ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES Knowledge to Innovation

Special Issue (April)

Online - 2455-3891
Print - 0974-2441

Full Proceeding Paper

A REVIEW ON IMAGE AND VIDEO COMPRESSION STANDARDS

MADHAVEE LATHA P*, ANNIS FATHIMA A

School of Electronics Engineering, VIT University, Chennai, Tamil Nadu, India. Email: madhaveelatha.p2016@vitstudent.ac.in

Received: 23 January 2017, Revised and Accepted: 03 March 2017

ABSTRACT

Nowadays, the number of photos taken each day is growing exponentially on phones and the number of photos uploading on Internet is also increasing rapidly. This explosion of photos in Internet and personal devices such as phones posed a challenge to the effective storage and transmission. Multimedia files are the files having text, images, audio, video, and animations, which are large and require lots of hard disk space. Hence, these files take more time to move from one place to another place over the Internet. Image compression is an effective way to reduce the storage space and speedup the transmission. Data compression is used everywhere on the internet, that is, the videos, the images, and the music in online. Even though many different image compression schemes exist, current needs and applications require fast compression algorithms which produce acceptable quality images or video with minimum size. In this paper, image and video compression standards are discussed.

Keywords: Image compression, Discrete cosine transform, Discrete wavelet transform, Video compression.

 DOI: http://dx.doi.org/10.22159/ajpcr.2017.v10s1.19760

INTRODUCTION

Compression is a process in which the required bits to store an image is reduced. Compression allows to store images in less storage space and to transmit through limited bandwidth channels. Compression is two types: Lossy and lossless.

In lossy compression, the reconstructed image contains degradations. With this perfect reconstruction of the image is sacrificed with respect to the original image, but higher compression ratio is achieved. This is used commonly in multimedia applications.

In lossless compression, the reconstructed image is an exact replica of the original image after compression. Only the modest compression rate can be achieved and used in medical applications. The categorization of compression is shown in Fig. 1.

IMAGE COMPRESSION STANDARDS

The general block diagram of Image Compression is shown in Fig. 2.

In Image compression, initially, the correlation between the pixels is reduced if the values of one pixel and its adjacent pixels are very similar. Various techniques such as predictive coding, orthogonal transform, and sub-band coding are used to reduce the correlation between pixels. Quantization is a lossy operation, which helps in reducing the number of bits required to represent the pixels. In entropy coding, the symbols are represented in variable length codes according to their probability.

JPEG baseline

The JPEG baseline given by Wallace [1] is the first international compression standard for continuous-tone still images. Discrete cosine transform (DCT) is used as transform coding for image compression. In JPEG standards, different modes such as sequential, progressive, and hierarchical modes are available. Here, the DCT is applied over fixed blocks of size 8×8.

A set of 64 DCT coefficients is produced. Quantization is applied on these coefficients, which are then zigzag scanned.

Huffman coding and arithmetic coding are used in JPEG. Huffman coding requires a predefined table for compression, whereas arithmetic coding requires no table. Arithmetic coding produces 5-10% better compression than Huffman coding.

JPEG 2000

JPEG2000 proposed by Christopoulos *et al*. [2] is an image compression standard compatible with multimedia technologies with the features such as superior bit rate performance, lossy and lossless compression, progressive transmission by pixel and by accuracy, coding of interested region, and not error prone to bit errors.

The first step of encoding process of the JPEG2000 is pre-processing which includes tiling, offset leveling, and irreversible color transform. Dividing the image as rectangular and non-overlapping blocks is called tiling. This reduces the quality and small tiles create more artifacts. The level offset pre-processing stage ensures the normal dynamic range centered about zero. No offset is added if the data are already signed.

The second step is the application of discrete wavelet transform (DWT) on all the blocks of the image. This decomposes tiles or blocks into different subbands of different resolution levels. The subbands of coefficients represent frequency characteristics of the tiles. Then, these subbands are scalar quantized. The quantization is used to reduce the precision of sub-band coefficients so that fewer bits are required to encode the transformed coefficients.

Embedded block coding with optimized truncation (EBCOT) is used as the encoding algorithm in JPEG2000. The subband is divided into groups of samples called code blocks. For each code block, the bitstreams are generated by EBCOT. Arithmetic coder entropy codes the bit planes of these codeblocks in three passes. The first pass is called significant propagation pass, in which bits are coded only if their locations are not significant but they have at least one significant neighbors in 8-connected neighbors. The second pass is called magnitude refinement pass, in which the bits not coded in the first pass and became significant in a previous bit plane are coded. The last pass is cleanup pass, where all the bits that are not coded in the first and second passes are coded. To code, the zeros efficiently run-length coding is performed. A code unit (CU) is formed in each coding pass called chunk. Here, arithmetic coder uses contextual information to generate a bitstream. For error, resilience markers are added in the bitstream.

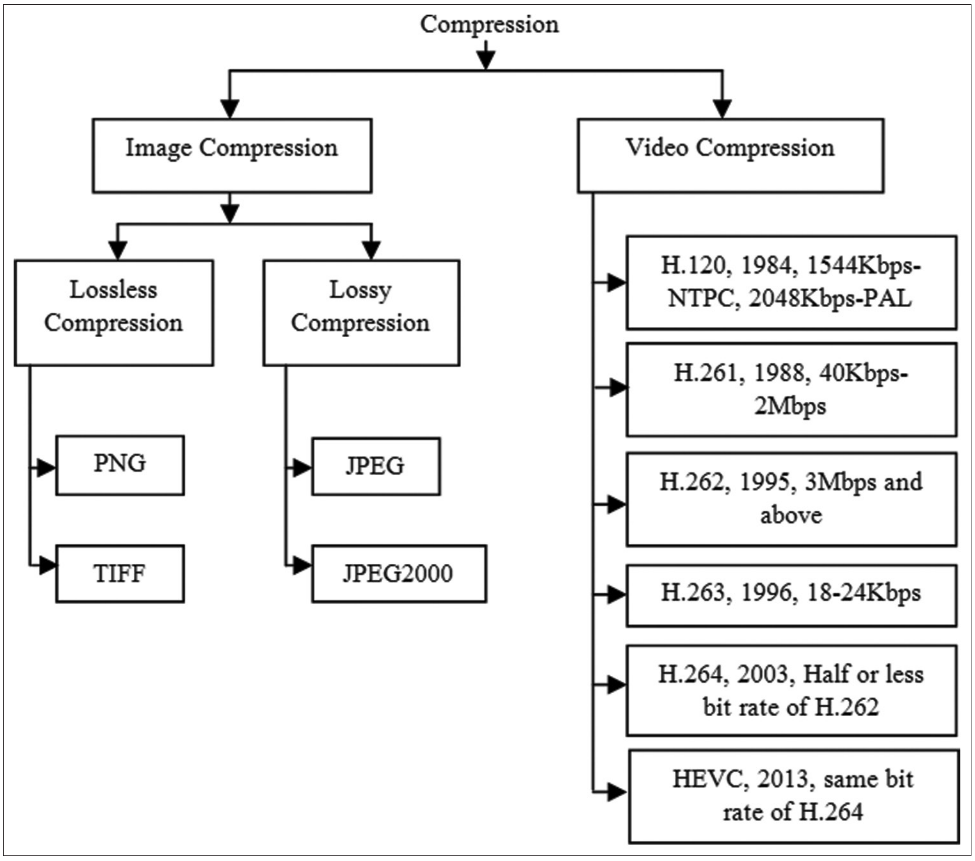

Fig. 1: Compression standards

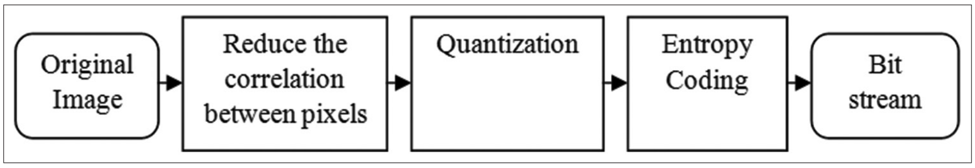

Fig. 2: The general encoding flow of image compression

Tagged image file format (TIFF)

The TIFF [3] is an image file format which is flexible in saving eight bits per color for 24 bit and 16 bits per color for 48 bit, respectively. It is of two types: Lossy and lossless. For lossless storage, it uses Lempel–Ziv–Welch technique as a compression algorithm. This format is used widely in printing business as photograph file standard and not used in browsers.

Portable network graphics (PNG)

PNG [4] is created as an open-source alternative to GIF. This file format supports 8-bit palette images and 24-bit true color or 48-bit true color. PNG exceeds JPEG when the image has large uniform colored areas. PNG supports grayscale, true color, indexed, grayscale and alpha, and true color and alpha images.

For example, consider the size of the true color image, grayscale image, and scanned document as 696 KB, 257 KB, and 1.1 MB, respectively. The compression ratios discussed by Aguilera [5] that is achieved for above images using different types of compression techniques are tabulated in Table 1.

For example, consider the size of a true color image as 768 KB. JPEG provides a well-compressed image with quality 50 of size 49.5 KB and it gives blocking artifacts for quality <50. Whereas in JPEG2000, the rectangular regions are smoothened for low-quality images, that is, the image with quality 1 of size 6.3 KB.

Table 1: Compression ratios of TIFF, PNG, and JPEG for different images

Type of image	Compression ratio achieved using lossless image format		
	TIFF	PNG	JPEG
True color image	2:1	2.7:1	16:1
Grayscale image	1:1	1.5:1	3.2:1
Scanned document	21.5:1	11.2	No compression

TIFF: Tagged image file format, PNG: Portable network graphics

VIDEO COMPRESSION STANDARDS

Video compression is the combination of spatial image compression and temporal motion compensation. Most of the video compression algorithms use lossy compression. Spatial image compression can be performed using transform coding and the temporal redundancies are removed using the block-based motion estimation and compensation techniques. The general video processing block diagram is shown in Fig. 3.

H.120

The first international digital video compression or coding standard is H.120 [6]. The techniques used in this standard are a switch for sampling, scalar quantization, differential pulse-code modulation

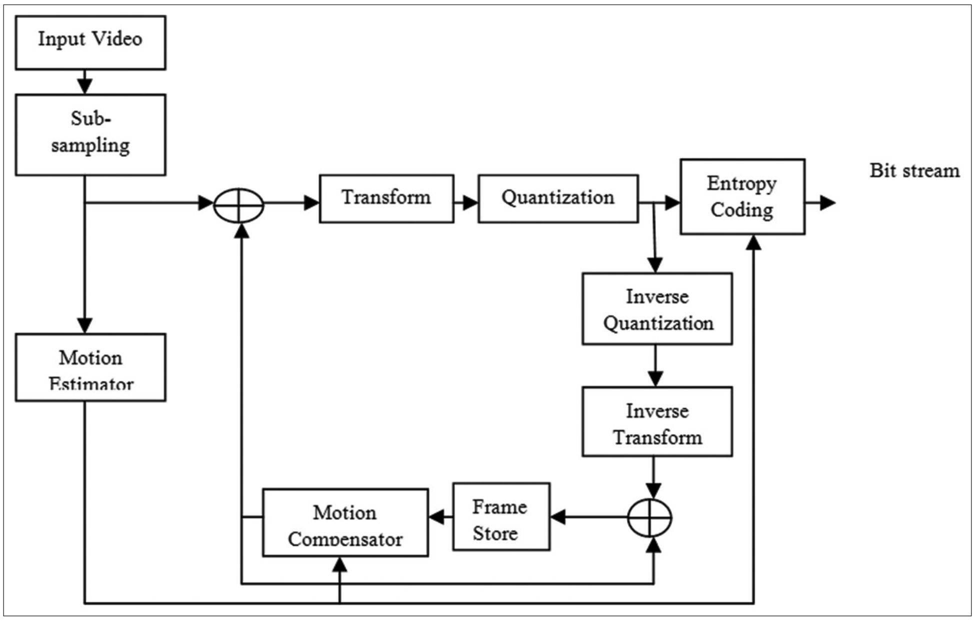

Fig. 3: Block diagram of general video coding scheme

(DPCM), and variable-length coding. Background prediction and motion compensation are added in version 2 of H.120. The video quality of H.120 is not good enough and this lead to block-based codecs.

H.261

H.261 [7] is a practical ITU-T video compression standard, which is developed for transmitting over integrated services digital network lines of data rates 64 Kbits/s. This operates at video bit rates between 40 Kbits/s and 2 Mbits/s. It supports the color space YC_bC_r with 4:2:0 chrominance subsampling. This coding standard is upgraded to video conference and video telephony applications. In this standard, a hybrid motion compensation is used for inter-picture prediction. The video is the first spatial transform coded, on which scalar quantization is applied. These quantized coefficients are zigzag scanned and then entropy coding is performed.

H.263

The video compression standard designed for a low bit-rate compression for videoconferencing is H.263 [8]. Compensation through variable block size and overlapped block motion compensations are the main features of H.263. This achieves better video at 18-24Kbps. H.263 supports the picture sizes 128×96 (Sub-QCIF), 176×144 (QCIF), 352×288 (CIF), 704×576 (4CIF), and 1408×1152 (16CIF).

H.263+

It is approved by ITU-T in 1998, and this offers a high degree of error resilience for wireless or packet-based transport networks.

H.264

Wiegand *et al.* [9] proposed H.264, which is designed to use in applications such as broadcast, wireless and mobile networks, video-on-demand or multimedia streaming services, and multimedia messaging services.

The features of this standard which enhances the efficiency of video coding are:

1. Variable block-size motion compensation
2. The picture boundary extrapolation technique
3. Multiple reference pictures for motion compensation purposes
4. This standard used "direct" motion compensation for B-plane coded areas
5. The deblocking filter is placed inside the motion compensated prediction loop
6. Representation of signals in a more locally adaptive fashion using 4×4 transforms
7. Exact-match inverse transform
8. Context-adaptive binary arithmetic coding (CABAC) is introduced as an entropy coding technique, which is a very good method created using arithmetic coding.

The coded video data are organized in packets that contain an integer number of bytes, which are called network abstraction layer (NAL) units. These units are divided into video coding layer (VCL) and non-video coding layer (non-VCL) NAL units. The sample values of the video pictures are represented in VCL NAL units and the information such as parameter set and enhancement information are included in non-VCL NAL units.

H.264/AVC uses a sampling structure 4:2:0 sampling with 8 bits of precision per sample. In this, a separable integer transform with similar properties as a 4×4 DCT is applied to 4×4 blocks. A quantization parameter is used for the quantization of transform coefficients that take 52 values. These quantized coefficients of each block are zigzag scanned and then entropy coded.

H.264 supports two entropy coding methods: Context adaptive variable length coding and CABAC.

High-efficiency video coding (HEVC)

HEVC is published by Sullivan *et al*. [10], which is designed to address increased video resolution and use of parallel processing architectures.

In HEVC, the various features of the VCL are:

1. The size of code tree unit is selected by the encoder
2. Each CU is partitioned into prediction units and a tree of transform units
3. Variable prediction block sizes from 64×64 down to 4×4 samples
4. Advanced motion vector prediction, improved skipped, and direct motion inference is used for motion vector signaling
5. Intrapicture prediction is performed using 33 directional modes, planar, and DC prediction modes
6. Uniform reconstruction quantization is used for quantization control
7. CABAC is used for entropy coding
8. A deblocking filter is used within the interpicture prediction loop
9. A sample adaptive offset is introduced after deblocking filter within the prediction loop to better reconstruct the original signal amplitudes.

RECENT WORK ON COMPRESSION METHODOLOGIES

The correlation between bands is very high due to improved resolution of hyper-spectral sensors. Qian *et al.* [11] presented a paper for the detection of hyper-spectral target based on unsupervised fully constrained least squares (UFCLS) linear spectral mixture analysis-based compression technique. In this technique, the data compression is achieved through the generation and coding of the interested targets in a scene. High compression ratio was achieved because the number of spectral bands is larger than the number of targets in hyper-spectral image. The spatial correlation among pixels was removed by UFCLSLU. The coding redundancy was removed by Huffman coding.

Hou *et al*. [12] published a wavelet packet coding algorithm. In this paper, the rich texture information of SAR image data is represented by the wavelet packet transform. The classification of these coefficients is performed using quad-tree technique. As a result, a high-rate distortion ratio is achieved. The coding achieved gains 0.49 dB and 0.59 in texture energy and entropy, respectively.

An efficient compression system for the hyper-spectral images was published by Christophe *et al.* [13], in which the attention is paid to optimize a compression system using wavelet for hyper-spectral images. In terms of rate-distortion criteria, an algorithm is found to get the 3D wavelet decomposition optimally and this decomposition is more useful in complexity issues. Two efficient compression methods embedded zerotree wavelet (EZW) and set partitioning in hierarchical trees (SPIHT) based on zero-tree coding are adapted. EZW gave performance similar to JPEG 2000 and also provides bitstream which is embedded fully. Good results are achieved using SPIHT; although, it is used without an arithmetic coder.

Contents can be identified easily by representing features of the image as signatures. Converting the real hashes of the image into signatures of binary type has been developed by Lv and Wang [14]. Hamming distance metrics are used for fast image identification. In image hash generation, the attacks information and the virtual image distortions information are explored. This algorithm makes use of the advantage of extended hash feature space to generate a signature for the image. The proposed method by Lv and Wang generated binary hashes of short length which are not error prone than that of conventional quantization methods. This method efficiently combined different types of hashes to give good overall performance for identification.

In 2014, Xu *et al.* [15] proposed a compressibility constrained sparse representation (CCSR) approach to image compression. This is performed using a complete dictionary of texture patches for low bit-rate compression. Convex relaxation methods are more stable for sparse representation but their coefficients are not always compressible. Here, a CCSR formulation is proposed to impose constraint on compression of the coefficients of sparse representation for image patches. The CCSR method greatly outperformed the conventional JPEG 2000 and recursive least squares dictionary learning algorithm at low bit-rates in compressing images.

Hussain *et al*. [16] published Hybrid predictive wavelet coding, which is a novel technique for image compression, in which the features of predictive coding are combined with the properties of discrete wavelet coding. The inter-pixel redundancy is removed in the pre-processing stage using predictive coding. The error values are obtained by taking the difference between the predicted values and the original values. DWT transform is applied on this error values. In the predictive coding system, a non-linear neural network predictor is utilized for prediction. At compression ratios such as 64:1 and higher, the results of this system have a significant improvement over JPEG2000. The quality of the reconstructed images using this method at a compression ratio of 1024:1 was higher than the quality of the images by JPEG2000 at a compression ratio of 64:1. Hussain *et al.* developed system got images with high quality for high decomposition levels of DWT.

A medical image compression approach was presented by Abo-Zahhad *et al.* [17]. In pre-processing, the image is passed through DPCM. Then, the application of wavelet transform is performed on the output of DPCM. These coefficients are huffman coded, resulting in threefold compression. The DPCM-DWT-huffman method performs well than the huffman, DPCM-Huffman and DWT-Huffman methods, and compression ratio provided by these methods were 6.48, 4.32, 2.27, and 1.2, respectively.

In 2015, Aulakh *et al.* [18] presented a paper, in which the image compression is done with hybrid DCT-DCT transform. This hybrid transforms utilized steganography process to compress the image. A hybrid approach, that is, DCT-DWT and steganography gave more compression in contrast to DCT and DWT and gave more storage saving and good quality of reproduced image.

Zhou *et al.* [19] published a hybrid method to compress an image by combining DPCM and vector quantization (VQ). First, the image is DCT transformed and the coefficients of DCT are zigzag scanned. The first coefficient is called DC coefficient and the rest are called AC coefficients. DPCM is performed on DC coefficient and multistage VQ is used to code the AC coefficients. Then, entropy encoding is performed. Zhou *et al.* proposed algorithm was superior in complexity, time-consuming, and compression ratio than hybrid DCT-VQ technique and conventional VQ algorithm. This scheme had a good PSNR value when compared to JPEG standard.

Shi *et al.* [20] proposed a novel scheme named content-based adaptive scanning (CAS) scheme for compression of onboard images. Initially, the wavelet transform is applied to organize the code stream according to the subbands and then, adaptive scanning is performed to preserve the texture information. At the end, the wavelet coefficients are coded by the binary tree codec. It effectively improved the coding performance and worked well at very low bit rate.

Compression of digital video results in video with different bit rate. Paek and Chang [21] presented a method to the retrieval of compressed video data from disk to digital video server effectively and also an algorithm for reservation of resources. By this method, the system is utilized efficiently.

Shimizu *et al.* [22] presented a novel multiview video coding scheme to enhance the compression efficiency to display video on different terminals. In this scheme, views are grouped into base and enhancement views. This scheme used video encoder to code the estimated view dependent geometry of the base view. Image-based rendering techniques are employed to generate prediction images of enhancement views and view-dependent geometry. The coding efficiency is increased by adopting three modifications such as interpolation, depth estimation, and quarter-pel accuracy.

A surveillance system is published by Liu *et al.* [23] (2009) with an encoder having low-complexity based on Wyner–Ziv coding. To get improvement in the coding efficiency a method is proposed called backward-channel aware Wyner–Ziv (BCAWZ) video coding, which minimizes the complexity of the encoder. And also for reliable object detection in real time an error resilience scheme is further proposed in the backward-channel. BCAWZ achieved more coding efficiency than H.264 for surveillance.

Kannangara *et al.* [24] presented a novel algorithm for describing the functionality of video decoder and structures named decoder description syntax (DDS). DDS is a syntax that defines a video decoder with the help of instructions independent of platforms. Using DDS, new coding algorithms are implemented and also multiple coding formats are implemented on one platform. A video decoder of prototype user visual design was implemented using DDS, which worked completely in real-time.

Liu *et al.* [25] proposed two novel depth compression techniques, trilateral filter, and sparse dyadic mode, which consider the structure similarity to minimize the coding artifacts. This method preserves edges while providing good compression for depth information. The depth having coding artifacts is filtered by trilateral filter based on pixel proximity, the depth samples similarity, and collocated pixels similarity in the video frame. The reconstruction of the depth map is performed by sparse dyadic mode. The synthesized video quality was improved by about 0.5-1.5 dB under same bit rate.

A novel lossless compression method is proposed by Wu *et al.* [26] in the year 2016 to achieve further compression of a group of JPEG compressed correlated images with the help of lossless compression techniques. These methods remove image redundancies such as intra- and inter-image redundancies in the feature domain, spatial domain, and frequency domains. When compared to JPEG coded image collections, this method achieved average bit savings of more than 31%.

CONCLUSION

Various image and video compression standards are discussed. For true color images and Gray-scale images, JPEG provides more lossless compression when compared to TIFF and PNG. For scanned document having lines and text, the TIFF format provides more compression and JPEG provides no compression. The TIFF and PNG provide no compression for a true color image with lots of bright colors and textures and for this type of images, JPEG2000 gives good images with low quality than JPEG. In video compression standards, H.264 supports various applications at low bit rates. HEVC is designed for the parallel processing at the same bit rate of H.264. To meet the current inflow of multimedia files on the social platform, advanced parallel techniques has to be employed, in which the computational speed is increased and also high-compression ratio is achieved while maintaining the minimum acceptable quality of Images.

REFERENCES

1. Wallace GK. The JPEG still picture compression standard. IEEE Trans Consum Electron 1992;38(1):xviii-xxxiv.
2. Christopoulos C, Skodras A, Ebrahimi T. The JPEG2000 still image coding system: An overview. IEEE Trans Consum Electron 2000; 46(4):1103-27.
3. Available from: https://www.en.wikipedia.org/wiki/TIFF.
4. Available from: https://www.en.wikipedia.org/wiki/Portable_Network_Graphics.
5. Aguilera P. Comparison of Different Image Compression Formats. Wisconsin College of Engineering, ECE 533; 2006.
6. Available from: https://www.en.wikipedia.org/wiki/H.120.
7. Available from: https://www.en.wikipedia.org/wiki/H.261.
8. Available from: https://www.en.wikipedia.org/wiki/H.263.
9. Wiegand T, Sullivan GJ, Bjontegaard G, Luthra A. Overview of the H. 264/AVC video coding standard. IEEE Trans Circuits Syst Video Technol 2003;13(7):560-76.
10. Sullivan GJ, Ohm JR, Han WJ, Wiegand T. Overview of the high efficiency video coding (HEVC) standard. IEEE Trans Circuits Syst Video Technol 2012;22(12):1649-68.
11. Qian D, Chang CI. Linear mixture analysis-based compression for hyperspectral image analysis. IEEE Trans Geosci Remote Sens 2004;42(4):875-91.
12. Hou X, Liu G, Zou Y. SAR image data compression using wavelet packet transform and universal-trellis coded quantization. IEEE Trans Geosci Remote Sens 2004;42(11):2632-41.
13. Christophe E, Mailhes C, Duhamel P. Hyperspectral image compression: Adapting SPIHT and EZW to anisotropic 3-D wavelet coding. IEEE Trans Image Process 2008;17(12):2334-46.
14. Lv X, Wang ZJ. Compressed binary image hashes based on semisupervised spectral embedding. IEEE Trans Inf Forensics Secur 2013;8(11):1838-49.
15. Xu M, Li S, Lu J, Zhu W. Compressibility constrained sparse representation with learnt dictionary for low bit-rate image compression. IEEE Trans Circuits Syst Video Technol 2014;24(10):1743-57.
16. Hussain AJ, Al-Jumeily D, Radi N, Lisboa PJ. Hybrid neural network predictive-wavelet image compression system. Neurocomputing 2015;151:975-84.
17. Abo-Zahhad M, Gharieb RR, Ahmed SM, Abd-Ellah MK. Huffman image comression incorporating DPCM and DWT. J Signal Inf Process 2015;6:123-35.
18. Aulakh NK, Kaur Y. Increasing image compression rate using (DWT+DCT) and steganography. Int J Emerg Res Manage Technol 2015;4(5)253-60.
19. Zhou X, Bai Y, Wang C. Image compression based on discrete cosine transform and multistage vector quantization. Int J Multimed Ubiquitous Eng 2015;10(6):347-56.
20. Shi C, Zhang J, Zhang Y. Content-based onboard compression for remote sensing images. Neurocomputing 2016;191:330-40.
21. Paek S, Chang SF. Video-server retrieval scheduling and resource reservation for variable bit rate scalable video. IEEE Trans Circuits Syst Video Technol 2000;10(3):460-74.
22. Shimizu S, Kitahara M, Kimata H, Kamikura K, Yashima Y. View scalable multiview video coding using 3-D warping with depth map. IEEE Trans Circuits Syst Video Technol 2007;17(11):1485-95.
23. Liu L, Li Z, Delp EJ. Efficient and low-complexity surveillance video compression using backward-channel aware Wyner-Ziv video coding. IEEE Trans Circuits Syst Video Technol 2009;19(4):453-65.
24. Kannangara CS, Philp JM, Richardson IE, Bystrom M, de Frutos Lopez M. A syntax for defining, communicating, and implementing video decoder function and structure. IEEE Trans Circuits Syst Video Technol 2010;20(9):1176-86.
25. Liu S, Lai P, Tian D, Chen CW. New depth coding techniques with utilization of corresponding video. IEEE Trans Broadcast 2011;57(2):551-61.
26. Wu H, Sun X, Yang J, Zeng W, Wu F. Lossless compression of JPEG coded photo collections. IEEE Trans Image Process 2016;25(6):2684-96.