation and contourlet transform. In: *Proceedings of Third Scientific Conference of Electrical Engineering (SCEE)*. IEEE (2018)

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